


Gestational weight gain is associated with childhood height, weight and BMI in the Peri/Postnatal Epigenetic Twins Study

Deborah N. Ashtree^{1,2} , Deborah A. Osborne², Amelia Lee^{3,4}, Mark P. Umstad^{5,6}, Richard Saffery^{7,8}, Jeffrey M. Craig^{7,9} and Katrina J. Scurrah^{1,2}

Original Article

Cite this article: Ashtree DN, Osborne DA, Lee A, Umstad MP, Saffery R, Craig JM, and Scurrah KJ. (2022) Gestational weight gain is associated with childhood height, weight and BMI in the Peri/Postnatal Epigenetic Twins Study. *Journal of Developmental Origins of Health and Disease* **13**: 757–765. doi: [10.1017/S2040174422000113](https://doi.org/10.1017/S2040174422000113)

Received: 24 October 2021

Revised: 15 December 2021

Accepted: 24 February 2022

First published online: 24 March 2022

Keywords:

Gestational weight gain; twins; child anthropometrics

Address for correspondence:

Katrina J. Scurrah, Centre for Epidemiology and Biostatistics, School of Population and Global Health, University of Melbourne, Parkville, VIC, Australia. Email: k.scurrah@unimelb.edu.au

¹Centre for Epidemiology and Biostatistics, School of Population and Global Health, University of Melbourne, Parkville, VIC, Australia; ²Twins Research Australia, Centre for Epidemiology and Biostatistics, School of Population and Global Health, University of Melbourne, Parkville, VIC, Australia; ³Nutrition Department, Royal Women's Hospital, Parkville, VIC, Australia; ⁴School of Exercise and Nutrition Science, Faculty of Health, Deakin University, Burwood, VIC, Australia; ⁵Department of Obstetrics and Gynaecology, University of Melbourne, Parkville, VIC, Australia; ⁶Department of Maternal Fetal Medicine, Royal Women's Hospital, Parkville, VIC, Australia; ⁷Epigenetics Group, Murdoch Children's Research Institute, Royal Children's Hospital, Parkville, VIC, Australia; ⁸Department of Paediatrics, University of Melbourne, Parkville, VIC, Australia and ⁹IMPACT – the Institute for Mental and Physical Health and Clinical Translation, School of Medicine, Faculty of Health, Deakin University, Waurn Ponds, VIC, Australia

Abstract

Multifetal pregnancies are at risk of adverse maternal, neonatal and long-term health outcomes, and gestational weight gain (GWG) is a potentially modifiable risk factor for several of these. However, studies assessing the associations of GWG with long-term health in twins are rare, and studies which do assess these associations in twins often do not account for gestational age. Since longer gestations are likely to lead to larger GWG and lower risk of adverse outcomes, adjusting for gestational age is necessary to better understand the association of GWG with twin health outcomes. We aimed to explore long-term associations of GWG-for-gestational-age with twin anthropometric measures. The Peri/Postnatal Epigenetic Twins Study (PETS) is a prospective cohort study, which recruited women pregnant with twins from 2007 to 2009. Twins were followed-up at 18 months and 6 years of age. GWG-for-gestational-age *z*-scores were calculated from pre-pregnancy weight and weight at delivery. We fitted regression models to assess associations of GWG with twin weight, height and BMI at birth, 18 months, and 6 years. Of the 250 women in the PETS, 172 had GWG measured throughout pregnancy. Overall, higher GWG-for-gestational-age *z*-scores were associated with higher birthweight (β : 0.32 *z*-scores, 95% Confidence Interval (95% CI): 0.19, 0.45), BMI (β : 0.29 *z*-scores, 95% CI: 0.14, 0.43) and length (β : 0.27 *z*-scores, 95% CI: 0.09, 0.45). However, these associations were not observed at 18 months or 6 years of age. GWG was associated with twin length, weight and BMI at birth but not during childhood. Further research is needed to determine the long-term effects of GWG on twin health outcomes.

Introduction

Gestational weight gain (GWG) has been identified as a potentially modifiable factor with important implications for maternal, neonatal and long-term health.¹ In both singleton and twin pregnancies, high GWG is associated with hypertensive disorders, weight retention and obesity in the mother, and infant mortality, fetal growth and obesity in the offspring.² Given the reported relationship between GWG and birthweight in singleton gestations,³ the Institute of Medicine (IOM) developed GWG guidelines to minimise the risk of low birthweight, achieve optimal fetal growth and minimise excessive GWG.^{3,4} For twin gestations, these recommendations are: 16.8–24.5 kg for women of normal weight pre-pregnancy, 14.1–22.7 kg for overweight women pre-pregnancy and 11.3–19.1 kg for obese women pre-pregnancy.

However, these most recent IOM guidelines for GWG in twin pregnancies are only considered 'provisional'.⁴ This is because they are based on the interquartile range from one study of 706 women pregnant with twins, in which mothers delivered twins with birthweight ≥ 2500 g at ≥ 37 weeks' gestation.^{4–6} These recommendations were also developed without data on long-term health outcomes for mothers and twins, and without appropriate consideration of gestational age.

In Australia, 64% of twins are delivered prior to 37 weeks and 54% have a birthweight of less than 2500 g.⁷ As such, women who deliver twins prior to 37 weeks' gestation would be likely to have had inadequate GWG according to the current 'term' recommendations. However, attempting to gain weight within the 'term' recommendations would likely lead to excessive

© The Author(s), 2022. Published by Cambridge University Press in association with International Society for Developmental Origins of Health and Disease. This is an Open Access article, distributed under the terms of the Creative Commons Attribution licence (<https://creativecommons.org/licenses/by/4.0/>), which permits unrestricted re-use, distribution, and reproduction in any medium, provided the original work is properly cited.

