THE EFFECT OF GESTATIONAL WEIGHT GAIN ON NEWBORN BODY COMPOSITION

Dr. Corah Ohadike

Candidate Number: 888073

St Catherine's College

University of Oxford

MSc (Res) Obstetrics and Gynaecology

Table of Contents

Word count: Approximately 23, 000.

Abstract

The Effect of Gestational Weight Gain on Newborn Body Composition

Corah Ohadike, St. Catherine's College, University Of Oxford
Thesis submitted for MSc (by research) in Obstetrics and Gynaecology
Trinity Term 2017

This thesis assesses how gestational weight gain (GWG) influences newborn body composition (NBC) in an ethnically diverse population and determines the modifiers on that relationship.

To answer the research question, it was important to understand what may be classified as optimal GWG. To achieve this, I carried out a systematic review of studies that produced GWG charts; this showed there was considerable heterogeneity in the quality of available GWG charts.

GWG was transformed into z-scores; this allowed analyses, which were independent of gestation length. Newborn body composition is considered a more accurate marker of neonatal size in anthropometric research when compared with birthweight. It is usually divided into body fat percentage (BF%) and fat free mass (FFM). NBC was measured using a PEAPOD® device which calculates BF% and FFM in infants up to the age of 6 months. It is considered the gold standard of measuring NBC.

My results showed that as GWG increased, both BF% and FFM increased. For each z-score unit increase in GWG, BF% increased by 1.8% and FFM by 292.7grams. The strongest modifiers on this relationship were BMI group, gender and GA at delivery. In addition, endocrine disorders and age at PEAPOD measurement affected only FFM and smoking affected only BF%.

The factors driving future morbidities are, of course, complex but GWG represents a modifiable factor that is easily measured and monitored hence its appeal. Public health issues such as childhood obesity, adulthood obesity, stunting and wasting and medical issues such as cardiovascular disease and diabetes appear to be affected by fetal environment and NBC. If it is at all possible to alter these, by influencing GWG through women's behaviour and clinical care we provide, the benefits in both the developed and developing world in the form of preventative medicine would be invaluable.

Acknowledgements

I dedicate this thesis to my father, Dr. Patrick Ohadike R.I.P. I hope you are looking down with pride. Thank you for always showing us the value of the pursuit of knowledge and setting me on the right path to achieve anything I wished to.

Special thanks to my amazing mother, Caroline, and my eldest sister, Pamela, *(the best deputy mum in the world)*, for always listening, inspiring, advising and comforting me.
Special thanks also to my sister, Yvonne, *(epidemiologist extraordinaire)* for enduring my long transatlantic phone calls during which she translated statistical 'speak' so I could remain sane and focussed.

Also Thanks to:

My brothers, Kanwulia and Menkiti, who have always encouraged their little sister. Thanks for cheering me on.

My closest friends, Hlupe, Chetan and Sharifah, who never fail to encourage me and have tolerated me being a hermit for last few months.

My supervisors: Jane Hirst and Leila Cheikh-Ismail, and advisors: Fabien Puglia and Eric Ohuma, for your invaluable input, guidance and help. You have made it possible for me to hopefully pursue a career as an clinical academic.

Professor Kennedy, who convinced me to apply for the clinical research fellow job via a Skype call, over 4 years ago, when I was at a crossroads. Thank you for always helping and advising me throughout my time at Oxford.

And finally to my wonderful husband Greg, my best friend: Thank you for your unwavering faith in me. Thank you for being the reason I smile and laugh everyday. I am so blessed to have you in my life.

List of figures and tables

Abbreviations

- 4-C: 4-compartment
- ADP: Air displacement plethysmography
- ARD: Absolute risk difference
- BF%: body fat percentage
- BMI: Body mass index
- CI: Confidence interval
- DEXA: Dual X-ray absorptiometry
- FFM: Fat free mass
- FM: Fat mass
- g: gram
- GA: Gestational age
- GDM: Gestational diabetes mellitus
- GTT: Glucose tolerance test
- GWG: Gestational weight gain
- INTERGROWTH 21st –The international fetal and newborn growth consortium for the 21st century
- IAEA: International Atomic Energy Agency
- IOM: Institute of Medicine
- Kg: kilogram
- LGA: Large for gestational age
- NBC: Newborn body composition
- NGT: Normal glucose tolerance
- OR: Odds ratio
- PPW: Pre-pregnancy weight
- PRISMA: Preferred Reporting Items for Systematic Reviews and Meta-Analysis
- QS: Quality score
- SD: Standard deviation
- SGA: Small for gestational age
- TOBEC: Total body electricity
- USS: Ultrasound scan
- WHO: World Health Organisation

Cover Letter

To the examiners,

We confirm that the re-use of text in the systematic review (chapter 2) of this thesis has adhered to the conditions set out by the Medical Sciences Division Graduates Studies Department as follows:

(1) The text has not been used in another thesis for obtaining another degree.

(2) It has been acknowledged clearly in the thesis what parts of the text had been previously been published, and where.

(3) We confirm that Dr. Corah Ohadike wrote the relevant text.

Yours Sincerely,

Dr Jane Elizabeth Hirst

Dr Leila Cheikh-Ismail

CHAPTER 1: INTRODUCTION

1. Introduction

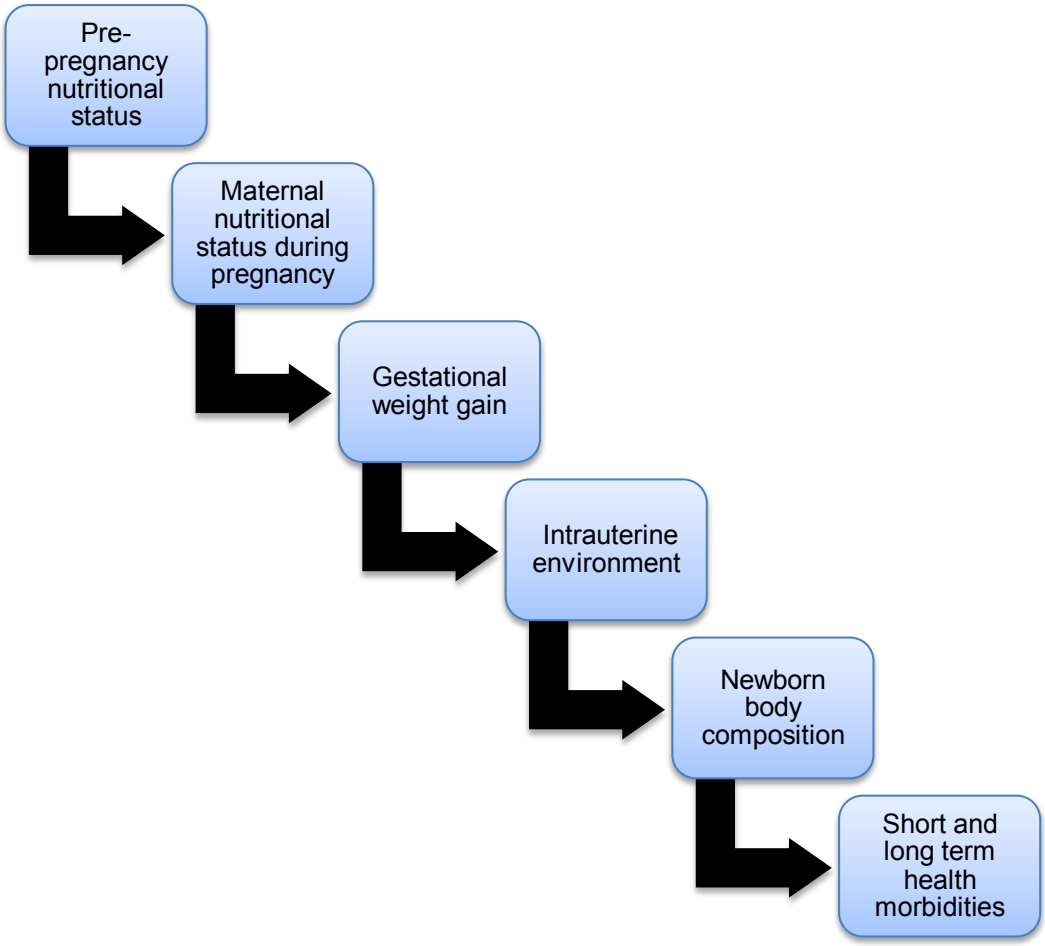
1.1 Rationale for investigating newborn body composition and gestational weight gain

There is a growing body of evidence that recognises the importance of early life exposures on childhood and adulthood disease (Barker, Osmond, Forsen, Kajantie, & Eriksson, 2005; Oken & Gillman, 2003; Osmond & Barker, 2000; Sayer et al., 2004). This is relevant because pregnancy may represent a crucial time to intervene and prevent significant morbidities in later life. Maternal body composition, weight gain and dietary intake have been shown to directly affect fetal size (Barker et al., 2005; Russo et al., 2013; Widen et al., 2015). The long-term issues, which have been related to small or large neonatal size, include diabetes, obesity and coronary heart disease (Barker et al., 2005; Oken & Gillman, 2003; Osmond & Barker, 2000; Poston, 2012).

Evidence supports an association between GWG and neonatal fat mass (Au, Raynes-Greenow, Turner, Carberry, & Jeffery, 2013; Crozier et al., 2010; Hull et al., 2011; Poston, 2012). This in turn has been linked to childhood obesity (Oken, Taveras, Kleinman, Rich-Edwards, & Gillman, 2007; Widen et al., 2015; Zilko, Rehkopf, & Abrams, 2010) with markers of cardiovascular risk detectable in childhood (Fraser et al., 2011; Oken et al., 2007).

Figure 1.1 summarises the proposed theories on the effect of GWG on future health.

Figure 1.1: Proposed role of GWG on future morbidity from synopsis of literature



Childhood and adolescent obesity is rapidly becoming a public health issue of great concern. Rates are rising worldwide, putting children at increased risk of obesity in adulthood (De Onis & Lobstein, 2010). In the UK, 19.8% of 10-11 year olds and 9.3% of 4-5 year olds were been found to be obese in 2015-2016 (Public Health England, 2016). Health Survey England has also provided separate data in 2015 showing that 28.2% of children aged 2-15 are overweight or obese (Boodnha, 2014). Whilst factors driving this epidemic are complex, it is likely that maternal nutrition and the intrauterine environment may represent modifiable early life influences (Black et al., 2013; Poston, 2012).

In order to determine the effects of GWG on future health risks, it is important to establish optimal GWG. The most commonly used GWG targets are those published by the Institute of Medicine (IOM). The IOM assessed the available evidence on GWG to produce guidelines in 1970, with subsequent updates in 1990 and 2009 (Medicine, Institute of, 2009). Whilst the IOM acknowledged serious limitations in available evidence, (Agency for Healthcare Research and Quality (AHRQ), 2008), these guidelines were released for clinical use and gave a target GWG ranges for women during pregnancy based on maternal pre-pregnancy BMI. These guidelines have been widely adopted in clinical practice, despite an acknowledgement that they were designed for women living in the USA and were based on a lack of randomised evidence demonstrating any improvements in outcomes.

In addition it is important to note that these guidelines were references as opposed to standards. A reference is defined as a tool for grouping and analysing locally acquired data; it is descriptive. A standard embraces the concept of a healthy or aspirational target and involves a value judgement; it is prescriptive. Whilst reference data have been widely, used to make inferences about the health or nutrition of populations and individuals, the World Health Organisation (WHO) have endorsed the use of standards for human anthropometry (World health Organisation, 1995; Villar, Altman, et al., 2013). This would mean weight gain is compared to a physiological ideal, enabling valid comparison between populations and direct comparison of the efficacy of nutritional interventions in pregnancy.[1]

[1] Taken from Ohadike et al. Adv Nutr 2016;7:313–22

In 2016, the International Fetal and Newborn Growth Consortium for the 21st Century (INTERGROWTH-21st) published the first international standards for GWG for women in the normal BMI range (18.5-24.99 kg/m^2) (Cheikh Ismail et al., 2016). INTERGROWTH-21st was a large population-based project that investigated growth and development from under 14 weeks' gestation, through pregnancy and in infants up to 2 years of age. One unique feature of this project was that it was conducted in eight geographically distinct populations around the world. Within these populations, women were selected to participate in the project who were at a low risk of adverse perinatal outcomes (Villar, Altman, et al., 2013). The INTERGROWTH 21st GWG standards provide equations for calculating mean and SD GWG for each gestational week, which allows calculation of z-scores in relation to an optimal standard (Cheikh Ismail et al., 2016).

Z-scores provide a way of assessing GWG independent of gestational length, allowing valid comparisons between women at any gestation. Z scores are commonly used in paediatric anthropometric data presentation, improving accuracy and interpretation of data independent of age or gender(Kirkwood & Sterne, 2003).

Traditionally, the assessment of neonatal size has been by measurement of birthweight. Birthweight however is a crude indicator, and alone cannot give any indication of the quality of the in utero growth, with babies that are normally grown and growth restricted having the same weight. Quality of growth is thought to be represented more accurately by quantifying body

composition, comprising fat mass (FM) and fat free mass (FFM) (Au et al., 2013; Catalano, Thomas, Huston-Presley, & Amini, 2003; IAEA, 2013; Poston, 2012). Neonates who may be born with similar measurements for length, weight, BMI or ponderal index can differ markedly in terms of FM and FFM (IAEA, 2013). FFM has a characteristic metabolic function with a rate of formation that increases predictably throughout gestation. FM, on the other hand, is variable between babies being influenced by nutrition and the in utero environment (Catalano et al., 2003; Catalano, Drago, & Amini, 1995). Neonatal FM within the first week of birth is not thought to be affected by postnatal influences and therefore may provide valuable insight into the effects of the intrauterine environment on childhood and adulthood disease (Starling et al., 2015).

GWG can be influenced by lifestyle interventions during pregnancy making it an appealing target to decrease short and long term adverse outcomes (Catalano et al., 1995; Russo et al., 2013; Widen et al., 2015). An in-depth understanding of the association between GWG, using the INTERGROWTH-21st standard, and FFM/FM as a marker of NBC could provide novel insights. If the association is confirmed, this information could help decide whether modifying GWG could be a possible intervention to improve NBC with potential benefits for the child in the future. The following sections of this thesis will explore these concepts in more detail, setting the scene for the original analyses presented in chapters 3-6 of this thesis. It is first important however to establish what is known about GWG, and what factors affect NBC and the effect of NBC on childhood outcomes.