GWG for gestational age, and could potentially lead to adverse health outcomes. Therefore, GWG recommendations for multi-fetal pregnancies should consider gestational age, and studies assessing the link between GWG and maternal and twin health should make appropriate adjustments for gestational age.

Many previous studies of GWG and twin health have focused on birth and neonatal outcomes, particularly size at birth. However, birth size has frequently been linked to long-term health outcomes, especially anthropometric and cardiometabolic health. Given that birth size should only be used as a proxy for events occurring in utero, and that GWG is one factor leading to size at birth, GWG may have long-term links to health outcomes, including catch-up growth. Understanding these long-term implications of GWG may be of particular importance for twin pregnancies, as women pregnant with twins gain more weight than women with singleton pregnancies, and are also at higher risk of adverse pregnancy outcomes.⁴ Despite this, studies exploring the association of GWG with long-term twin health outcomes are lacking, and a recent systematic review (originally published in 2014⁸ and updated in 2021⁹) found most studies assessing GWG and pregnancy outcomes in twin gestations suffer from methodological flaws, including failure to: adjust for confounders such as smoking and socio-economic position; adjust for twin chorionicity; include long-term health outcomes; determine whether there was enough statistical power to detect associations and/or had a small sample size; and adjust for gestational age, potentially leading to reverse causation.⁸

Therefore, we aimed to address these limitations by assessing the association of GWG-for-gestational age *z*-scores (to account for gestational age) with twin height (length at birth), weight and BMI at birth, 18 months and 6 years of age.

Method

Cohort selection

The Peri/Postnatal Epigenetic Twins Study (PETS) is a Melbourne-based prospective twin cohort study, established in 2007. Details of the methods of recruitment, data collection and ethics approval have been reported previously.¹⁰ Briefly, women pregnant with twins were recruited during their second trimester between 2007 and 2009 from three pregnancy clinics in Melbourne, Australia. Ethics approval was obtained from each pregnancy clinic – the Royal Women's Hospital, Mercy Hospital for Women and Monash Medical Centre. Women were excluded if they planned to leave the area before delivery or if they had limited English language skills. Of the 287 women recruited during pregnancy, 250 mothers and their twins remained in the study at birth, of which 172 twin-pairs had full exposure and outcome measures recorded.

Exposure

Maternal pre-pregnancy weight was self-reported through questionnaires. Gestational weight was recorded during study visits at ~12, ~24 and ~36 weeks' gestation. For women recruited after 12 weeks' gestation and who missed the 12-week study visit, weight at 12 weeks was collected retrospectively via maternal recall (*n* = 7). GWG was calculated by subtracting a woman's pre-pregnancy weight from her last measured weight before delivery, usually at 36 weeks' gestation. Pre-pregnancy BMI was calculated as weight (kg) divided by height (m²), then categorised as underweight, normal weight, overweight or obese according to the

World Health Organization guidelines.¹¹ These categories were used to calculate pre-pregnancy BMI- and gestational age-standardised *z*-scores, based on GWG growth charts for twin pregnancies.⁵ These *z*-scores account for all gestational ages, and do not assume linearity. Growth charts for women underweight prior to pregnancy are currently unavailable, so *z*-scores could not be calculated for these women, and this group were excluded from all analyses.

Outcome

Neonatal anthropometric measures, including birthweight, were measured twice: once at birth by clinical and delivery staff; and again by PETS staff, usually at delivery, but at most, within 72 h. Measurements recorded by PETS staff were used in these analyses. Twenty-six twins could not be measured by PETS staff, due to requirements for special or intensive care. For these twins, birth measurements were recorded by medical personnel at the time of birth or accessed via birth records. In cases where clinical staff measurements and PETS staff measurements differed substantially (likely due to delays in PETS measurements), the PETS team discussed and determined which measurement to use, on a case-by-case basis. Small-size-for-gestational-age (SGA) was assessed as birthweight below the 10th percentile, based on Australian birthweight percentile reference charts for twin pregnancies.¹²

Follow-up of the twins occurred at ages 18 months and 6 years of age, when mothers completed a questionnaire on the health, development and nutritional history of the twins. The twins also had anthropometric measurements taken by a trained research assistant. Weight was recorded using digital weight scales, and height was recorded using a stadiometer. BMI was calculated as weight (kg) divided by height (m²). Height, weight and BMI *z*-scores, accounting for twin sex and age, were calculated using Australian twin birthweight reference charts,¹² and UK anthropometric charts for 18 months and 6 years of age, using the Zanthro package in Stata 15.

Childhood growth was calculated as the change in *z*-score between two time periods, including birth to 18 months, 18 months to 6 years and birth to 6 years.

Other variables

Confounders were identified *a priori* based on knowledge of the subject area and through reviewing recent evidence.⁸ All models were adjusted for maternal age at delivery, pre-pregnancy BMI (continuous), smoking during pregnancy (any vs none), socio-economic status (SES) at birth, gestational age, twin sex and chorionicity.

Socio-economic indexes for areas (SEIFA) scores at the time of birth were used as an indicator of women's SES. SEIFA scores are created using information from the Australian census about people and households in a defined small geographical area (postal area in this analysis, containing approximately 8000 people per postcode in Australia in 2006), to measure relative socio-economic advantage and disadvantage, economic resources, education and occupation.¹³ Higher SEIFA scores indicate a higher SES and lower disadvantage. We used the Index of Relative Socio-Economic Disadvantage in this analysis.

Statistical analysis

Means and standard deviations were calculated for continuous demographic variables, and percentages were calculated for categorical variables.

Table 1. Descriptive statistics as mean (standard deviation) or frequency [percent] for mothers and twin children in the PETS

	Missing GWG	Observed GWG	P-value for difference
Mothers (n = 250)	n = 78	n = 172	
Age at delivery (years)	31.92 (5.32)	33.04 (4.95)	0.107
Smoked (at all) during pregnancy	10 [12.8%]	26 [15.1%]	0.499
Alcohol (at all) during pregnancy	18 [23.1%]	60 [34.9%]	0.009
Gestational age (weeks)	33.82 (2.71)	36.78 (1.31)	<0.001
Pre-pregnancy BMI (kg/m ²)	24.96 (5.61)	24.77 (5.43)	0.800
Children (n = 500)	n = 156	n = 344	
Preterm birth ^a	136 [87.2%]	102 [29.7%]	<0.001
Delivery type (caesarean)	122 [78.2%]	203 [59.4%]	<0.001
Zygosity (monozygotic)	88 [56.4%]	120 [34.9%]	<0.001
Chorionicity (monochorionic)	64 [41.0%]	74 [21.5%]	<0.001
Sex (male)	71 [45.5%]	173 [50.3%]	0.322
Birthweight (g)	2079.47 (595.56)	2645.28 (403.95)	<0.001
SGA	15 [9.6%]	19 [5.5%]	0.096
BMI (kg/m ²) at age six ^b	14.69 (2.10)	16.57 (1.77)	0.290

BMI, body mass index; GWG, gestational weight gain; SGA, small for gestational age.

^aDefined as delivery prior to 37 weeks' gestation.

^bResults presented are median (interquartile range) as this variable was non-normal and right-skewed.

Associations of GWG-for-gestational-age *z*-scores with neonatal and childhood health outcomes were assessed by fitting linear regression models using generalised estimating equations to account for the correlation between twins in each pair. Our main model included GWG for gestational age *z*-score as a continuous exposure (Model 1). Results assessing associations of GWG per kilogram increase, with and without adjusting for gestational age, have been included in Supplementary Table S1 to allow comparison of our results with previous studies, and to assess the impact of not adjusting for gestational age (Model 2).

Simes-adjusted *q*-values were estimated for all adjusted regression models to account for multiple testing.¹⁴ Normality of residuals and linearity of the relationship were assessed, and variables were transformed where necessary. Linearity of GWG *z*-score with twin anthropometric outcomes was assessed visually using scatter plots with predicted linear and lowess lines, and confirmed by using quadratic terms in linear regression models. The influence of residuals was assessed by plotting the distribution using a histogram, and removing the highly influential values in a sensitivity analysis, to assess whether these values altered the results.

In addition to performing complete-case analyses, we implemented inverse probability weighting (IPW) for women with observed or missing GWG in the PETS, to assess the impact of bias due to missingness. Methods for IPW have been described in detail elsewhere.¹⁵⁻¹⁷ Briefly, we compared the demographic characteristics of women who had GWG measures with those of women missing GWG data, then used IPW to weight the observed GWG according to the probability of a woman having GWG measurements. We used stabilised weights, as this better accounts for the large weights due to participants with a low probability of having complete data.¹⁵ The stabilised weights, *sw*, were calculated as:¹⁷

$$sw = \frac{1}{Pr} \text{ when } r = 1 \text{ (no missing data)}$$

$$sw = \frac{1 - Pr}{1 - P} \text{ when } r = 0 \text{ (missing data)}$$

where *Pr* is the probability of having complete data, when considering other covariates, and *P* is the probability of having complete data without considering other covariates. We then conducted a sensitivity analysis using this weighted GWG variable. Detailed IPW methods applied to the PETS data have been provided in Supplementary Material.

All analyses were performed using Stata 15 (StataCorp. 2017. *Stata Statistical Software: Release 15*. College Station, TX: StataCorp LLC.).

Results

Table 1 compares the demographic and clinical characteristics of the 172 mothers (344 twin children) included in this study with the 78 mothers with missing GWG. Details of attrition and loss to follow-up are shown in Fig. 1. There were 152 twin-pairs included in the adjusted regression analyses.

For women and twins included in this study, mean gestational age (36.8 weeks) was slightly higher than the mean for Australian twins (35.2 weeks),¹⁸ and mean pre-pregnancy BMI was 24.8 kg/m². Mean birthweight for the twins (2.6 kg) was slightly higher than the mean for Australian twins (2.4 kg),¹⁸ and 6% of twins were considered SGA according to twin guidelines.¹²

Of the 78 women missing total GWG, 68 women delivered twins preterm. Women with spontaneous preterm labour (*n* = 26) did not have their weight measured prior to delivery, and so 'total GWG' was not available for these women (Fig. 2). Of the remaining