Chapter 2 presents a systematic review of the methodological quality of studies aimed at creating charts for GWG. This review adapted standard methodology used by the INTERGROWTH-21st Consortium to assess the risk of bias of published studies of growth charts. This review was published in Advances in Clinical Nutrition: ***Adv Nutr 2016;7:313–22; doi:10.3945/an.115.010413*** (Ohadike et al., 2016)***.*** I led the writing, developed and conducted the searches of electronic databases, performed the assessment of risk of bias and synthesis of the results in duplication with my supervisor (Dr J Hirst). As the published review only included articles until April 2014, for this thesis I have repeated the search up to September 2017, and include a critique of the 2 additional studies published after April 2014. One of the new studies provided the GWG standards that were used to create the z-scores utilised in the regression analyses. Short sections of the literature review in chapter 1 concerning GWG are also taken from this publication and are indicated with footnotes.

1.2 Gestational weight gain

1.2.1 Definition

Gestational weight gain is defined as the total weight gained during pregnancy. Ideally this should be the last measured weight prior to delivery minus measured weight at a pre-conception visit (Institute of Medicine 2009). Monitoring maternal GWG has been part of antenatal care since the early 20th century, despite lack of agreement on the optimal range of GWG (Agency for Healthcare Research, 2008; Crane, White, Murphy, Burrage, & Hutchens, 2009; Poston, 2012; Rasmussen et al., 2010; Scholl, Hediger, & Schall, 1995)[2].

1.2.2 Measuring gestational weight gain

GWG includes not only maternal weight but also the fetus, placenta, amniotic fluid, plasma volume, uterine and breast tissue (Agency for Healthcare Research and Quality (AHRQ), 2008; IAEA, 2013). Despite GWG being a non-specific marker, research comparing the use of GWG and other measures such as BMI, maternal height or mid-upper arm circumference has concluded that GWG is a suitable measure of maternal anthropometry. It is also a consistent marker in terms of its predictive value for perinatal outcomes (World Health Organisation, 1995; Russo et al., 2013; Widen et al., 2015). Using GWG as part of the monitoring of women throughout pregnancy has other advantages; it is a readily available, cheap and easy to perform making it suitable for use around the world.

[2] Taken from Ohadike et al. Adv Nutr 2016;7:313–22

However, there are several challenges in GWG research. Obtaining an objective measure of pre-pregnancy weight is technically and ethically challenging. Thus it is generally accepted that this information should be obtained either by maternal recall or by measuring first trimester maternal weight, meaning that various studies may not be using the same baseline. This is also a problem when assessing final weight in pregnancy, with some studies also using maternal recall for this measure (Agency for Healthcare Research and Quality (AHRQ), 2008; Medicine, Institute of, 2009). This over reliance on self-reported weight represents a considerable methodological limitation. Another issue that arises is the way GWG is defined in terms of patterns of gain across pregnancy, with either gain during each trimester or gain in the first and second halves of pregnancy presented (Medicine, Institute of, 2009).

GWG recommendations are available in various formats: longitudinal charts (Calvo et al., 2009; Guelinckx, Beckers, Vansant, & Devlieger, 2010), single weight gain targets for the entire pregnancy (Luke, Hediger, & Scholl, 1996; Scholl et al., 1995), targets based on body mass index (BMI) (Institute of Medicine, 2009), and proportional increases based on antenatal 'weight-for-height' (Mardonnes &Rosso, 1997; Rosso, 1985). This makes comparisons and the process of evidence synthesis of methodological quality challenging. In addition, due to the varying definitions and standardisation of data collected, it makes comparisons between studies difficult.[3]

[3] Taken from Ohadike et al. Adv Nutr 2016;7:313–22

Longitudinal presentation of GWG in the form of charts, could offer advantages. Firstly, it would enable women to track their own weight change visually and instigate health interventions as appropriate. Secondly, it would provide clinicians with a more precise tool for monitor GWG taking into account length of gestation and the rate of weight gain throughout pregnancy. However there has been little attention to studies creating GWG charts or their methodology.[4] A consistent approach to the assessment and presentation of GWG would aid women and clinicians in interpreting this data and allow valid comparisons between studies.

[4] Taken from Ohadike et al. Adv Nutr 2016;7:313–22

1.2.3 Gestational weight gain standards

In 2009, the Institute of Medicine (IOM) published the most comprehensive synthesis of GWG recommendations to date (Institute of Medicine, 2009). These guidelines were specific for BMI categories and independent of age, parity, smoking and ethnicity. Women, who are underweight (BMI < 18.5 kg/m^2), are advised to gain between 12.7 and 18.1 kg; normal weight (BMI between 18.5 to 24.9 kg/m^2) 11.3-15.9 kg; overweight (BMI of 25-29.9 kg/m^2) 6.8 - 11.3 kg and obese (BMI > 30 kg/m^2) 5.0 - 9.1 kg (Institute of Medicine, 2009).

In 2016, the INTERGROWTH-21st Project published the first international GWG standards for women with a normal BMI. This project is a large-scale, well-designed, multi-ethnic and international study. Women were weighed every 5±1 weeks during pregnancy using a strict anthropometric protocol (Cheikh Ismail et al., 2016). The results showed that women from ethnically and geographically diverse populations who were free from nutritional, social and medical constraints on pregnancy demonstrated remarkably similar ranges of GWG (Cheikh Ismail et al., 2016). Along with the overall GWG standards, GWG targets for specific gestation week windows ($14\text{-}18^{+6}$, $19\text{-}23^{+6}$, $24\text{-}28^{+6}$; $29\text{-}33^{+6}$, and $34\text{-}40^{+0}$ weeks) were also published. The total GWG standard at 40 weeks was 13.7±4.5 kg.

1.2.4 Association of excessive and insufficient gestational weight gain and perinatal outcomes

Associations between insufficient GWG and perinatal outcomes

Insufficient GWG is associated with an increased risk of delivering a small-for-gestational-age baby (SGA), delivery by caesarean, preterm birth and lower success of breastfeeding (Agency for Healthcare Research and Quality (AHRQ), 2008; Institute of Medicine, 2009; Nohr et al., 2009; Zilko, Rehkopf, & Abrams, 2010). A recent large systematic review and meta-analysis by Goldstein et al. (2017) evaluated the association between insufficient and excessive GWG and perinatal outcomes. 5354 studies were identified of which 23, with 1,309,136 participants, met the inclusion criteria. They found that compared with women gaining weight appropriately, GWG below the IOM recommendations was associated with higher risk of SGA (OR, 1.53 [95%CI, 1.44-1.64], absolute risk differences (ARD), 5% [95%CI, 4%-6%]) and preterm birth (OR, 1.70 [1.32-2.20]; ARD, 5%[3%-8%]), a lower risk of large-for-gestational-age (LGA) (OR, 0.59 [0.55-0.64]; ARD, −2% [−10% to −6%]) and macrosomia (OR, 0.60 [0.52-0.68]; ARD, −2%[−3%to −1%]). The authors did not find a difference in caesarean delivery (OR, 0.98 [0.96-1.02]; ARD, 0%[−2%to 1%]) (Goldstein et al., 2017).

A search of the literature revealed there are few published studies on the effects of insufficient GWG on NBC. Most work about insufficient GWG is in the form of observational studies, which utilise birthweight alone as a marker for neonatal size. However a recent study by Catalano et al. (2014) shows that low GWG between 0-5kg was associated with significantly lower

birthweight, FM and FFM compared to women who gained over 5kg: birthweight (3258±443 vs. 3467±492 g, p<0.0001), FM (403±175 vs. 471±193 g, p<0.0001), and FFM (2855±321 vs. 2995±347g, p<0.0001) (Catalano et al., 2014).

Associations between excessive GWG and perinatal outcomes

Excessive GWG has been associated with a range of conditions including gestational diabetes (Agency for Healthcare Research, 2008; Edwards, Hellerstedt, Alton, Story, & Himes, 1996; Kabiru & Denise Raynor, 2004; Institute of Medicine, 2009; Saldana, Siega-Riz, Adair, & Suchindran, 2006), pregnancy-induced hypertension, pre-eclampsia (Agency for Healthcare Research, 2008; Bodnar, Catov, Klebanoff, Ness, & Roberts, 2007; Crane et al., 2009; Poston, 2012; Rasmussen et al., 2010; Scholl et al., 1995; Sibai et al., 1997; Thadhani et al., 1999), macrosomia (Agency for Healthcare Research, 2008; Edwards et al., 1996; Kabiru & Denise Raynor, 2004; Institute of Medicine, 2009; Saldana et al., 2006), increased risk of Caesarean delivery (Joseph, 2003; Rosenberg, Garbers, Lipkind, & Chiasson, 2005; Sherrard et al., 2007; Witter, Caulfield, & Stoltzfus, 1995) and postpartum weight retention(Luke et al., 1996; Scholl et al., 1995).[5]

In the systematic review by Goldstein et al. (2017), the authors found that gestational weight gain above the IOM recommendations was associated with lower risk of SGA (OR, 0.66 [0.63-0.69]; ARD, −3%; [−4%to −2%]) and preterm birth (OR, 0.77 [0.69-0.86]; ARD, −2%[−2%to −1%]) and higher risk of LGA (OR, 1.85 [1.76-1.95]; ARD, 4%[2%-5%]), macrosomia (OR, 1.95 [1.79-

[5] Taken from Ohadike et al. Adv Nutr 2016;7:313–22

2.11]; ARD, 6%[4%-9%]), and caesarean delivery (OR, 1.30 [1.25-1.35]; ARD, 4%[3%-6%]) (Goldstein et al., 2017).

1.2.5 Association of excessive and insufficient gestational weight gain and longer term outcomes

Effects of insufficient GWG on long term outcomes

Few studies have robustly examined the association of insufficient GWG with long-term outcomes for the mother or child. Zilko et al. (2010) found that insufficient GWG reduced the chance of postpartum weight retention (OR, 0.65; 95% CI, 0.48–0.89) when compared with women who had adequate GWG (Zilko, Rehkopf, & Abrams, 2010), however longer-term child outcomes were not examined.

Oken et al. (2007) demonstrated GWG is associated with significant differences in child BMI at three years of age. In this study of 1044 women and children, adequacy of GWG was determined according to the IOM 1990 guidelines. When compared with adequate or excessive GWG (0.47 z-score units [95% CI, 0.37, 0.57] and 0.52 z-score units [95% CI, 0.44, 0.61]), women with inadequate GWG had children with lower BMI z-scores at age 3 (0.17 z-score units [95% CI, 0.01, 0.33]) (Oken et al., 2007)This was one of the first studies to recognize the possibility that adiposity of the child is adversely affected by insufficient GWG. It is hypothesised that reduced adiposity at birth may mimic the findings of wasting (Victora et al., 2015) and this would potentially increase the risk of metabolic and cognitive dysfunction in later life (Dewey & Begum, 2011; Black et al., 2013).

Effects of excessive GWG on long term outcomes

Evidence is accumulating that higher GWG puts children at risk of obesity, with the potential for adult cardiovascular disease and obesity (Poston, 2012). In order to try to identify causal pathways between GWG and childhood obesity and to distinguish these from the influence of environment and postnatal diet, the association of NBC with GWG may be important.

Crozier et al. (2010) demonstrated that excessive GWG is associated with increased FM at birth. The authors followed 948 children who had body composition measurements at birth, 4 years and 6 years. They found that offspring of women with excessive GWG had higher FM at birth, 4 years and 6 years compared to offspring from mothers with adequate GWG (increase in FM by 7% (95% CI: 1, 14), 4% (95% CI: 0, 9), and 10% (95% CI: 4, 17) (Crozier et al., 2010).

These observations of increased childhood adiposity are supported in other studies on the effects of maternal GWG on offspring longer-term outcomes (Diesel et al., 2015; Widen et al., 2015; Zilko, Rehkopf, & Abrams, 2010). Diesel et al. (2015) found that after adjusting for covariates such as pre-pregnancy BMI and smoking status, a change GWG z-scores from – 1.0 SD to 1.5 SD resulted in a significant increase in the risk of child obesity at 16 years. In addition, compared with a GWG z-score of 0 SD a GWG z-score of + 1.5 SD was associated with a 1.81-fold increase in the risk of child obesity at age 16 (Diesel et al., 2015). It is worth noting, that the sample of women in this study were all from low income households, however the authors state

they adjusted for household income and paternal education in their analyses.

Widen et al. (2015) also showed that increasing GWG was associated with increased BMI in childhood. It was demonstrated that a five kg increase in total GWG was associated with a 0.11 higher child BMI z-score. Excessive GWG was associated with an almost 300% increased risk of childhood obesity and a higher BMI z-score {0.44 z-score units [(95% CI): 0.2, 0.7], p=0.001} BF% at 7 years [β co-efficient: 2.2% (95% CI: 1.0, 3.5), p=0.001)] and obesity [OR: 2.93 (95% CI: 1.5, 5.8), p=0.002] (Widen et al., 2015). The study population was limited to Dominican and African-American children whose mothers had high rates of obesity; this may explain why there was a much higher risk of childhood obesity (300%) compared with other studies. These findings therefore need to be interpreted with care, as they are likely to lack external validity.

There is also observational evidence of much longer-term effects on obesity. In one study that followed over 2400 children from birth to age 42 in Denmark, GWG was observed to be associated with offspring obesity in adulthood. When comparing high GWG (>16kg) to low GWG (<6kg) there was a 2.36 (1.08-5.15) fold risk of obesity and 1.28 (0.89-1.85) fold risk of overweight in adulthood (Schack-Nielsen, Michaelsen, Gamborg, Mortensen, & Rensen, 2009).