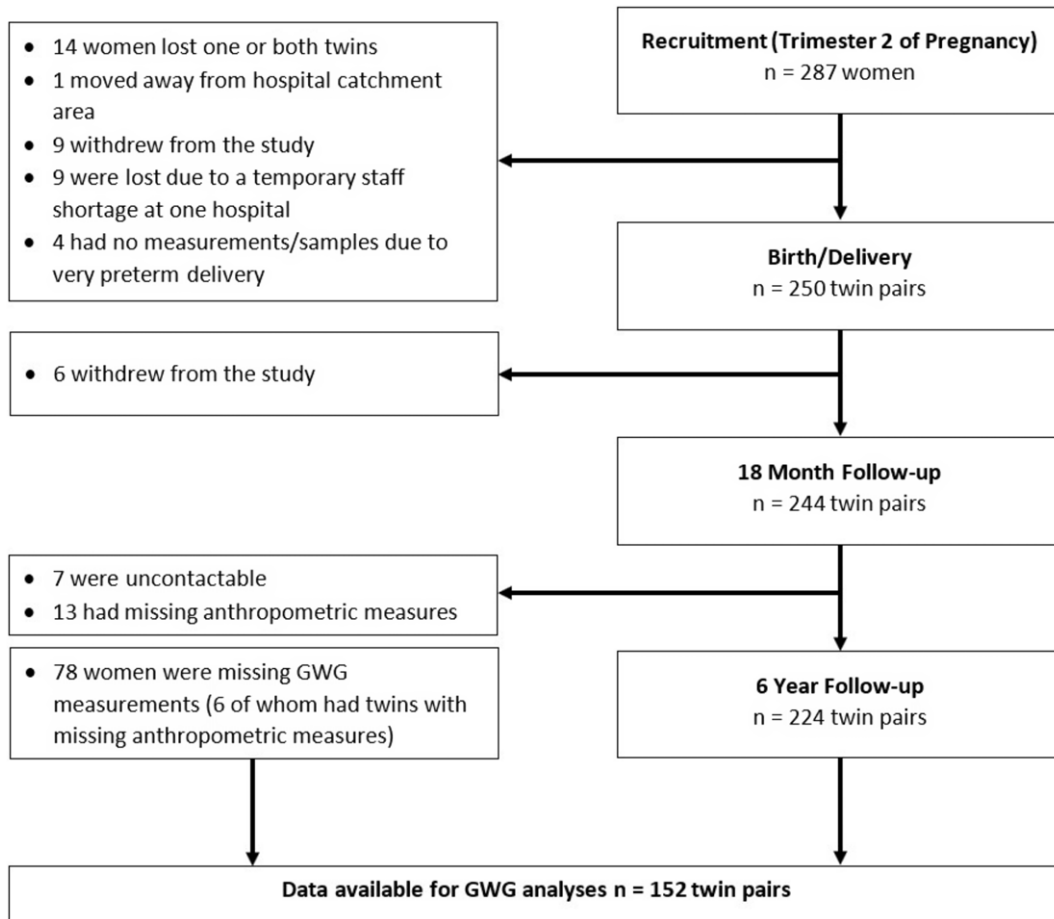


Figure 1. Flow chart showing loss to follow-up between recruitment and the 6-year follow-up in the PETS.

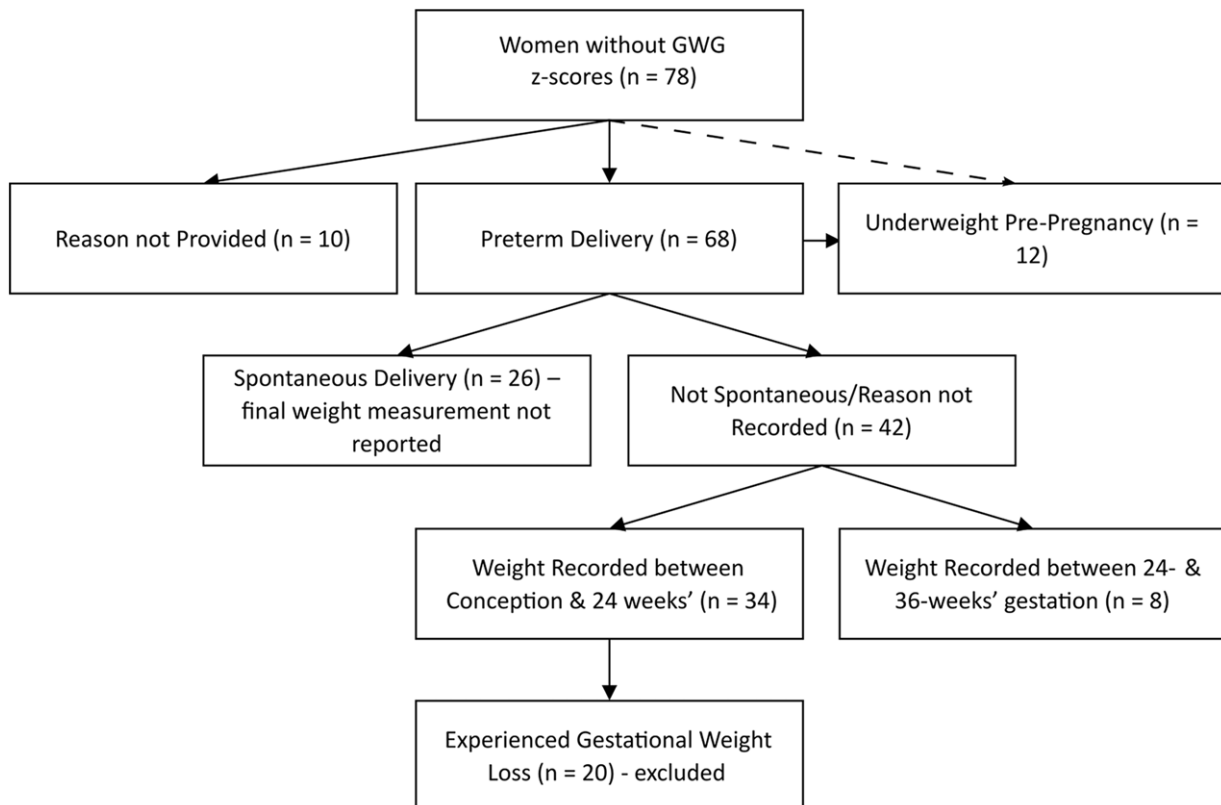


Figure 2. Flow chart showing the reasons for women missing GWG measurements.

Table 2. Results from the unadjusted and adjusted linear regression models assessing the associations of GWG-for-gestational-age z-scores with birth outcomes and early-life anthropometrics in the children

	Unadjusted β -coefficient (95% CI), <i>p</i> -value	Adjusted ^a β -coefficient (95% CI), <i>q</i> -value ^b
Birth outcomes		
Birthweight (g; n = 304)	92.77 (25.57, 159.97), 0.007	117.30 (71.93, 162.68), 0.001
Birthweight (z-score; n = 304)	0.32 (0.16, 0.48), <0.001	0.32 (0.19, 0.45), 0.001
BMI (kg/m ² ; n = 279)	0.19 (0.01, 0.36), 0.033	0.29 (0.14, 0.45), 0.001
BMI (z-score; n = 279)	0.18 (0.03, 0.34), 0.023	0.29 (0.14, 0.43), 0.001
Length ^c (cm; n = 279)	0.44 (0.05, 0.82), 0.027	0.51 (0.18, 0.85), 0.003
Length (z-score; n = 279)	0.23 (0.07, 0.38), 0.004	0.27 (0.09, 0.45), 0.003
18-month outcomes		
Height ^c (cm; n = 268)	0.44 (0.05, 0.82), 0.027	0.46 (−0.05, 0.97), 0.076
Height ^c z-score (n = 268)	0.07 (−0.08, 0.21), 0.373	0.04 (−0.17, 0.25), 0.741
Weight ^d (kg; n = 267)		
Linear term	0.06 (−0.15, 0.27), 0.563	0.17 (−0.03, 0.37), 0.094
Quadratic term	−0.17 (−0.28, −0.06), 0.003	−0.11 (−0.21, −0.01), 0.035
Weight ^d (z-score; n = 267)		
Linear term	0.02 (−0.16, 0.20), 0.827	0.07 (−0.11, 0.25), 0.457
Quadratic term	−0.13 (−0.23, −0.03), 0.012	−0.13 (−0.23, −0.03), 0.010
BMI ^d (kg/m ² ; n = 267)		
Linear term	−0.08 (−0.28, 0.13), 0.459	0.02 (−0.21, 0.24), 0.887
Quadratic term	−0.25 (−0.37, −0.13), 0.001	−0.22 (−0.34, −0.09), 0.001
BMI ^d (z-score; n = 267)		
Linear term	−0.03 (−0.18, 0.13), 0.744	0.03 (−0.14, 0.20), 0.735
Quadratic term	−0.18 (−0.28, −0.07), 0.001	−0.17 (−0.27, −0.07), 0.001
6-year outcomes		
Height ^{c,e} (cm; n = 206)	0.58 (−0.55, 1.75), 0.319	0.69 (−1.04, 2.42), 0.434
Height ^c (z-score; n = 194)	−0.005 (−0.18, 0.17), 0.958	0.01 (−0.23, 0.25), 0.923
BMI ^f (kg/m ² ; n = 206)	0.07 (−0.19, 0.34), 0.605	0.001 (−0.09, 0.09), 0.992
BMI (z-score; n = 194)	0.05 (−0.12, 0.23), 0.558	0.01 (−0.14, 0.15), 0.975
Weight ^g (kg; n = 206)	0.35 (−0.23, 0.96), 0.241	0.08 (−0.11, 0.29), 0.679
Weight (z-score; n = 194)	0.02 (−0.14, 0.19), 0.771	−0.06 (−0.16, 0.14), 0.975

BMI, body mass index; CI, confidence interval; GWG, gestational weight gain.

^aAdjusted for maternal age at delivery, maternal pre-pregnancy BMI, maternal smoking status during pregnancy, socio-economic status, twin sex, chorionicity, and gestational age.

^bAdjustment for multiple testing, by using the Simes method with *p*-values from all adjusted models presented in this paper.

^cAlso adjusted for maternal height.

^dQuadratic regression was a better fit for the data, so the model also included a squared term.

^eInverse cube transformation, presented on original scale for ease of interpretation.

^fInverse square transformation, presented on original scale for ease of interpretation.

^gInverse transformation, presented on original scale for ease of interpretation.

women who delivered twins preterm, only eight had GWG recorded between 24 weeks' and 36 weeks' gestation. Since including these values did not alter the mean and distribution of GWG z-scores, or the estimated regression coefficients (data not shown), we included these eight GWG measures in the final models. The remaining women who delivered preterm had at least one GWG measurement recorded between conception and 24 weeks' gestation. However, most of these (n = 20) women experienced gestational weight loss between conception and 24 weeks, and so their calculated GWG z-scores would likely be unrepresentative of their GWG at delivery. As such, we excluded this group of women from further analyses. Twelve women who delivered

preterm were underweight prior to pregnancy, but had total GWG in kilograms available. However, since growth charts for women in this BMI category are unavailable, we were unable to calculate GWG z-scores for these women.

For every standard deviation increase in GWG-for-gestational-age z-score, birthweight increased by 117.30 g on average (95% CI: 71.93, 162.68), and increased by 0.32 standard deviations in z-score (95% CI: 0.19, 0.45, Table 2). Larger GWG-for-gestational-age z-scores were associated with higher weight at 18-months, but we observed no associations of GWG-for-gestational-age z-score with 6-year anthropometric measures. In contrast, when using GWG in kilograms (without adjusting for gestational age), the

Table 3. Results from the unadjusted and adjusted linear regressions assessing the associations of GWG-for-gestational-age z-scores with early-life growth^a

	Unadjusted β -coefficient (95% CI), p -value	Adjusted ^b β -coefficient (95% CI), q -value ^c
Growth between birth and 18-months		
Height (n = 255)	-0.17 (-0.36, 0.01), 0.067	-0.15 (-0.33, 0.03), 0.260
Weight (n = 267)	-0.21 (-0.43, 0.02), 0.072	-0.16 (-0.38, 0.06), 0.342
BMI (n = 248)	-0.04 (-0.26, 0.18), 0.735	-0.06 (-0.29, 0.17), 0.867
Growth between 18-months and 6-years		
Height (n = 184)	-0.20 (-0.32, -0.08), 0.002	-0.20 (-0.31, -0.09), 0.001
Weight (n = 184)	-0.22 (-0.39, -0.04), 0.015	-0.26 (-0.44, -0.09), 0.022
BMI (n = 184)	-0.12 (-0.35, 0.12), 0.323	-0.18 (-0.42, 0.05), 0.305
Growth between birth and 6-years		
Height (n = 178)	-0.26 (-0.44, -0.08), 0.006	-0.23 (-0.42, -0.03), 0.079
Weight (n = 194)	-0.28 (-0.51, -0.06), 0.015	-0.30 (-0.51, -0.09), 0.030
BMI (n = 178)	-0.10 (-0.32, 0.13), 0.401	-0.19 (-0.42, 0.03), 0.252

BMI, body mass index; CI, confidence interval; GWG, gestational weight gain.