Excessive GWG has also been associated with increased risk of raised blood pressure in childhood (Fraser et al., 2011; Oken et al., 2007). Oken et al. (2007) found that systolic blood pressure was an estimated 0.60 mmHg (95%

CI: 0.06, 1.13) higher per 5 kg of gestational weight gain. This association reduced to 0.34 mmHg (95% CI: 0.19, 0.87) after additional adjustment for child BMI (Oken et al., 2007). Fraser et al. (2011) found the mean systolic BP in children born to women gaining excessive weight compared to women with recommended weight gain was 1.25 mmHg higher (95%CI :0.604, 1.896). Whilst these increases in systolic BP are not likely to cause clinically significant changes in childhood, they could be markers of children that will become adults at increased risk of cardiovascular disease later in life (Fraser et al., 2011).

The afore mentioned study by Zilko et al. (2010) also found that women with excessive GWG were found to have significantly decreased odds of a SGA infant (OR, 0.56; 95% CI, 0.40–0.77) and increased odds of a LGA infant (OR, 2.15; 95% CI, 1.57–2.95) and child overweight (OR, 1.27; 95% CI, 1.10 – 1.48). The population in this study was more diverse so the results may have greater external validity than other studies where the population was limited to certain ethnic groups (Widen) or had less diverse populations (Oken et al., 2007; Schack-Nielsen et al., 2009).

GWG has been linked to various childhood and adulthood outcomes. However, what remains to be conclusively determined is how much of a role NBC, environmental factors and epigenetics play. Well-designed long-term studies are required in order to determine the pathways of causation and make any definitive conclusions.

1.3 Newborn body composition

1.3.1 Definition

There are five levels of body composition: Atomic (level I), Molecular (level II), cellular (level III), tissue system (level IV) and whole body (level (V). For the purposes of nutrition and anthropometry the molecular level is used; it is also known as the nutrition level. This divides the body into the following components: water, lipid, protein and other (IAEA, 2013). Models to describe NBC are divided into two main categories: 2-compartment and multi-compartment (IAEA, 2013; Müller, Braun, Pourhassan, Geisler, & Bosy-Westphal, 2016; Withers et al., 1998). The 2-compartment model divides the body into FM and FFM. FFM comprises total body water, bone minerals and protein whereas FM comprises chemical fat with energy stores (IAEA, 2013; Müller et al., 2016).

The principles underlying the different compartment models are explained as follows. In a 2-compartment model, FM is calculated using densitometry techniques and FFM then obtained by subtraction of FM from birthweight. A 3-compartment model elaborates upon the 2-compartment model, by measuring total body water and FM and determining FFM by subtraction. The 4-compartment model measures total body water, FM, and bone mineral density and residual mass is then calculated by subtraction (Withers et al., 1998). Multi-compartment models are of limited use in clinical and research settings because of the need for specialised equipment, long time to perform assessments and requirements of high technical expertise (Eriksson, Löf, & Forsum, 2010; IAEA, 2013; Müller et al., 2016). In addition, although the

assessment of body composition using multi-compartment models has been found to be more accurate in some studies (Withers et al., 1998), for the purpose of nutritional assessment and practicality, the use of 2-compartment models have been found to be acceptable and reliable (IAEA, 2013; Müller et al., 2016).

1.3.2 Measuring newborn body composition

Birth weight has been and continues to be the most commonly used marker to assess intrauterine growth. It is simple and readily available. As mentioned earlier, birthweight does not differentiate body composition (Catalano et al., 2003; IAEA, 2013). High birthweight may be due to any number of reasons: high fluid retention, fat content or increased lean body mass. Length measurements have also been used but require specific techniques that can be difficult to achieve in neonates (Bettiol, 2003). In addition, other anthropometric proxy measurements such as skinfold thickness, head circumference and mid-arm circumference are prone to errors with analysis(B. Eriksson et al., 2010; IAEA, 2013).

Numerous methods of measuring body composition exist: total body potassium, magnetic resonance imaging, acoustic plethysmography, underwater weighing, computed tomography, total body electricity (TOBEC), dual X-ray absorptiometry (DEXA) and air-displacement plethysmography (ADP) (Au et al., 2013; B. Eriksson et al., 2010; IAEA, 2013; Ma et al., 2004; Müller et al., 2016; Yao, Nommsen-Rivers, Dewey, & Urlando, 2003). The majority of these are risky, expensive, or unreliable for use in newborns. In addition, commonly used methods of measuring body composition such as stable isotope dilution and dual x-ray absorptiometry pose multiple practical challenges (Hawkes et al., 2011). As detailed below (Table 1.1), ADP is comparable to other highly accurate methods such as DEXA.

Two-compartment body composition is most accurately determined by densitometric methods (Urlando, Dempster, & Aitkens, 2003). Densitometry can be assessed by hydrodensitometry or air displacement plethysmography (ADP). Hydrodensitometry techniques are performed in water whereas ADP is performed in air (Urlando, Dempster, & Aitkens, 2003). Hydrodensitometry is unsuitable for use in the infant population due to subjects needing to be completely submerged in water (IAEA, 2013; Yao et al., 2003). ADP is now viewed as the gold standard method for measuring NBC (Au et al., 2013).

1.3.3 Air displacement plethysmography and PEAPOD

The PEAPOD® Infant Body Composition Tracking system (COSMED USA Inc.) launched in 2004, is the first commercially available device to assess ADP in infants between birth and 6 months age. PEAPOD is more amenable for use in large studies and provides reliable measurements of BF% (Hawkes et al., 2011; Hull et al., 2011). It is simple to use, acceptable to parents and well tolerated by infants (Au et al., 2013; Yao et al., 2003). Detailed information about the function of the PEAPOD® machine and the measurement procedures can be found in Appendix 1 and Figures (i) to (iii)

PEAPOD measurement validity

ADP measurements using PEAPOD have been validated against more traditional methods of assessing body composition measurements including the 4-compartment (4-C) model (Ellis et al., 2007) and deuterium dilution for body water technique (Ma et al., 2004). Table 1.1 shows comparisons between the assessment procedures. Mean estimations of BF% were similar when comparing the 4-C method (16.3 ± 7.2%) and ADP (16.9±6.5%) (Ellis et al., 2007). Mean estimations of BF% were also similar when comparing the deuterium method (20.3±6.9%) and ADP (20.4±6.7%) (Ma et al., 2004). The differences between the techniques were not statistically significant.

Table 1.1: Comparison of PEAPOD against 4-C and deuterium body composition assessments (Ellis et al., 2007; Ma et al., 2004)

	4-C Method	Deuterium	PEAPOD
Measurements	DEXA scan Deuterium Heel prick test x2	Heel prick test x2 Deuterium	Weight and BF% by machine
Calculations/Technical requirements	Complex calculations	Spectrometer required Complex calculations	Automated by machine
Time	Blood samples- 3 hours Whole body counting-15 minutes DEXA- 5 minutes Possibly +1 hour for neonate to sleep in order to do DEXA scan	Blood samples- 2 hours Centrifuge of samples and transfer- ± 3 hours	3 minutes
Acceptability/Ease of use	Parents refused heel prick test so changed to urine test Neonate had to be asleep to avoid movement	Invasive	No concern from parents Immediate output Not affected by neonate movement

1.3.4 Newborn body composition estimates

The mean neonatal BF% at term (≥37 weeks gestation) described in the literature (Au et al., 2013; Lee et al., 2009; Roggero et al., 2009; Starling et al., 2015) ranges from 8.6±3.7% (Roggero et al., 2009) to 10.6±4.6% (Lee et al., 2009). All of these studies used ADP to measure NBC.

Comparison of FFM between studies is difficult as some report FFM% rather than FFM in grams. The values are generally similar; 2947±342g in Au et al. (2013) and 2851±325g in Starling et al. (2015) Part of the variation in the results may be because measurements were taken at different points in the postnatal period (ranging form 0-10 days). As mentioned earlier, in the first week following delivery, FM is thought to remain constant but vary thereafter. FFM on other hand seems to decrease in the initial postnatal period (Institute of Medicine, 2009; Starling et al., 2015; Villar et al., 2017). Therefore, the commonly observed reduction in birthweight after delivery is likely to be predominantly related to decrease in FFM (Institute of Medicine, 2009; Starling et al., 2015; Villar et al., 2017).

There are few large studies that have measured NBC by ADP with the aim to create standards as opposed to using other indirect methods such as skinfold thickness. The numbers of newborns in these studies ranged between 17 and 110 (Eriksson et al., 2010; Lee et al., 2009; Ma et al., 2004; Roggero et al., 2009; Yao et al., 2003). Where large studies have been performed, the aim was to determine predictors or associated factors with NBC (Au et al., 2013; Starling et al., 2015). A recent large study (1019 newborns), as part of the INTERGROWTH 21st project, sought to directly measure and describe NBC

after birth (Villar et al., 2017). The mean BF% in all newborns at term was 10.2 ±4.0% and FFM 2850±421g. In the lowest risk newborns, BF% was similar 10.3±3.9% but FFM was higher 2968±336g. It was noted that females had significantly higher BF% (10.7±4.0% vs. 9.6±4.0%) and males had higher FFM (2965±422g vs. 2739±390g). These gender differences are consistent with previous studies (Au et al., 2013; Catalano et al., 1995; Simon et al., 2013). The mothers in this study were at low risk of adverse outcomes therefore the results of this study could be viewed as NBC standards. However as the population was limited to the UK, it may not be appropriate to assume these standards can be applied to international populations.

1.3.5 Association between increased and reduced newborn body fat on newborn and long-term outcomes

Reduced body fat

Reduced BF% is indicative of intrauterine wasting (Victora et al., 2015). Victora et al. (2015) assessed the factors associated with stunting and wasting as defined by the INTERGROWTH 21st standards in 51,200 babies enrolled in the INTERGROWTH 21st project. Wasting was defined as 'low weight for length or low BMI for age representing recent weight loss' (Victora et al., 2015). Compared with stunting, wasting was associated with higher risk of neonatal complications such as admission to neonatal intensive care OR: 6.7 (5.5-8.1); respiratory distress syndrome, 4.0 (3.3-4.9); transient tachypnoea, 2.1 (1.5-2.9)]; and no oral feeding for >24 hours, 5.0 (3.9-6.5). In addition, severe wasting was associated with a higher risk of mortality and poorer outcome with all assessed morbidities (Victora et al., 2015).

There is some evidence to show that 'thinness' at birth, although assessed by different methods, is significantly associated with coronary events, stroke, type 2 diabetes mellitus, adiposity, metabolic syndrome, and osteoporosis in adult life (Barker et al., 2005; Gluckman, Hanson, Cooper, & Thornburg, 2008; Osmond & Barker, 2000). This evidence suggests that low body fat may play a role in the future development of these conditions.

Long-term studies where a reliable assessment of newborn BF% is used may provide further insight into the consequence of reduced BF% on future morbidity. This is important, as there evidence is growing that low GWG has

been shown to be associated with low BF% and FFM, and if the link were proven to be causal this could provide a target for future interventions (Catalano et al., 2003; Oken et al., 2007; Zilko, Rehkopf, & Abrams, 2010).

Increased body fat

Increased BF% is thought to be a more important predictor of childhood obesity compared to birthweight. In their study to establish the effect of perinatal factors on childhood obesity, Catalano et al. (2009) demonstrated that BF% was associated with increased weight in childhood, however birthweight was not (Catalano et al., 2009). Raised fat mass at birth was also associated with metabolic dysfunction in childhood (mean age 8.8±1.8y). Furthermore, newborns in the upper tertile for percentage body fat had greater BF% in childhood (%) (39.3±4.3 vs. 19.7±2.6 vs. 28.2±2.6; p=0.0001), higher insulin resistance as assessed by the homeostasis model of insulin resistance model (3.42±1.72 vs. 1.55±0.56 vs. 2.22±1.15; p=0.002), higher triglycerides (mmol/L) (1.23±0.77 vs. 0.63±0.29 vs. 0.72±0.32 P = 0.009) and leptin concentrations (ng/ml) (15.9±7.0 vs. 2.5±0.6 vs. 7.6±4.7; p=0.0001) than children in tertiles 1 and 2 (Catalano et al., 2009).

Crozier et al. (2010) also demonstrated that increases in neonatal fat mass were associated with increases in childhood fat mass at ages 4 y (correlation co-efficient = 0.24) and 6 y (correlation co-efficient = 0.19). However no further analyses were presented to determine if this correlation was significant (Crozier et al., 2010).

Although there is limited long-term data available on the effects of raised BF% at birth, initial findings suggest that there is an increased risk of obesity and markers of morbidity from early childhood into adulthood. Well-designed long-

term studies are still required to understand the mechanism of these effects and how that knowledge can be utilised to reduce future morbidity.

1.3.6 Association between reduced newborn fat free mass on newborn and long-term outcomes

Stunting is defined as 'short length for age implying linear growth restriction' (Victora et al., 2015). Although causes of stunting are complex, maternal undernutrition, intrauterine growth restriction (Black et al., 2008), recreational drug use, smoking and small maternal stature are well-recognised associations (Victora et al., 2015). Whilst stunting is often not diagnosed until the age of 2-3 years, the process is thought to start in-utero. Stunting detectable at birth is thought to result from reduced fetal accumulation of FFM (Dewey & Begum, 2011; Catalano et al., 2014; Villar et al., 2017). The long-term effects of stunting include: short adult stature, reduced lean body mass, cognitive impairment, poor school achievement and reduced adult earnings (Black et al., 2008; Dewey & Begum, 2011). Recommendations are that interventions to prevent stunting should be implemented from pre-conception to age of 2 years in order to prevent long-term effects (Black et al., 2008; Dewey & Begum, 2011).

1.4 Known effects of gestational weight gain on newborn body composition

Ten studies examining the association between GWG and NBC were identified (Table 1.2) with a total of 4014 mother-infant pairs. These studies ranged in size from 172 to 836 participants. All studies were from high-income countries with seven of the studies were from North America, and one study from the UK, Australia and Denmark respectively. In general, increased GWG was associated with higher BF% and FFM. It was difficult to directly compare the effect size of GWG on NBC between studies as exposures and outcomes were presented in different formats. Weight gain was presented as: total GWG as a continuous variable, rate of weight gain per week, or as a categorical variable “normal’ or “excessive” as defined by the IOM guidelines. The methods of assessing newborn composition also varied. Skinfold thickness, being the cheapest and most readily available, was the most commonly used method (Catalano et al., 1995; Davenport, Ruchat, Giroux, Sopper, & Mottola, 2013; Waters, Huston-Presley, & Catalano, 2012). This was followed by ADP (Au et al., 2013; Hull et al., 2011; Starling et al., 2015), DEXA (Carlsen et al., 2014; Crozier et al., 2010) and TOBEC (Catalano et al., 2003). One study used two methods; both skin fold and TOBEC (Sewell, Huston-Presley, Super, & Catalano, 2006). The effect of GWG on NBC was also presented in different formats: Mean ± SD, variance, regression coefficients or stratified by BMI group or diabetic status. In addition, some studies presented FM and not BF%.