^aCalculated as $z - \text{score}_{\text{time2}} - z - \text{score}_{\text{time1}}$.

^bAdjusted for maternal age at delivery, maternal pre-pregnancy BMI, maternal smoking status during pregnancy, socio-economic status, twin sex, chorionicity, and gestational age.

^cAdjustment for multiple testing, by using the Simes method with p -values from all adjusted models presented in this paper.

associations of GWG with height and weight were present at 18-months and 6-years of age (Supplementary Table S1).

Higher GWG-for-gestational-age z-score was associated with lower growth rate (change in z-score) between 18-months and 6-years, and between birth and 6-years (Table 3).

Results of regression models which assessed associations of GWG with birth, neonatal and childhood anthropometric outcomes were robust to influential observations; results not shown. Further sensitivity analyses were performed to assess the impact of using GWG observations from women who self-reported weight at 12 weeks, using BMI as a continuous versus categorical variable, and with versus without using gestational age as a covariate in the GWG z-score models. Inferences for each of these sensitivity analyses were consistent with the original model; data not shown. Although we found differences in the demographic characteristics of women with versus without GWG measurements, inferences were consistent between the unweighted and IPW results (Supplementary Tables S5 and S6). This indicates that missing GWG was unlikely to influence the associations of GWG with twin anthropometric outcomes at birth, 18-months and 6-years of age, or postnatal growth of the twins.

Discussion

Overall, results from this paper indicated that GWG z-scores in twin pregnancy are associated with height, weight and BMI at birth in twins, but these associations had weakened by 6 years of age. However, GWG z-scores were associated with childhood growth until age six. Although limited by sample size, we found evidence that GWG above and below the current recommendations may be associated with higher weight and BMI in childhood, indicating that weight gain within an 'appropriate' range may be beneficial for long-term health outcomes. However, evaluating the appropriateness of current GWG recommendations was outside the scope of this paper.

Previous studies assessing the association of GWG in twin pregnancy with neonatal and childhood health outcomes have often not adjusted for smoking, socio-economic position or twin

chorionicity; have not included long-term health outcomes; or have not appropriately adjusted for gestational age. In this paper, we have addressed these limitations and found that when adjusting for gestational age using z-scores in the PETS, higher GWG were associated with higher anthropometric measures at birth. Few studies of GWG in twin pregnancies have assessed association with long-term health outcomes in the twins, but we found that the linear association of GWG with anthropometric measures weakened at 18-months and had disappeared entirely by 6 years. This indicates that the impact of GWG on offspring height, weight and BMI may be short-term. However, the lack of observed associations at 6 years may be due to low power to detect associations. We also found evidence that maternal GWG is associated with twin growth until 6 years of age, suggesting a potential longer-term association. Since the PETS has only followed the children to 6-years of age, though data collection at an 11-year follow-up visit is underway, further exploration of the association of GWG with twin growth is warranted. Though we found evidence for an association of GWG with birth outcomes, for example, an increase of one GWG z-score (~6 kg) results in an increase of 117 g, the magnitude of these associations may not be clinically relevant. Categorising outcomes to indicate whether an association was 'adverse' or 'beneficial' was not possible in this study, as this would have led to further loss of information and/or power. Such adverse outcomes could include being born small-for-gestational-age, or childhood obesity. Given the limitations of our study to determine whether GWG was associated with adverse or beneficial outcomes, the results from this study alone may only have limited clinical implications.

In contrast, previous studies of GWG have found strong links between 'optimal' GWG and improved birth and pregnancy outcomes, such as appropriate-for-gestational-age birth size and lower risk of neonatal death. However, these studies often do not appropriately account for gestational age.^{19,20} There is a strong correlation between gestational length and total GWG: women with longer pregnancies have more time to gain weight.⁸ Therefore, studies which explore associations of GWG with twin outcomes using total GWG are unable to separate the effects of

preterm delivery from low GWG. Previous studies have attempted to disentangle gestational duration from GWG by using rate of GWG, limiting analyses to term pregnancies, or by using gestational age as a covariate in regression models.

The first of these solutions assumes that GWG is linear throughout pregnancy.⁸ However, women tend to gain less weight in the first trimester than in the second trimester,⁸ and fetal growth in twin pregnancy slows from around 32 weeks until term,¹² so this assumption is unlikely to be met. Studies which limit analyses to term pregnancies are likely to have an unrepresentative sample, given that more than half of twin pregnancies are delivered preterm.⁷ Finally, studies which simply include gestational age as a covariate do not consider the collinearity between GWG and gestational age, so are likely to have biased results. In fact, when using GWG in kilograms as an exposure without adjusting for gestational age, we could conclude that GWG was associated with twin weight and height at 18 months and 6 years of age in the PETS. Yet, when we appropriately adjust for gestational age by using GWG *z*-scores, we do not observe this association. Although using raw GWG and *z*-scores addresses different research questions, the differences in conclusions between the models indicates the importance of accounting for gestational age for future studies assessing associations of GWG with twin health outcomes, though the best method depends on the specific research question being addressed. Since recent GWG growth charts have been created for twin gestations, these should be used to create GWG-for-gestational-age *z*-scores.⁵