Table 1.2: Studies showing the association between GWG, other predictors and newborn body composition

Study details	Inclusion/exclusion criteria	Method of assessing NBC	Method of assessing GWG	Covariates considered	GWG effect
Starling 2015 **Prospective** **Single site** **Diverse ethnicity** **826 participants** **USA**	Inclusion: live birth >37weeks; NBC measurements within 3 days of birth; Exclusion: multiple pregnancy; previous still birth; pre-existing diabetes; asthma; steroid treatment, cancer; psychiatric illness; age<16years; >24 weeks gestation	ADP	Rate per week	Maternal age Household income Education Smoking Parity Maternal race GA BMI	An increase in GWG of 0.1kg per week was associated with an increase in BF% of 0.55% (95% CI:0.37,0.72) and FFM of 34.0g (95% CI:21.4, 46.6).
Carlsen 2014 **Prospective** **Single site** **311 participants** **Denmark**	Inclusion: Low risk, singleton, >37weeks; normal BMI Exclusion: chronic disease leading to intrauterine growth restriction; NICU admission; congenital disease	DEXA	Continuous	BMI Parity GA Gender	A 1kg increase in GWG was associated with an increase in BF% of 0.2% [(95%CI:0.1, 0.3) p<0.001] and FFM of 13g [(95% CI:3,9) p<0.001]
Davenport 2013 **Prospective** **Single site** **172 participants** **Canada**	Inclusion: BMI>18.5kg/m^2; gestation; 18 years old Exclusion: Multiple pregnancy; GDM; pre-eclampsia; contraindications to exercise	Skinfold measurements	Categorical as per IOM guidelines	BMI Maternal age GA Gender	Across all BMI groups, women with excessive GWG delivered newborns with a mean BF% of18.7±3.3% and those with appropriate GWG delivered newborns with a mean BF% of 13.2±4.1% (p<0.01)

Study details	Inclusion/exclusion criteria	Method of assessing NBC	Method of assessing GWG	Covariates considered	GWG effect
Au 2013 **Prospective** **Single site** **599 participants** **Australia**	Exclusion: Congenital anomalies, pre-existing diabetes, admission to NICU>48 hours	ADP	Continuous	Gender Ethnicity GA BMI Parity Hypertension GDM	A 1kg increase in GWG was associated with an increase in BF% of 0.13% (95CI:0.07,0.19) (p<0.01)
Waters 2012 **Prospective** **Single site** **439 participants** **USA**	Inclusion: Low risk; singleton; delivered>36 weeks; normal GTT Exclusion: Pre-existing diabetes; hypertension, Pre-eclampsia; renal disease; congenital anomaly	Skinfold measurements	Continuous	Smoking Gender Maternal age Maternal race GA BMI	3.2% (P<0.0001) of the variation in BF% and 0.8% (p<0.0001) of variation in FFM was explained by GWG

Study details	Inclusion/exclusion criteria	Method of assessing NBC	Method of assessing GWG	Covariates considered	GWG effect
Hull 2011 **Prospective** **Single site** **Diverse ethnicity** **306 participants** **USA**	Inclusion: healthy newborns; singleton birth; >37weeks delivery Exclusion: birth defects; congenital abnormality; admission to neonatal intensive care unit; gestational diabetes; hypertension; pre-eclampsia	ADP	Categorical as per IOM guidance	Gender GA Infant age Maternal race Maternal age BMI	In women with excessive GWG obese women compared with normal BMI women delivered newborns with a higher mean BF% (14.2±0.8% vs.11.8± 0.5) (p=0.011) -Overweight women also delivered newborns with higher mean BF% compared with women with normal BMI (13.7 ± 0.7 vs.11.8± 0.5) (p=0.019) Across all BMI groups women with excessive GWG delivered infants with higher FFM compared with appropriate GWG: (2912g vs. 2797g) (p=0.013)
Crozier 2010 **Prospective** **Single site** **564** **UK**	Inclusion: Southampton women's survey prior to pregnancy; >37week delivery; Exclusion: seen after last menstrual period; perinatal death; congenital growth abnormalities	DEXA	Continuous	Smoking Maternal age Height Parity Education	A 1kg increase in GWG was associated with an increase in FM of 100g [(95% CI 40,150) P=0.0004] and FFM of 30g (-10, 70) p=0.14

Study details	Inclusion/exclusion criteria	Method of assessing NBC	Method of assessing GWG	Covariates considered	GWG effect
Catalano 2003 **Prospective** **Single site** **415 participants** **USA**	Inclusion: 195 participants- GDM; 220 participants- NGT Exclusion: birthweight<2kg; multiple pregnancy; congenital abnormalities	TOBEC	Continuous **(**GWG was not a predictor for GDM women)	Maternal age Height Weight GA Gender Family history of diabetes Race Smoking status	24% (p=0.006) of the variation in FM and of variation in FFM 24% (p=0.05) was explained by GWG
Sewell 2006 **Prospective** **Single site** **220 participants** **USA**	Exclusion: birthweight<2000g; congenital anomalies; multiple pregnancy; pre-existing diabetes; GDM;	Skinfold measurements; TOBEC	Continuous	GA GTT result Gender Smoking Ethnicity Parity GWG	**In women with a BMI< 25kg/m^2** 19.2% (p=0.041) of the variation in FFM was explained by GWG **In women with a BMI>25kg/m^2**: 13% (p=0.002) of variation in FM was explained by GWG
Catalano 1995 **Prospective**; **Single site**; **183 participants** **USA**	Inclusion: Low risk women; USS before 20 weeks/ certain LMP; non-smokers; Normal GTT	Skinfold measurement	Continuous	GA Parity	16% (p<0.05) of the variation in FM was explained by GWG 23% (p<0.05) of the variation in FFM was explained by GWG

[a]: ***Abbreviations***: *BMI: Body Mass Index; GA: Gestational age; GDM: Gestational diabetes mellitus; GTT: Glucose tolerance test; GWG: gestational weight gain; NBC: Newborn body composition NGT: Normal glucose tolerance; PPW: pre-pregnancy weight; TOBEC: total body electricity; USS-ultrasound scan)*

CHAPTER 2: SYSTEMATIC REVIEW OF THE METHODOLOGICAL QUALITY OF STUDIES AIMED AT CREATING GESTATIONAL WEIGHT GAIN CHARTS

2. Systematic review of the methodological quality of studies aimed at creating gestational weight gain charts

Introduction

The goal of this systematic review was to evaluate the methodological quality of studies that aimed to create GWG charts by scoring them against a set of predefined, independently agreed criteria. The number and quality of studies of GWG charts is unknown. Therefore, the aims of this study were to identify and compare the methodological quality of published GWG charts and then make recommendations for the essential reporting criteria for future research in this field. A synthesis of available literature concerning GWG standards and outcomes related to excessive and insufficient GWG are presented in Chapter 1.2.3-1.2.5.

Methods

This systematic review of observational studies was conducted and reported using the checklist proposed by the Preferred Reporting Items for Systematic Reviews and Meta-Analysis- PRISMA (Moher, Liberati, Tetzlaff, Altman, 2009). Electronic databases (MEDLINE, EMBASE and Web of Science) were systematically searched to identify studies aiming to create GWG charts between 1960 and September 2017. Prior to 1960, patient characteristics, data analysis and study design were deemed unlikely to be comparable with contemporary studies. The search was limited to full articles in English as it was not possible to identify two independent reviewers with a third mediator in other languages.

Primary research articles were included if: 1) participants were pregnant women; 2) weight was reported in standard units, pounds (lb.) or kilograms (kg); 3) the study had a longitudinal or cross-sectional design, and the main objective was to create a GWG chart describing changes throughout gestation. Articles were excluded if: 1) they were validation studies of previously described charts or review articles or recommendations synthesized from several sources (e.g. the IOM 2009 guidelines (Medicine, Institute of, 2009)); 2) the aim was to describe GWG limited to a very specific population (e.g. adolescents, women with medical conditions), 3) GWG recommendations were not provided in a format describing change throughout pregnancy or 4) the primary outcome was not maternal GWG.
A keyword search strategy was conducted in consultation with a specialist medical librarian, using Medical Subject Headings (MeSH) terms related to

pregnancy (“gestational” or “pregnancy” or “pregnancy outcomes” or “pregnant women”), AND weight (“body weight” or “body weight changes” or “ideal body weight chart” or “weight gain”), AND charts (“charts” or “body weight charts” or “reference values”. Combinations of the following keyword search terms were also included: “curves”, “centiles” “maternal weight”, “pregn*”, “chart*”, “gestational weight gain”, “range”, “reference”, “trend” and “gest*”.

One reviewer screened the titles and abstracts of all identified citations and selected potentially eligible studies for which full text articles were obtained and screened further for eligibility. Reference lists of retrieved articles were examined for additional and relevant citations. The list of articles was assessed for duplications and these removed. Two reviewers independently checked and scored the methodological quality of each eligible article against the criteria specified in Table 2.1. Discrepancies in scoring were discussed and a third reviewer consulted if needed. Statistical criteria were independently checked and scored by a medical statistician.

Methodological quality criteria specific for studies examining growth in the perinatal period have been developed and validated by the INTERGROWTH 21st group (Ioannou et al., 2012; Napolitano et al., 2014). These criteria were modified and adopted by the reviewers in advance of this review (Ohadike et al., 2016). There were 23 quality criteria within 3 domains consisting a total of 34 items: study design (12 criteria/ 21 items), statistical methods (7 criteria/

7items) and reporting methods (4 criteria/ 6 items). For each item, studies were scored as either high (score=0) or low (score=1) risk of bias (Table 2.1). The data extracted from studies are detailed in Table 2.1 along with the rationale for selecting these criteria. All scoring and study details were entered into a Microsoft Excel database (Microsoft Excel for Mac 2011 Version 14.4.7).

The total quality score (QS) for each study was defined as the percentage of low risk of bias scores out of the total possible score. Mean scores for each domain were also calculated. If it was inappropriate make a judgment about an item due to study design or if scoring an item would unfairly downgrade a study, then this item was excluded from the denominator when calculating the QS. The QS was also calculated using all items as a denominator and these studies were not unfairly upgraded either. Finally, studies were ranked by QS, and the similarities and differences amongst the recommendations of the top ranking studies were compared.

Results

Basic results and demographics

The search yielded 1,804 citations, from which 120 full-text articles were obtained. Of these, 14 met the inclusion criteria. The reasons for exclusion are given in Figure 2.1. The 14 studies published between March 1985 to May 2016 (Abrams, Carmichael, & Selvin, 1995; Calvo et al., 2009; Carmichael, Abrams, & Selvin, 1997; Cheikh Ismail et al., 2016; Dawes & Grudzinskas, 1991; Guelinckx et al., 2010; Hutcheon et al., 2013; Johansson, Hutcheon, Stephansson, & Cnattingius, 2016; Ochsenbein-Kölble, Roos, Gasser, & Zimmermann, 2007; Rosso, 1985; Straube, Voigt, Briese, Schneider, & Voigt, 2008; Theron & Thompson, 1990; W. Wong, Tang, Lau, & Wong, 2000; Xu, Luntamo, Kulmala, Ashorn, & Cheung, 2014) provided data from 2,414,930 pregnancies. The fourteen studies originated from 15 countries in 5 regions (Argentina, Brazil, Belgium, China, Germany, Hong Kong, India, Kenya, Malawi, Oman, South Africa, Switzerland, Sweden, UK and USA). Only one study was performed in multiple countries (Brazil, China, India, Italy, Kenya, Oman, UK and USA) (Cheikh Ismail et al., 2016)

Characteristics of studies

QS ranged from 26% to 94%. Five studies had QS above 70%: Cheikh-Ismail et al. 94% (Cheikh Ismail et al., 2016), Xu et al. (Xu et al., 2014) 85% Calvo et al. (Calvo et al., 2009) 79%, Hutcheon et al. (Hutcheon et al., 2013) 72% and Guelinckx et al. (Guelinckx et al., 2010) 71%. A summary of the assessed studies is shown in Table 2.2. Overall risk of bias scores across the studies for each of the criteria are shown in Figure 2.2.

Main sources of bias

The most common potential source of bias was in the lack of objective measurement of pre-pregnancy maternal weight, used to calculate the baseline for GWG. In all studies but Cheikh-Ismail et al. (2016) this was obtained through maternal recall and never actually measured.

A further common potential source of bias was in the absence of anthropometry quality control, with failure to report how measurements were validated. Only one study performed more than one measurement at each visit to minimise intra-observer error (Cheikh Ismail et al., 2016). Four studies named the instruments used to weigh women and described the weighing technique (e.g. with light clothing, shoes off). Choices of weighing instrument and calibration frequency were not consistently described: seven studies provided no information about how this was performed.

Study design domain

Of the 14 studies, five were prospective and planned. The remaining nine presented data collected from other studies or were retrospective analyses of existing databases and birth records. Eleven studies were longitudinal and three cross-sectional in design. Within the longitudinal studies, seven documented the number of weight measurements taken throughout pregnancy, which ranged from two to eight. Seven (63%) longitudinal studies specified an appropriate method of data analysis that accounted for repeated measurements within subjects.

There was a range in both the accuracy of gestational age assessment and the gestation at the first weight was recorded in each study. Gestational age was determined by a first trimester ultrasound scan in nine studies, by participant recall of the date of the last menstrual period in two, and undocumented in the remaining three. Four studies did not state the gestational age at which maternal weight was first measured. In the remaining studies it ranged from 5 to 26 weeks of gestation. Weight was measured monthly in six studies, however in the remaining eight, the frequency of measurements was unclear or undocumented.

Inclusion and exclusion criteria were provided in eleven studies; however, there was little consensus on what these criteria should be. Women were most commonly excluded for multiple pregnancy or pre-existing hypertension or diabetes, however none of the studies excluded all these conditions. Nine studies were limited to women deemed 'low-risk', defined either prospectively (e.g. women without any notable risk factors in their pregnancy or medical history), or retrospectively (i.e. women selected after delivery deemed to have experienced an uncomplicated pregnancy with a 'healthy' term baby). Cheikh-Ismail et al. (2016) prospectively enrolled women with low risk pregnancies, optimal health, nutrition and socioeconomic status. Baseline characteristics of women included in the study sample were presented in ten out of fourteen studies.

Statistical methods domain

In two studies (Carmichael et al., 1997; Dawes & Grudzinskas, 1991), since GWG recommendations were based on crude statistical analytical summaries

of observed associations between adverse neonatal or maternal outcomes and weight gain, scoring using the pre-specified criteria was not possible. Of the remaining 12, eight studies clearly described the statistical methods used; with nine of these also reporting the covariates in the analyses. Smooth centiles were created in nine studies, with seven reporting information on estimate precision (i.e. standard errors or confidence intervals).

Reporting methods domain

Regression equations were provided for five studies. One study reported both conditional and unconditional standards (Xu et al., 2014). The distribution of GWG at each gestation was presented as z-scores in three studies (Hutcheon et al., 2013; Johansson et al., 2016; Xu et al., 2014) with an additional three providing enough information to calculate z-scores (Abrams et al., 1995; Calvo et al., 2009; Cheikh Ismail et al., 2016) .

Assessed studies recommendations

GWG recommendations from each study are provided in Table 2.3. Five studies (Calvo et al., 2009; Carmichael et al., 1997; Dawes & Grudzinskas, 1991; Rosso, 1985; W. Wong et al., 2000) based GWG recommendations on adverse clinical outcomes however the outcomes selected varied. For example, Wong et al. (2010) selected birth weight, term delivery, absence of gestational diabetes and pregnancy-induced hypertension whereas others selected only birth weight (Calvo et al., 2009; Rosso, 1985).

Discussion

This systematic review demonstrates considerable heterogeneity in the methodology of studies used to create GWG charts. The median QS was 56%, with 5 studies scoring below 51% (Carmichael et al., 1997; Dawes & Grudzinskas, 1991; Rosso, 1985; Straube et al., 2008; Theron & Thompson, 1990) and five scoring above 70%(Calvo et al., 2009; Cheikh Ismail et al., 2016; Guelinckx et al., 2010; Hutcheon et al., 2013; Xu et al., 2014). Given the wide range of scores (26% to 94%), it is likely that some of the observed variations in recommendations may be due to methodological differences rather than underlying discrepancies in GWG patterns between populations.

The review has numerous strengths: 1) A systematic literature search was conducted using explicit and reproducible methodology; 2) Well-established tools (including a PRISMA checklist and flowchart) were used to guide the review process to ensure reliability; 3) By using a predefined set of methodological quality criteria, studies could be compared objectively; 4) These criteria were formulated *a priori* and adapted from previously validated methods (Ioannou et al., 2012; Napolitano et al., 2014); 5) Historical quasi-scientific studies were excluded; 6) Efforts were made to reduce bias when scoring studies: for example, unfair downgrading of studies was avoided by omitting inapplicable criteria from the score calculation; 7) The scoring criteria were designed to identify important factors for judging the validity and reliability of research guiding GWG recommendations, and 8) To minimize potential unconscious bias, the scoring system was deliberately kept simple, with each item scored only as low or high risk of bias.

Limitations of the study include the decision to exclude non-English language publications which resulted in the exclusion of some important studies in this field (Mardonnes & Rosso, 1997; Grandi & Luchtenberg, 2007; Vila-Candel & Hevilla-Cucarella, 2009). As this review aimed to assess the methodological quality of GWG charts, we feel a sufficient number of studies were included to provide a range of authors, locations, times and populations, and highlight the main methodological strengths and weaknesses of research in this field. A further limitation of methodological quality scoring systems is the assumption that each item has equal importance with the same potential to produce biased results. This is clearly an oversimplification as, for example, poor study design could substantially reduce the overall quality of the study from the outset, despite all other elements being performed appropriately.

This review has demonstrated that GWG studies are prone to errors in measurement of baseline (usually pre-pregnancy) weight and gestational age estimation, both of which are essential for accurate determination of GWG. Maternal recall of pre-pregnancy weight reporting has potential to cause bias. Russell et al. (2013) reported that women systematically under-report their pre-pregnancy weight; however, the issue is how big the magnitude of bias is. An alternative would be to use weight measured in the first trimester as the baseline thereby providing an objective and replicable measurement (Fattah et al., 2010). Accurate knowledge of gestational age is also essential for determining the trend in GWG, which may not be linear. Without an accurate gestational age, there may be a failure to recognise women delivering preterm, thus underestimating total GWG (Hutcheon et al., 2013). We found

that five studies used methods of gestational age assessment prone to error and bias (Carmichael et al., 1997; Ochsenbein-Kölble et al., 2007; Straube et al., 2008; W. Wong et al., 2000), with three providing no information about how gestational age assessment was performed. This finding is consistent with those of a recent systematic review on fetal growth assessment in pregnancy (Ioannou et al., 2012) and highlights a significant potential source of bias in pregnancy research. Of note, however, is the study by Cheikh-Ismail et al. (2016) had an objective measure of gestational age and weight in the first trimester. These were the most reliable anthropometric measures in any of the studies.

Four of five studies with quality scores above 70%, were from various single-site locations: Argentina (Calvo et al., 2009), Belgium (Guelinckx et al., 2010), Malawi (Xu et al., 2014) and the United States of America (Hutcheon et al., 2013), which arguably limits their potential for international use. International studies are required in order to determine if geographic or ethnic differences exist in patterns of GWG (World Health Organisation, 1995). Again, the study by Cheikh-Ismail et al. (2016) fulfilled these criteria as their study was conducted in eight geographical regions.

Z-scores were presented only in studies conducted after 2013 (Hutcheon et al., 2013; Johansson et al., 2016; Xu et al., 2014), indicating that research in this field is continuing to develop. The use of Z-scores is well established in fetal, infant and child studies for classifying growth during and after pregnancy (IAEA, 2013; World Health Organisation, 1995; Villar et al., 2013a) They

provide a method of describing weight gain independently of gestation (Hutcheon et al., 2013; Xu et al., 2014) and make it easier to describe the severity of abnormal weight gain patterns. Z-scores are also useful for epidemiological analyses as their statistical characteristics make them less prone to bias and difficulties in interpretation due to non-linearity (Allison, Paultre, Goran, Poehlman, & Heymsfield, 1995; Calvo et al., 2009).

Furthermore, using a Z-score allows adherence to the reference distribution, provides a linear scale permitting summary statistics, has uniform criteria across indices and is useful for detecting changes at extremes of distributions. Application of GWG charts in clinical practice implies that there are known thresholds above or below which the risk of adverse outcomes increases. Ideally, these thresholds should only be determined by assessing the association with maternal and neonatal outcomes that are specifically related to excessive or insufficient GWG. We found that several studies included outcomes unlikely to be causally linked to GWG (e.g. stillbirth), which makes it difficult to compare recommendations across studies.

Of the studies that produced charts of GWG as a continuous variable through pregnancy, the majority are references as opposed to standards. The INTERGROWTH 21st study was the only one to create GWG standards because their population was at low risk of adverse outcomes (Cheikh Ismail et al., 2016).

Future studies need to move towards creating standards instead of references to enhance quality and applicability and recognising the use of reference values to make value judgements and recommend interventions.
We observed no common exclusion criteria across studies. Moreover, few studies excluded women with well-recognised factors to adversely affect GWG such as pre-existing diabetes, smoking, alcoholism, illicit drug use and inflammatory bowel disease (Agency for Healthcare Research, 2008). Agreement on standard criteria to define healthy, low-risk pregnancies would enable valid comparisons between studies. The INTERGROWTH-21st Project Consortium has provided a detailed list of such criteria, which could be applied for future research in this field. Rather than exclude all possible risk factors, it was decided to limit them to factors that would prevent women being classified as low-risk for antenatal care (Villar et al., 2013).

From this review, we recommend that the future reporting of studies GWG should adhere to a minimum set of criteria (Table 2.4). It is recommended that anthropometric studies should apply standardised internationally agreed protocols for measuring techniques to ensure that results are accurate, reliable and comparable (de Onis, Onyango, Van den Broeck, Chumlea, & Martorell, 2004; Cheikh Ismail, Knight, Ohuma, Hoch, & Chumlea, 2013; World Health Organisation, 1995).

The proposed minimum requirements for creating future GWG charts should be: 1) Data collection is population-based and prospective; 2) There is detailed description of and attention to reliable evaluation of gestational age

by ultrasound examination before 14 weeks gestation; 3) The selection and measurement of the baseline weight is fully described; 4) Comprehensive descriptions of measurement procedures, instruments and protocols including efforts to minimise intra and inter-observer bias; 5) Methods to calculate of adequate sample size are provided; and 6) Appropriate longitudinal statistical modelling techniques used with creation of smoothed centiles and z scores. Adherence to these criteria would represent significant progress in the rigour of reporting in this field and help to address the question of how much of the differences in GWG charts are due to local factors. The study by the INTERGROWTH 21st group is a step in the right direction to developing GWG standards with strict anthropometric protocols, objective and rigorous assessment of GWG and creation of regression equations to enable calculation of z-scores (Cheikh Ismail et al., 2016). However, as their population had normal BMI, it would be essential to extend the development of standards to those with underweight, overweight and obese BMI.

In conclusion, there is a lack of international agreement on what constitutes adequate GWG. We have demonstrated substantial heterogeneity in the methodological quality of studies used to make recommendations for GWG. It is possible that this heterogeneity has contributed to some of the variation in current GWG recommendations around the world. Further high-quality international studies at low-risk of methodological bias are needed to guide future clinical recommendations and counselling of women.

Table 2.1: Methodological quality criteria

Domain	Low risk of bias (score = 1 for each point)	High risk of bias (score =0 for each point)	Rationale
Study design (maximum points):			
1.1 Aim of publication (2)	The aim is clearly stated to produce a chart or recommendations for maternal Aims in abstract and paper are the same weight gain during pregnancy	Aim not clearly defined Different aims in abstract and paper	More reliable charts will be obtained if the primary purpose of the study was to create them
1.2 Definition of target population (1)	Target population clearly defined, i.e. geographical location, ethnic group	Not clearly defined	Determines to which population the chart can be appropriately applied
1.3 Definition of reference or standard (1)	Authors state if the trend in weight gain is a reference or standard	Not clearly defined	Reference and standards charts are different tools and, therefore, users must know which they are
1.4 Design (1)	Clearly described and longitudinal	Not reported or not longitudinal	Good scientific practice
1.5 Sample selection (6)	Population-based Clearly defined inclusion and exclusion criteria Women selected and enrolled consecutively and in accordance with target population Planned study before enrolment	Not population Not clearly defined inclusion or exclusion criteria -based Convenience sampling or not described Unplanned study: retrospective study or secondary analysis of data collected for another research study	Good scientific practice

Domain	Low risk of bias (score = 1 for each point)	High risk of bias (score =0 for each point)	Rationale
1.6 Pre-pregnancy weight and measurements during pregnancy (2)	Pre-pregnancy weight is recorded Each woman's weight measured at least every month	No pre-pregnancy weight taken or based on maternal recall alone No clear documentation of intervals between weight measurements or less than monthly	Poor reliability of measurements increases dispersion
1.7 Inclusion/ exclusion criteria (1)	Prescriptive approach to sample selection, i.e. only women at low risk of complications in a defined population included	Not clearly documented General population sample	Enable creation of standard
1.8 Sample size (1)	Documented determination/calculation of sample size and justification	No documented determination/calculation of sample size and justification	Good scientific practice More reliable charts created
1.9 Data collection (2)	Prospective study Data collected specifically for creating GWG charts	Retrospective study	Good scientific practice
1.10 Weight evaluation (3)	Weight measured using standardized instruments Instruments and protocols reported in detail e.g. calibration schedule, training of staff Gestation ages at which measurements taken documented	Standardized instruments not used or not described Not described No reported	Poor reliability of measurements increases dispersion
1.11 Number of observations for each stage of pregnancy (1)	Reported	Not reported	Precision of estimates increases with the number of observation
1.12 Gestational age determination (1)	Clearly described and reliable, i.e. last menstrual period confirmed by ultrasound scan at 9-14 weeks	Not described	Charts will be based on gestational age
2. Statistical methods:			
2.1 Number of measurements taken at each visit (1)	More than one measure of weight per woman at each visit	Single measure of weight per woman at each visit	Precision of estimates increases with the number of observation
2.2 Statistical methods (1)	Clearly described	Not clearly described/performed	Crucial to use an adequate statistical model for creating reliable charts
2.3 Assessment of goodness of fit (1)	Test of goodness of fit reported	Goodness of fit test not reported	Allows the reader to evaluate the adequacy of the model used to trace charts
2.4 Methods used to estimate the reference intervals /standards (1)	Mean and standard deviation (SD) model LMS method	Inadequate	Measure of reliability of results

Domain	Low risk of bias (score = 1 for each point)	High risk of bias (score =0 for each point)	Rationale
2.5 Precision of estimates (1)	Standard errors or confidence limits reported	Not reported- single mean or median value only reported	Measure of reliability of results
2.6 Covariates (1)	Charts stratified by pre-pregnancy BMI	Not clearly described	Some covariates may exert effects on GWG
2.7 Smooth centiles (1)	Smooth centiles created	Raw centiles only presented	Smoothing reduces the fluctuations observed in raw centiles due to sampling variability
3. Reporting methods:			
3.1 Characteristics of study population (1)	Presented clearly or in a table	Not presented	Determines if the population studied belongs to the target population
3.2 Presentation of recommendations (1)	Presented as gestation specific mean/SD in the form of centiles charts or Z-score charts	Not clearly described or single recommendation for whole of pregnancy or trimester specific recommendations	Aids interpretation of results
3.3 Report of regression equations for the mean and SD if relevant (1)	Reported	Not reported	Use of appropriate statistical model for creating reliable charts
3.4 Chart presentation (3)	Values of the at least 10^{th} 50^{th} and 90^{th} centiles or parameters that allow them to be computed are reported Z-scores presented or computable Conditional and unconditional standards included	Not reported Not reported Not reported	Graphical method has a lower accuracy than use of numerical values Use of z-scores increases the reliability of size evaluation. It also allows comparison among different subjects or population.