A recent study using these GWG *z*-scores found that higher GWG was associated with post-partum weight retention and childhood obesity at age five.²¹ The study concluded that the current GWG recommendations may be inappropriate and may be contributing to adverse long-term health outcomes for mothers and twins.²¹ However, this study only focussed on the BMI or educational outcomes of twins, so further work is needed to understand the links between GWG and other long-term outcomes. Our study expands on this paper by examining other anthropometric measures and at different ages throughout childhood. This previous paper also examined the association of GWG *z*-scores with BMI for the upper and lower limits of the current IOM recommendations, and concluded that even gaining weight within the recommendations led to adverse outcomes. Initially, we had intended to perform similar additional analyses to determine whether associations of GWG *z*-scores with twin anthropometric outcomes differed according to pre-pregnancy BMI or IOM category (below, within or above current recommendations). However, due to sample size limitations, these analyses were underpowered, so we have only included a graphical representation of these results (Supplementary Figs. S1 and S2) as an indicator of potential associations. Since we were unable to comprehensively evaluate whether associations differ for women who gain within or outside the recommendations, or whether the results differ according to pre-pregnancy BMI, we recommend caution when interpreting results. However, these results indicate that GWG *z*-scores may be associated with twin weight and BMI at 18 months or 6 years for women who gain above or below the current IOM recommendations but not for women who gain within the recommendations (Supplementary Fig. S1). We also find that these associations may differ according to pre-pregnancy BMI (Supplementary Fig. S2). Though these results indicate differences in twin outcomes for women with different

pre-pregnancy BMI or who gain weight within or outside IOM categories, more research with larger sample sizes are needed to confirm these observations.

Limitations

Though the current study found that GWG was associated with adverse outcomes for the twins, the results must be considered within the limitations of the study. Firstly, the PETS was not originally established to evaluate GWG in twin pregnancies. As such, the sample size may not be large enough to evaluate the associations between GWG and child health outcomes with adequate power. Some of the results may have been false-positives arising from multiple testing, so to assess whether multiple testing influenced our results, we implemented the Simes method of *p*-value adjustment.¹⁴ The original *p*-values and adjusted *q*-values resulted in similar inferences for most models. We recognise the limitations of frequentist/hypothesis testing approaches, described in detail elsewhere,²² and instead recommend focusing on the effect size, rather than *p*-values for significance, and comparing our results with other similar studies.²³

SES was determined by an area-level estimate, so may not be representative of individual SES. Although the PETS recruited from different hospitals across Melbourne, women with limited English and women planning to leave the hospital catchment area before delivery were excluded. As such, this may have introduced an English language or a sampling bias, so may not be representative of twin pregnancies across Australia. Maternal pre-pregnancy weight was assessed through maternal recall, potentially leading to recall bias. We found differences in the demographic characteristics of women who had GWG measures compared to those with missing GWG, which may also have contributed to selection bias. Women without GWG measures typically gave birth to twins with a lower birthweight and were more likely to have a caesarean delivery than women with GWG recorded, suggesting these women may be more likely to have a complicated pregnancy. However, when using IPW to determine what impact these missing data had on our results, we found consistent regression estimates between the unweighted and weighted models, indicating that missing data may not be an issue in this study. However, we recognise that it is impossible to know whether all potential confounders were measured or included in our analyses, which may have led to incorrect model specification.

We adjusted for several confounders in our analyses, however, it may be important to account for additional factors. Postnatal factors, such as breastfeeding, childhood nutrition and physical activity may all be important for twin anthropometric outcomes in later life. However, including these additional variables would have reduced our sample size further, and potentially led to over-adjustment in our models. Larger twin studies of GWG should account for additional postnatal factors.

We recognise that pre-pregnancy BMI may play an important role in the association of GWG with twin health outcomes; however, we were unable to comprehensively evaluate whether outcomes or associations differ for twins according to maternal pre-pregnancy BMI. We also recognise that the association of GWG with twin health outcomes may differ according to sex (of individual twins and of twin-pairs). Though we included an adjustment for sex, due to sample size restrictions, we were unable to comprehensively assess whether associations differ for male and female twin-pairs. However, in a supplementary subgroup

analysis, we found evidence that the association of GWG *z*-scores with 18-month weight and BMI may differ for male and female twins (Supplementary Table S8).

Finally, GWG as a rate per week may be important for maternal and offspring outcomes, and the IOM do provide recommendations for GWG per week. However, these recommendations assume linear GWG, and since fetal growth in twin pregnancy slows from around 32 weeks until term,¹² this was a somewhat unrealistic assumption we were unwilling to make in this study. Additionally, assessing the timing of GWG, for example, comparing early-mid to mid-late GWG, may also provide additional insights into the links between maternal GWG and child anthropometric and cardiometabolic health. However, due to sample limitations, we chose to focus on total GWG in our main analyses. We have provided an additional analysis in supplementary material, and although results should be interpreted with caution, we found evidence that timing of GWG may have a differential effect on twin health outcomes and may have long-term associations with twin anthropometric outcomes. Furthermore, in the PETS, GWG was collected at three time points only (usually at 12, 24 and 36 weeks' gestation), and so further exploration of timing of GWG, such as per week, was not possible from our study. Given that the IOM recommendations are currently only available to women who deliver twins at or after 37 weeks' gestation, comparing our results to the IOM recommendations is also limited.

Strengths

Unlike the study informing the current guidelines, this study did not restrict the sample to women delivering twins with a birth-weight of ≥ 2500 g at ≥ 37 weeks. We also addressed the confounding effect (and potential for reverse causation) of gestational age, by calculating GWG-for-gestational-age *z*-scores. Finally, our study addresses some of the limitations of previous studies,⁸ by adjusting for smoking status, SES and chorionicity of the twins, and by considering long-term health outcomes of the twin children. To our knowledge, this is the first Australian study to assess the association of GWG *z*-scores in twin pregnancy with childhood anthropometric measures.

Conclusions

Higher GWG *z*-scores were associated with higher anthropometric measures at birth. This study has added to the current body of knowledge by examining GWG in an Australian twin cohort, appropriately adjusting for gestational age, and considering the long-term health of the twin children.