Table 2.2: Summary of assessed studies

Reference	Country	Sample size	Study design	Gestation first weight (weeks)	Number of measurements	Data	Quality Score (%)
Cheikh-Ismail (2016)	Multiple (8)	4607	Longitudinal	<14	6	Prospective	94
Xu et al (2014)	Malawi	358	Longitudinal	14-25	1733 in total	Prospective	85
Calvo et al (2009)	Argentina	1090	Longitudinal	12	8 per participant	Prospective	79
Hutcheon et al (2013)	USA	648	Longitudinal	6	≥5 per participant	Retrospective	72
Guelinckx et al (2010)	Belgium	605	Longitudinal	<12	5341 in total	Retrospective	71
Johansson (2016)	Sweden	141,767	Longitudinal	<14	711,615 in total	Retrospective	62
Ochsenbein-Kolble et al (2007)	Switzerland	4034	Cross sectional	>5	Not applicable	Retrospective	59
Wong et al (2000)	Hong Kong	504	Longitudinal	ND	6 per participant	Retrospective	53
Abrams et al (1995)	USA	10418	Cross sectional	14-26	Not applicable	Retrospective	53
Theron et al (1990)	South Africa	1003	Longitudinal	<24	Not documented	Prospective	50
Carmichael et al (1997)	USA	7002	Longitudinal	ND	2 per participant	Retrospective	50
Rosso et al (1985)	USA	262	Longitudinal	12-14	Not documented	Prospective	48
Dawes et al (1991)	UK	1145	Longitudinal	<20	6 per participant	Retrospective	46
Straube et al (2008)	Germany	2241487	Cross Sectional	ND	Not applicable	Retrospective	26

Table 2.3: Recommendations/Results of assessed studies

Study	Format of chart	Recommendation[1] (in kg) Underweight	Normal	Overweight	Obese	All women	QS[8]
Cheikh-Ismail (2016)	Smoothed centiles	N/A	13.7±4.5	N/A	N/A		94
Xu et al (2014)	Conditional and unconditional percentiles	-	-	-	-	GWG[7]: 6kg (unconditional chart)	85
Calvo et al (2009) [2]	Body weight and BMI[6] centiles	12.2 ± 3.9	12.1 ± 4.3	12.1 ± 5.0	10.2 ± 4.8	Mean weight gain 11.9 ± 4.4	79
Hutcheon et al (2013) [3]	Smoothed means, SD[9], centiles in all women	15.7 ± 4.6	15.5 ± 5.3	15.1 ± 6.2	9.9 ± 6.5	-	72
Guelinckx et al (2010) [2]	Centiles in all BMI categories	15.4 ± 4.1	15.1 ± 4.5	13.7 ± 5.3	12.0 ± 5.9	Mean weight gain 14.8 ± 4.7	71
Johansson (2016)	Smoothed centiles; 50th centiles reported	13.7	14.3	13.8	11.6(class1) 9.7(class2) 8.1(class3)		
Ochsenbein-Kolble et al (2007)	Centile charts for Caucasian, Asian and Black ethnic groups	-	-	-	-	Mean weight gain 15.5±5.9 5th centile 5.7 95th centile 25.4	59
Wong et al (2000)[3]	Centiles in all women and in underweight, normal and overweight BMI groups	15.1 ± 3.8	13.8 ± 4.2	11.2 ± 5.2	-	-	53
Abrams et al (1995)	Raw and fitted regression average Centiles in Asian, Hispanic, Black, White ethnic groups						53
	1st trimester/ week					0.169 ± 0.268	
	2nd trimester/week	-	-	-	-	0.563 ± 0.236	
	3rd trimester/ week	-	-	-	-	0.518 ± 0234	

Study	Format of chart		Recommendation[1] (in kg) Underweight	Normal	Overweight	Obese	All women	QS[8]
Theron et al 91990)	Centile chart for all women, for women>25 years +BMI>24 and women <25years or BMI<24		-	-	-	-	0.45	50
Carmichael et al (1997)[4]	Centiles in all BMI groups	-	-	-	-			50
	1st trimester total	1.92 ± 3.06	2.19 ± 3.47	2.16 ± 3.95	1.65 ± 3.94			
	2nd trimester/ week	0.57 ± 0.20	0.58 ± 0.20	0.51 ± 0.24	0.41 ± 0.27			
	3rd trimester/ week	0.48 ± 0.19	0.51 ± 0.21	0.49 ± 0.22	0.47 ± 0.24			
Rosso et al (1985)[5]	Maternal body weight target as percent of calculated 'standard weight'	11.7 ± 8.3	10.4 ± 6.3	7.3 ± 6.6				48
Dawes et al (1991)	Weight gain mean +/- 1SD in all women Weight gain in women with infants of birthweights >90th centile compared to 10-90th centile	-	-	-	-		Mean maternal weight gain 10.71±4.3 Mean weekly gain 0.38±0.16	46
Straube et al (2008)	Centiles for individualized weight, height phenotype	-	-	-	-	Centile charts provided for 12 groups of women. Largest group women 161-171cm In height and weighing ≤64kg: 5th centile 6.5; 50th centile 14.0; 95th centile21.0		26

[1] All units of weight in kilograms or kilograms ± 1 Standard deviation [2] BMI categories as per IOM (Medicine, Institute of, 2009)- Underweight: <19.8; Normal: 19.8-26; Overweight: 26-29; Obese: >29 [3] BMI categories as per Asian BMI standard (W. Wong et al., 2000)- Underweight: <19; Normal: 19-23.5; Overweight: >23.5; [4] BMI categories are as per WHO (Organisation, 1995)- Underweight: <18.5; Normal: 18.5-24.9; Overweight: 25-29.9; Obese: >30 [5] Weight categories as per Rosso 1985 (Rosso, 1985)- Underweight:<89% of standard weight; Normal: 90-110% standard weight; Overweight:>111% standard weight [6] BMI- body mass index [7] GWG- gestational weight gain [8] QS- quality score [9] SD- standard deviation

Table 2.4: Recommendations to create accurate Gestational weight gain charts in future[6]

Domain	Requirements
Study design	Prospective study Planned study before enrolment Women enrolled consecutively and in accordance with target population Population based study with efforts to exclude those with conditions known to affect GWG Longitudinal design Multiple measurements from each individual on multiple occasions through the pregnancy
	Reliable evaluation of pre-pregnancy weight Weight objectively measured with standardised scales by trained data collectors Weight taken with light indoor clothing- Footwear, coats, sweaters and heavy clothing removed
	Reliable evaluation of gestational age Early pregnancy (9-13^{+6} weeks) scan to determine gestational age
	Detailed description of measurement procedures, protocols Standardised training of data collectors Data collected specifically for creation of GWG charts Standardised instruments with specific calibration schedule Gestation at which measurements taken accurately recorded
Statistical methods	Creation of smooth centiles Use of appropriate statistical methods to convert raw centiles to smooth centiles Sample size Adequate sample size for each range of measurements to allow presentation by z scores and centiles Regression based methods used
Reporting methods	Centile charts or z scores Charts presented with at least 10th, 50th and 90th centiles or parameters that allow them to be computed Z scores presented or computable

[6] Taken from Ohadike et al. Adv Nutr 2016;7:313–22

Figure 2.1: PRISMA flowchart

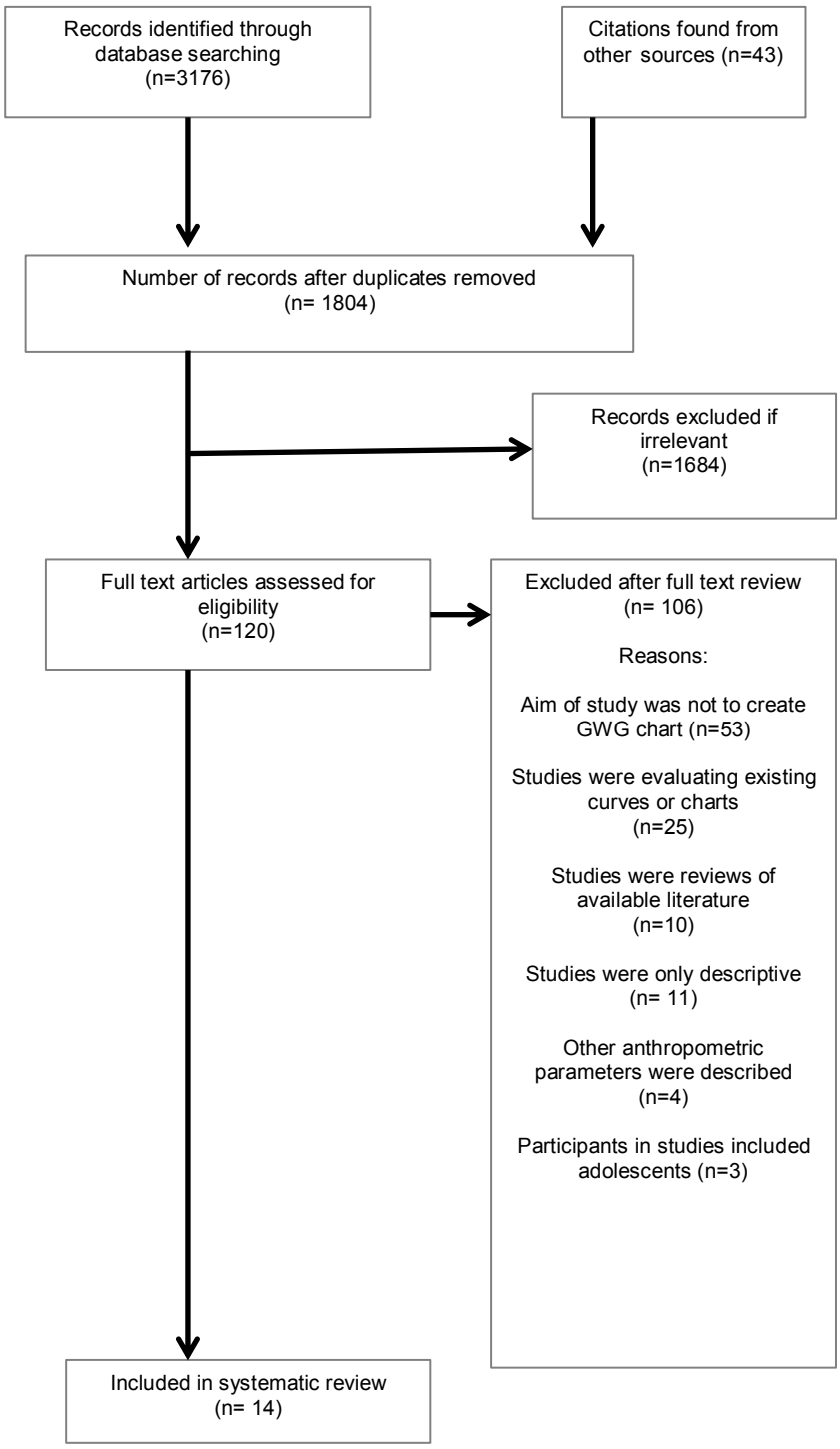
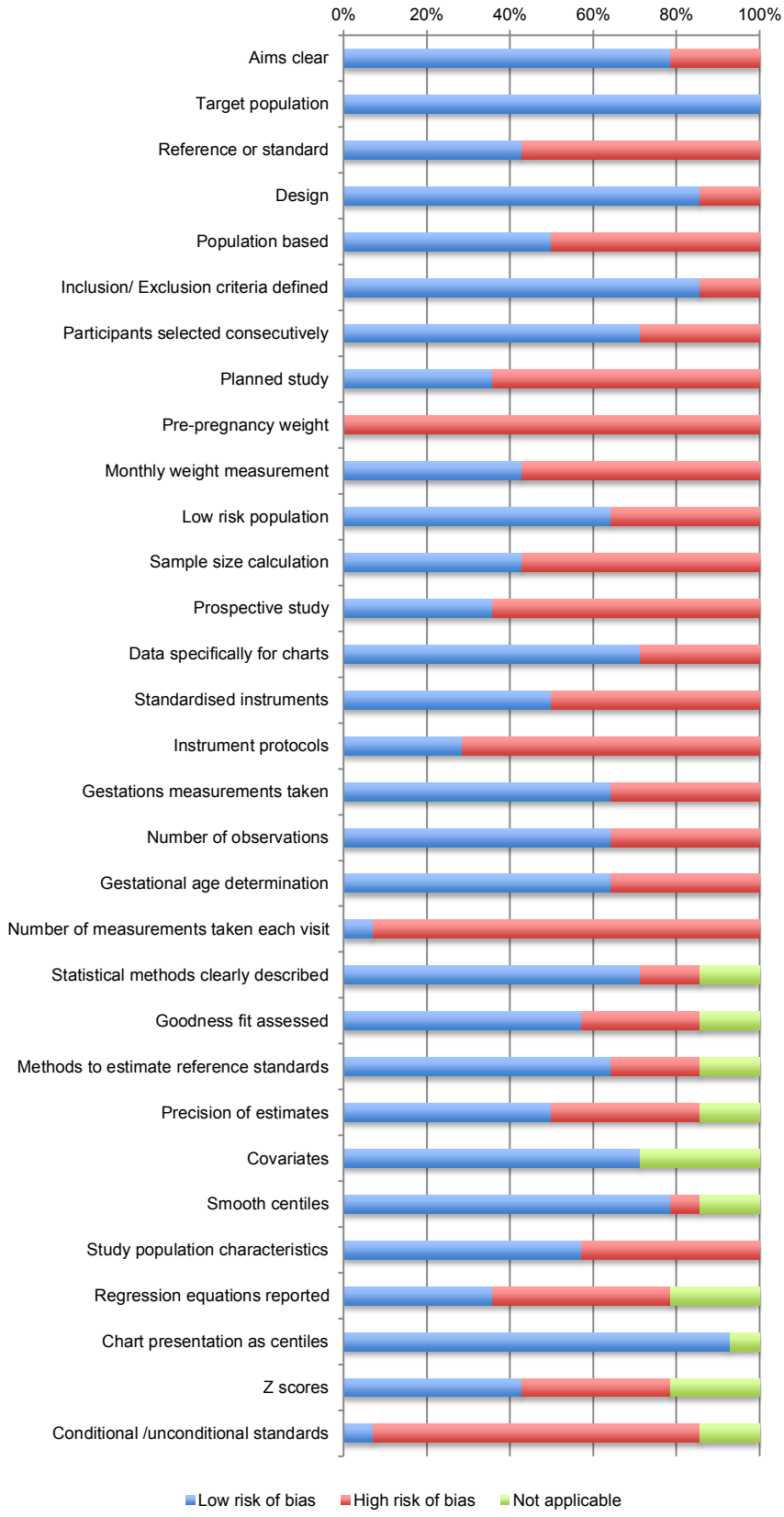


Figure 2.2: Overall methodological quality of included studies assessed against quality criteria (corresponds to criteria in TABLE 2.1)



AIMS AND OBJECTIVES

Aims and Objectives

The systematic review in chapter 2 provides an assessment regarding the issues and drawbacks surrounding available GWG charts by scoring them according to pre-defined criteria. These criteria were designed to address the ideal characteristics of any study seeking to create GWG standards and z-scores.

This thesis aims to explore the association between GWG z-scores, newborn body composition (BF% and FFM) and the strongest modifiers on this relationship, in ethnically and geographically diverse, moderate to high-risk populations.

Hypotheses:

1. GWG z-scores are significant predictors of newborn BF% and FFM
2. BMI, gestational age, gestational diabetes, and gender represent potential confounders of the relationship between GWG z-scores and newborn body composition.

CHAPTER 3: METHODS

Methods

Introduction

This is a secondary analysis of data collected in the INTERBIO-21st Study, an extension of the INTERGROWTH-21st Project. As mentioned before, INTERGROWTH-21st is a large international population-based project that investigates growth and development from <14 weeks' gestation up to 2 years of age. Whereas INTERGROWTH-21st populations were typically low risk in order to create new prescriptive growth standards, populations in INTERBIO-21st were at higher health risks such as malnutrition, infections, smoking and obesity. The INTERBIO-21st Study sites were: Brazil (Pelotas); Kenya (Kilifi); Kenya (Nairobi); Thailand (MaeSot); Pakistan (Karachi); South Africa (Soweto) and the UK (Oxford). This analysis includes data from the UK, Thailand and South Africa, as only these sites had access to the equipment to assess NBC (INTERBIO 21st Consortium Study Protocols, 2012).

Study sites

The UK site was based at the John Radcliffe Hospital in Oxford, which serves a population of approximately 667,000 within Oxfordshire and neighbouring counties. The John Radcliffe Hospital is a large tertiary teaching hospital with approximately 7300 deliveries per year.

The South African site was Baragwanath Hospital, Soweto, Johannesburg. Baragwanath Hospital has a secondary and tertiary care obstetric unit, attached to the University of the Witwatersrand and serves a population of over 2 million in the township of Soweto. It has an average of 22,750 deliveries per year.

The Thailand site is centred around the Shoklo Malaria Research Unit, which is a field station of the Mahidol-Oxford Research Unit of Mahidol University in Bangkok and is located on the Thailand-Myanmar border. The site composed of three clinics, which deal with a total of approximately 2,100 deliveries per year. The data was obtained from the Maela camp, the largest refugee camp along the border.

Data collection

Trained anthropometrists at all study sites collected maternal and newborn anthropometric data, demographic data and variables using strict and standardised questionnaires and protocols (INTERBIO 21st Consortium Study Protocols, 2012). For the purpose of this analysis, we used the following variables: maternal age (years); weight at first assessment (kg); height at first assessment (cm); BMI at first assessment (kg/m^2); weight at each antenatal visit; parity; total GWG (kg); smoking (Y/N); recreational drug use (Y/N); gestational diabetes (Y/N); pre-existing diabetes (Y/N); pre-existing hypertension (Y/N); pregnancy-induced hypertension (Y/N); pre-eclampsia (Y/N); gestational age at delivery (weeks); neonate gender (M/F); birthweight (g); birth length (cm); newborn head circumference (cm); newborn body fat percentage; age at PEAPOD measurement (weeks) and gestational age at booking (weeks) and delivery (weeks).

Selection criteria

Women aged 18 years and over attending antenatal care with naturally conceived singleton pregnancies were included in this analysis. This was irrespective of their risk profile for adverse pregnancy or neonatal outcomes.

The median number of ultrasound scans (excluding dating scan) was 5 (range from 1 to 7 scans) performed every 5 weeks (± 1 week) from 14 weeks to 42 weeks of gestation. Maternal weight was measured at each ultrasound visit.

Measurement of newborn body composition

Birth length and head circumference were measured in all neonates enrolled in INTERBIO 21st. In a select number of cases, birthweight and newborn body fat percentage were measured using the PEAPOD® device (see chapter 1.3.3). All measurements were performed within one week of birth (INTERBIO 21st Consortium Study Protocols, 2012).

Assessment of maternal BMI and GWG

BMI and weight were measured at the first antenatal appointment and weight at each subsequent study appointment thereafter. More information about the anthropometric protocols can be found in the INTERGROWTH-21st supplement (Cheikh Ismail, Knight, Bhutta, Chumlea et al. 2013a; Cheikh Ismail et al., 2013b). Total GWG was calculated as the weight at first assessment subtracted from the weight taken at the final scan appointment before delivery.

A summary of the measurement protocols is shown in Appendix 2

Ethical approval

All women were eligible to participate after giving written informed consent. This project was approved as an extension of INTERBIO-21st study, which had been approved by the Oxfordshire Research Ethics Committee 'C' (reference: 08/H0606/139) and research ethics committees of the individual participating institutions and corresponding health authorities where the

project was implemented. The INTERGROWTH-21st/INTERBIO-21st Committee approved this secondary analysis.

Data analysis

Outcome variable

The primary outcome was NBC, which is divided into newborn body fat percentage and fat free mass (g). Newborn body fat percentage and weight were measured directly by the PEAPOD® and used to calculate fat free mass (g) using the following formula: Birth weight (g) - (newborn body fat percentage x birthweight in (g)).

Exposure variables

These variables were selected, as they were plausibly associated with GWG and NBC from the literature review conducted prior to this study (see Table 1.2). They included parity, smoking, recreational drug use, pre-existing diabetes, endocrine disorders, gestational diabetes, essential hypertension, pregnancy induced hypertension, pre-eclampsia, age at PEAPOD® measurement, gestational age at delivery and newborn gender. Maximum, minimum, mean and standard deviations were calculated for continuous variables. All binary variables were coded as (0 or 1), with absence of the condition designated as the referent group. BMI was used as a continuous variable and also recoded as a categorical variable in keeping with international categorisations of underweight, normal weight, overweight and obese used in clinical practice. In order to account for differences in gestational age at measurement, z-scores of GWG were calculated according

to the GWG standards from the INTERGROWTH-21st Project (Cheikh Ismail et al., 2016).

Statistical analysis

The data was screened for missing values, outliers and implausible values.

Univariable associations

The effect of the following variables on newborn body fat percentage and fat free mass were assessed using simple linear regression analyses: total GWG; GWG z-score; weight; BMI at first assessment (BMI groups: Underweight (<18.50 kg/m$^{2)}$, Normal (18.50-24.99 kg/m^2), Overweight (25.00-29.99 kg/m^2), Obese (>30.00 kg/m^2 with the normal BMI as the referent group)); smoking; recreational drug use; pre-existing diabetes; endocrine disorders; gestational diabetes; essential hypertension; pregnancy-induced hypertension; pre-eclampsia; gestational age at delivery; age at PEAPOD® measurement and newborn gender. Statistical significance was set at a p-value <0.10 to select variables to be included in the multiple regressions. All analyses were adjusted for study site

Multivariable associations

Many factors are known to affect NBC and neonatal size. In order to establish whether there was any measurable confounding on the effect of GWG on NBC components, multiple regression analyses were performed to control for potential measured confounders. Two multiple regression models were fitted with BF% and FFM as the dependent variables and GWG z-scores as the predictor. Model 1 was adjusted for BMI group (as above) only. Model 2 included the following variables as covariates: age at PEAPOD®

measurement (weeks); BMI groups; pre-eclampsia; pregnancy induced hypertension; maternal smoking; recreational drug use; pre-existing diabetes; gestational diabetes; endocrine disease; newborn gender and gestational age at delivery (days). Again, all analyses were adjusted for study site.

The analyses were performed using IBM SPSS Statistics for Macintosh (IBM Corp. Released 2013. IBM SPSS Statistics for Macintosh, Version 22.0. Armonk, NY: IBM Corp.).

CHAPTER 4: RESULTS

Results

A total of 535 women and their newborns were included in this analysis: 407 from the UK, 90 from Thailand and 38 from South Africa.

Anthropometric and Baseline characteristics

Maternal and neonatal baseline characteristics are shown in Table 1. Overall, women had their first antenatal appointment early (12.1 ± 1.6 weeks) (mean±SD) and delivered at full term (39.7 ± 1.3 weeks). The majority of participants were aged between 26 and 35 years and those who were under 20 or over 40 years of age accounted for less than 0.05%. Overall, there were low rates of co-morbidities in women in the study. The commonest medical complication was pregnancy-induced hypertension, affecting 9.0% of participants. Sixty-seven participants (12.5%) were smokers and five (0.9%) used recreational drugs.

Most women (60.0%) had a normal BMI at the first study visit. Fewer women were underweight (3.2%) compared to the proportion that were overweight and obese, 26.4% and 10.5% respectively. Mean GWG was 11.7 ± 4.6kg, however there was a wide range from weight loss in 3 participants to a gain of 25.2 kg (range - 2.1 to 25.2 kg). The three participants who lost weight (0.6%), with losses of 2.1kg, 1.4kg and 1.3kg.

A boxplot of the distribution of GWG according to 1st trimester BMI (Figure 4.1a) demonstrates that obese women on average gained less weight than women who were underweight, normal weight or overweight. Figure 4.1b and

4.1c show that newborn BF% and fat free mass were highest in the obese women and lowest in those who were underweight. A one-way ANOVA showed that these differences were significant for both BF% and FFM (BF%: $F=3.788$, $p=0.005$; FFM: $F=3.693$, $p=0.006$).

There was a wide range in the birthweight of newborns from 2.1 to 6.1 kg (3.3 ± 0.5kg). All PEAPOD measurements were taken within the first week of birth (mean 0.2 ± 0.2weeks) with the majority (92.5%) taken within 72 hours of birth. The mean newborn body fat percentage (BF%) was 10.5 ± 3.9% (range 1.2 to 25.5%) and mean fat free mass (FFM) was 2923 ± 383 g (range 2031 to 4783 g).

Table 4.1: Baseline characteristics of all participants

Variable	Values
Maternal age (y)[1]	30.6±5.3
Maternal height (cm)[1]	162.4±8.1
Maternal weight in first trimester (kg)[1]	64.1±13.3
Maternal BMI in first trimester (kg/m^2)[1]	24.2±4.1
GA at first assessment (weeks)[1]	12.1 ± 1.6
GA at delivery (weeks)[1]	39.7 ± 1.3
Total GWG (kg)[1]	11.7±4.6
Birthweight (g)[1]	3279.3±503.1
Birth length (cm)[1]	49.1±2.2
Birth head circumference (cm)[1]	34.1±1.3
Body fat percentage (%)[1]	10.5±4.0
Fat free mass (g)[1]	2923.1±383.3
Previous pregnancies[2]	
None	207 (38.7)
1	216 (40.4)
≥2	112 (13.5)
Smoking[2]	67 (12.5)
Recreational drugs[2]	5 (0.9)
History of diabetes[2]	3 (0.6)
Endocrine disorders[2]	18 (3.4)
Gestational diabetes[2]	20 (3.7)
Pre-existing Hypertension[2]	30 (5.6)
Pregnancy induced hypertension[2]	48 (9.0)
Pre-eclampsia[2]	8 (1.5)
Newborn gender[2]	
Female	257 (48.0)
Male	278 (52.0)
BMI group[2]	
Under 18.50	17 (3.2)
18.50-24.99	321 (60.0)
25.00-29.99	141 (26.4)
30 and over	56 (10.5)
Total GWG (kg)[2]	
-2.1 to 0.0	3 (0.6)
0.1 to 5.0	34 (6.4)
5.1 to 10.0	162 (30.3)
10.1 to 15.0	213 (39.8)
15.1 to 20	98 (18.3)
20.1 to 25.2	25 (4.7)

Sample size: 535
[1] Mean ±SD
[2] Number (%)

In order to investigate any population differences the baseline characteristics by study site (UK, Thailand and South Africa) were analysed. These are shown in Table 4.2. Women in the UK had the highest maximum GWG attained (25.2kg) and also the lowest GWG (-2.1kg). BF% was highest in the South African babies women compared with those born in the UK or Thailand population (11.5±4.0% vs. 10.7±3.8% vs. 9.1±4.3%). Conversely, FFM was lowest in the South African population compared with the UK and Thailand population (2664.4±263.4g vs. 2979.7±375.1g vs. 2776.5±382.0g).

Table 4.2: Baseline characteristics by study site

Variable	UK (n=407)	Thailand (n=90)	South Africa (n=38)
Maternal age (y)	31.3±4.7	27.4±6.3	30.2±5,8
Maternal weight in first trimester (kg)	67.2±12.3	49.3±7.6	66.0±10.0
Maternal BMI in first trimester (kg/m^2)	24.6±4.1	21.4±3.1	26.3±3.9
Birthweight (g)	3350.2±492.3	3067.2±510.1	3023.0±368.4
Total GWG (kg)	12.3±4.5	9.9±3.8	9.7±5.6
GA at delivery (weeks)	39.7±1.5	39.7±1.0	39.2±1.4
Body fat percentage (%)	10.7±3.8	9.1±4.3	11.6±4.0
Fat free mass (g)	2979.7±375.1	2776.5±382.0	2664.4±263.4

All results Mean±SD

Univariable analyses

GWG, expressed as a z-score, was significantly associated with both newborn BF% and FFM. For each unit increase in z-score BF% increased by 1.24% (95%CI: -0.13, 2.61, p=0.08) and FFM by 346.5g (95%CI: 219.53,473.47, p<0.01) (Table 4.3).