Supplementary materials. For supplementary material for this article, please visit <https://doi.org/10.1017/S2040174422000113>

Acknowledgements. We wish to thank all the twins, mothers and families involved in the PETS. We also wish to thank the entire PETS research team, in particular, Dr Cong Sun, Dr Pamela Leong, Dr Jane Loke and Dr Mihiri Silva for their help in accessing and cleaning the required data.

Financial support. This work was supported by the Australian Government Research Training Programme Scholarship (D.N.A); Twins Research Australia, which receives support from the National Health and Medical Research Council through a Centre of Research Excellence Grant, administered by the University of Melbourne (K.J.S); the Australian National Health and Medical Research Council (J.M.C, grant numbers 437015 and 607358); the Bonnie Babes Foundation (J.M.C., grant number BBF20704); the Financial

Markets Foundation for Children (J.M.C, grant number 032-2007); and the Victorian Government's Operational Infrastructure Support Programme.

Ethical standards. This work contains observational data from a cohort of women pregnant with twins. The authors assert that all procedures contributing to this work comply with the ethical standards of the Australian Code for the Responsible Conduct of Research (2018), and that the research was approved by the following institutional ethics boards: the Royal Women's Hospital ethics committee (project # 06/26), Mercy Hospital for Women (project # R06/30) and then Monash Medical Centre (project # *06117C).

Conflicts of interest. None.

References

- Lustiv O, Hulman A, Woolcott C, *et al.* Examining the provisional guidelines for weight gain in twin pregnancies: a retrospective cohort study. *BMC Pregnancy Childbirth.* 2017; 17(1), 330.
- Herring SJ, Oken E. Weight gain during pregnancy: importance for maternal and child health. *Ann Nestlé Engl.* 2010; 68(1), 17–28.
- Goodnight W, Newman R. Optimal nutrition for improved twin pregnancy outcome. *Obstet Gynecol.* 2009; 114(5), 1121–1134.
- Institute of Medicine (US) and National Research Council (US) Committee to Re-examine IOM Pregnancy Weight Guidelines. Composition and components of gestational weight gain: physiology and metabolism. In *Weight Gain During Pregnancy: Re-examining the Guidelines* (eds. Rasmussen KM, Yaktine AL), 2009; pp. 71–110. National Academies Press.
- Hutcheon JA, Platt RW, Abrams B, *et al.* Pregnancy weight gain by gestational age in women with uncomplicated dichorionic twin pregnancies. *Paediatr Perinat Epidemiol.* 2018; 32(2), 172–180.
- Luke B, Hediger ML, Nugent C, *et al.* Body mass index-specific weight gains associated with optimal birth weights in twin pregnancies. *J Reprod Med.* 2003; 48, 217–224.
- Australian Institute of Health and Welfare. Australia's mothers and babies 2015 - in brief. Canberra: AIHW; 2017. <https://www.aihw.gov.au/reports/mothers-babies/australias-mothers-babies-2015-in-brief/summary>.
- Bodnar LM, Pugh SJ, Abrams B, Himes KP, Hutcheon JA. Gestational weight gain in twin pregnancies and maternal and child health: a systematic review. *J Perinatol.* 2014; 34(4), 252–263.
- Whitaker KM, Ryan R, Becker C, Healy H. Gestational weight gain in twin pregnancies and maternal and child health: an updated systematic review. *J Womens Health (Larchmt).* 2022; 31(3), 362–381.
- Saffery R, Morley R, Carlin JB, *et al.* Cohort profile: the Peri/Post-Natal Epigenetic Twins Study. *Int J Epidemiol.* 2012; 41(1), 55–61.
- World Health Organization. *Obesity: Preventing and Managing the Global Epidemic, Report of a WHO Consultation on Obesity*, 2000. World Health Organization.
- Li Z, Umstad MP, Hilder L, Xu F, Sullivan EA. Australian national birth-weight percentiles by sex and gestational age for twins, 2001–2010. *BMC Pediatr.* 2015; 15(1), 148.
- Pink B. *Information Paper: An Introduction to Socio-Economic Indexes for Areas (SEIFA)*, 2006. Australian Bureau of Statistics.
- Simes RJ. An improved Bonferroni procedure for multiple tests of significance. *Biometrika.* 1986; 73(3), 751–754.
- Austin PC, Stewart E. Moving towards best practice when using inverse probability of treatment weighting (IPTW) using the propensity score to estimate causal treatment effects in observational studies. *Stat Med.* 2015; 34(28), 3661–3679.
- Seaman SR, White I. Review of inverse probability weighting for dealing with missing data. *Stat Med.* 2011; 22(3), 278–295.
- Xu S, Ross C, Raebel MA, *et al.* Use of stabilized inverse propensity scores as weights to directly estimate relative risk and its confidence intervals. *Value Health.* 2010; 13(2), 273–277.
- Laws PJ, Hilder L. *Australia's Mothers and Babies 2006, 2008*. Australian Institute of Health and Welfare.

19. Fox NS, Rebarber A, Roman AS, *et al.* Weight gain in twin pregnancies and adverse outcomes: examining the 2009 Institute of Medicine guidelines. *Obstet Gynecol.* 2010; 116(1), 100–106.
20. González-Quintero VH, Kathiresan ASQ, Tudela FJ, *et al.* The association of gestational weight gain per Institute of Medicine Guidelines and pre-pregnancy body mass index on outcomes of twin pregnancies. *Am J Perinatol.* 2012; 29(6), 435–440.
21. Bodnar LM, Cartus AR, Parisi SM, *et al.* Pregnancy weight gain in twin gestations and maternal and child health outcomes at 5 years. *Int J Obes.* 2021; 45(7), 1382–1391.
22. Ranganathan P, Pramesh C, Buyse M. Common pitfalls in statistical analysis: the perils of multiple testing. *Clin Res.* 2016; 7(2), 106–107.
23. Sjölander A, Vansteelandt S. Frequentist versus Bayesian approaches to multiple testing. *Eur J Epidemiol.* 2019; 73, 809–821.