When assessed individually, heavier maternal weight and higher BMI, later GA at delivery and the presence of a maternal endocrine disorder were also associated with increases in BF% and FFM. In contrast, babies born to mothers who smoked had both lower BF% and FFM. Maternal diabetes, both pre-pregnancy and gestational diabetes, was associated with a higher BF% but not FFM. In contrast, maternal recreational drug use was associated with a lower FFM but not BF%. The later the PEAPOD® measurement was performed from birth, the lower the FFM, however there was no effect on BF%. BF% was found to be lower in males, whilst FFM was higher [BF%: -1.82% (95%CI: -2.47, -1.18; p<0.01); FFM 154.79g (95%CI: 93.26, 216.31; p<0.01].

When compared to women with normal BMI, newborns born to underweight women had both a lower BF% and FFM [-1.83% (95%CI: -3.74,0.08; p=0.06); -292.98g (95%CI: -478.51, -107.44; p<0.01)], whilst newborns born to overweight and obese women only had higher BF% [0.77% (95%CI: -0.01, 1.54; p0.05) and 1.47% (95%CI: 0.36, 238; p=0.01) respectively]. The strongest positive correlation between BMI and BF% was in obese women, with an increase of 1.5% compared to women in the normal BMI range.

Table 4.3: Univariable* associations with newborn body composition

Variable	Body fat percentage Coefficient (95%CI)	P value	Fat free mass (grams) Coefficient (95%CI)	P value
GWG z-score	***1.24 (-0.13,2.61)***	***0.08***	***346.50 (219.53, 473.47)***	***<0.01***
Total GWG (kg)	0.08 (0.01,0.16)	0.03	24.63 (17.96, 31.30)	<0.01
Weight at first assessment (kg)	0.04 (0.02, 0.07)	<0.01	6.39 (3.70, 9.07)	<0.01
BMI at first assessment (categorical) (kg/m^2)				
Normal BMI is referent group				
Underweight (under 18.50)	**-1.83 (-3.74, 0.08)**	**0.06**	**-292.98 (-478.51, -107.44)**	**<0.01**
Overweight (25.00-29.99)	**0.77 (-0.01, 1.54)**	**0.05**	29.09 (-46.23, 104.41)	0.45
Obese (30.00 and over)	**1.47 (0.36, 2.38)**	**0.01**	88.53 (-19.43, 196.49)	0.11
Smoking	**-0.90 (-1.90, -0.10)**	**0.08**	**-111.74 (-206.20, 17.272**	**0.02**
Newborn gender (Male)	**-1.82 (-2.47, -1.18)**	**<0.01**	**154.79 (93.26, 216.31)**	**<0.01**
GA at delivery (days)	**0.06 (0.03, 0.09)**	**<0.01**	**19.56 (16.81, 22.30)**	**<0.01**
Recreational drug use	0.78 (-2.68, 4.23)	0.66	**-333.07 (-659.02, -7.13)**	**0.05**
Pre-existing diabetes	**4.61 (0.18, 9.04)**	**0.04**	-63.57 (-484.70, 357.56)	0.77
Endocrine disorders	**1.74 (-0.10, 3.6)**	**0.06**	**157.66 (16.32, 331.64)**	**0.08**
Gestational diabetes	**1.86 (0.12, 3.61)**	**0.04**	24.83 (-140.83, 190.49)	0.77
Pre-existing hypertension	0.60 (-0.85, 2.04)	0.42	-48.39 (-185.41, 88.63)	0.49
Pregnancy induced hypertension	-0.02 (-1.18, 1.15)	0.98	-44.70 (-154.98, 65.58)	0.43
Pre-eclampsia	0.62 (-2.11, 3.36)	0.66	-111.49 (-370.32,147.34)	0.40
Age at PEAPOD measurement (weeks)	0.25 (-1.93, 2.44)	0.82	**-244.11 (-450.00, -38.23)**	**0.02**

*All analyses controlled for study country; Sample size for each variable 535; Significance- P≤0.1

Multivariable regression

In Model 1, adjusting for BMI group with the normal weight group as the reference group, this resulted in a modest change in the effect on BF% [1.41% (95%CI: 0.03,2.79; p=0.05)] and FFM [365.10g (95%CI: 237.74,492.45; p<0.01)] (Table 4.4).

The final stepwise regression (model 2) is shown in Table 4.5. Appendix 3 shows each stage of the stepwise regression. The association between GWG and NBC remained significant. Adjusting for these factors, for each unit increase in GWG z-score, newborns had an increase of 1.82% (95%CI: 0.48, 3.15; p<0.01) in their BF% and 292.71g (95% CI: 186.70, 398.73; p<0.01) in their FFM.

When compared to mothers with a normal BMI, newborns born to obese mothers had a statistically significantly higher BF% and FFM. All included covariates increased BF% and FFM except smoking and underweight BMI. In addition, older newborn age at PEAPOD® measurement reduced FFM and not BF%.

Table 4.4: Model 1-Multiple regression* assessing the effect of GWG z-score on newborn body composition adjusted for BMI groups

Variables	Body fat percentage Coefficient (95% CI)	P value	Fat free mass (grams) Coefficient (95% CI)	P value
Final GWG z-score	**1.41 (0.03, 2.79)**	**<0.01**	**365.10 (237.74, 492.45)**	**<0.01**
Adjusted for:				
BMI at first assessment (categorical) (kg/m^2)				
Normal BMI is referent group				
Underweight	-1.22 (-3.16, 0.72)	0.22	**-231.52 (-410.01, -53.03)**	**0.01**
Overweight	0.49 (-0.30, 1.28)	0.22	21.48 (-51.08, 94.04)	0.56
Obese	**1.32 (0.19, 2.45)**	**0.02**	**123.36 (19.09, 227.63)**	**0.02**

*All analyses adjusted for site of delivery; Sample size for each variable: 535; Significance- $p<0.05$

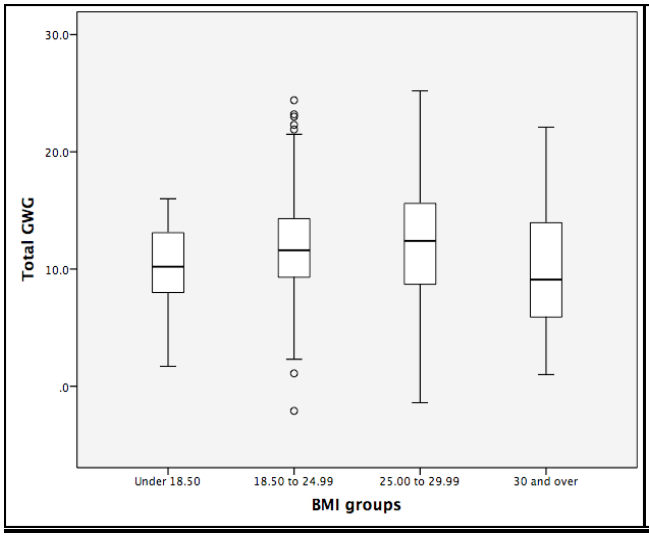
Table 4.5: Model 2- Stepwise multiple regression* assessing the effect of GWG z-score on newborn body composition with adjusted confounders.

Variables	Body fat percentage Coefficient (95% CI)	P value	Fat free mass (grams) Coefficient (95% CI)	P value
Final GWG z-score	***1.82 (0.48, 3.15)***	***<0.01***	***292.71 (186.70, 398.73)***	***<0.01***
Adjusted for:				
BMI at first assessment (categorical) (kg/m^2)				
Normal BMI is referent group				
Underweight	-1.13(-2.98, 0.720)	0.23	-124.73 (-271.69, 22.24)	0.10
Overweight	0.51 (-0.25, 1.28)	0.19	38.73 (-22.10, 99.56)	0.21
Obese	**1.13 (0.04, 2.23)**	**0.04**	**114.24 (27.78, 200.70)**	**0.01**
Newborn gender (Male)	**-1.92 (-2.55, -1.29)**	**<0.01**	**146.89 (96.58, 197.21)**	**<0.01**
GA at delivery (days)	**0.06 (0.02, 0.09)**	**<0.01**	**19.06 (16.49, 21.63)**	**<0.01**
Pre-existing diabetes	4.05 (-0.36, 8.46)	0.07	N/A	
Gestational diabetes	1.48 (-0.28, 3.24)	0.10	N/A	
Endocrine disorders	0.78 (-1.11, 2.67)	0.42	**153.48 (12.94, 294.00)**	**0.03**
Smoking	**-0.96 (-1.92, 0.00)**	**0.05**	-74.42 (-151.14, 2.29)	0.06
Age at PEAPOD measurement (weeks)	N/A	N/A	**-180.52 (-345.97, -15.06)**	**0.03**
Recreational drug use	N/A	N/A	-128.69 (-392.46, 135.08)	0.34

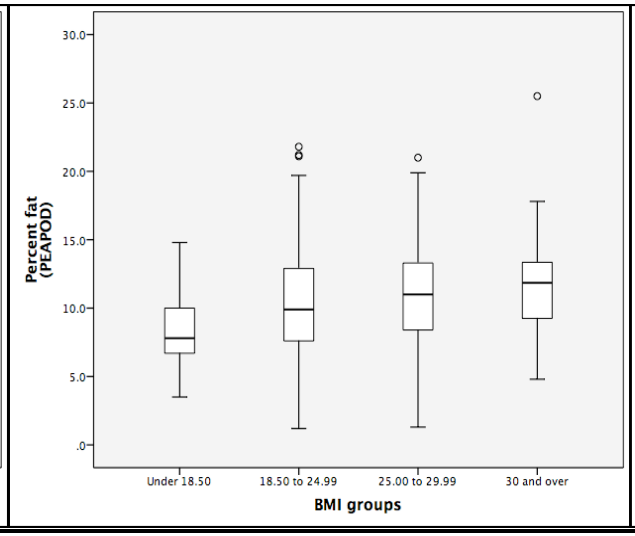
*All analyses adjusted for site of delivery; Sample size for each variable: 535; Significance- $p<0.05$

Figure 4.1: Boxplots of GWG, newborn body fat percentage and newborn fat free mass in BMI groups

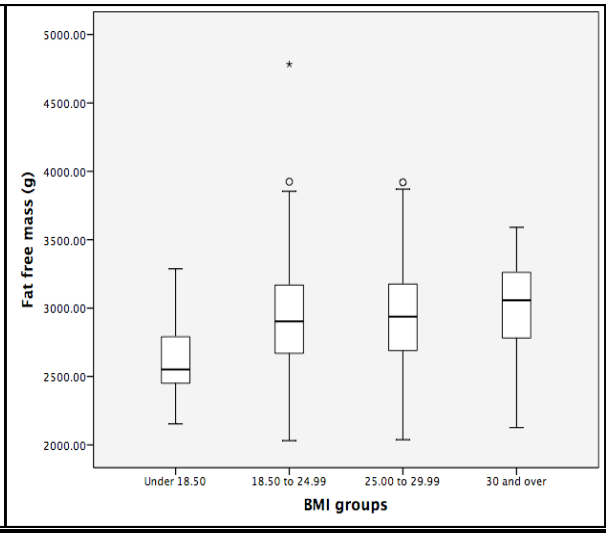
4.1a: Boxplot of Total GWG against BMI group



4.1b: Boxplot of body fat percentage against BMI group



4.1c:Boxplot of fat free mass against BMI group



CHAPTER 5: DISCUSSION AND CONCLUSIONS

Discussion

This study shows that greater maternal GWG is associated with increases in newborn BF% and FFM in a multi-ethnic, geographically diverse population.

Baseline characteristics

The study population characteristics were generally comparable to previous literature with moderate to high-risk participants. There were some differences that are highlighted below. Participants in this analysis were similar in age (30.6±5.3 years) on average to previous studies by Hull (2011), Au (2013) and Davenport (2013): 31±5.9; 33±5.0 and 32±4.0 years respectively. In addition, variables such as mean maternal height, weight BMI, gestational age at delivery, birth weight and birth length were similar. Mean birthweight (3279.3±503.1g) was also in a similar range than that reported in Hull et al. (2011), Au et al. (2013), De Cunto et al. (2013) and Davenport et al. (2013) (3250.7±451.8g; 3417±475g; 3220±640g; 3628±415g respectively).

Mean total GWG (11.7±4.6kg) was lower in this analysis when compared to Hull et al. (2011) and Au et al. (2013) (15.9±4.9kg; 13.1±5.9kg). It is worth noting that the other studies used participant recall to determine the pre-pregnancy weight as opposed to the objective measure used in this analysis. Evidence has shown that participants will consistently report a lower starting pregnancy weight than their measured weight (Russell, Gillespie, Satya, & Gaudet, 2013). This means the resultant GWG will be higher and this is the likely explanation for the observations increased GWG reported in other studies. Another possibility is that GWG was underestimated in the INTERBIO

21st study because first trimester weight was used instead of pre-pregnancy weight. However, there is evidence showing first trimester weight is suitable to use in lieu of pre-pregnancy weight as there is no significant difference (Fattah et al., 2010).

The mean total GWG in this analysis was 11.7±4.6kg and this is 2 kg lower than the standard of 13.7 ±4.5 kg reported for normal BMI women in Cheikh-Ismail et al (2016). This suggests that women in a population with higher rates of co-morbidities do gain less weight, however further work is required to assess if the clinical implications and the causality. Environmental factors may have also played a role as the women from the Thailand site were refugees and those from South Africa were from an economically deprived area.

Newborns in this analysis had lower levels of mean newborn body fat percentage than those reported in the study by Hull et al. (2011) (10.5±4.0 vs. 12.3±4.4%) but slightly higher than those in Au et al. (2013) and Starling et al. (2015) (9.2±4.4% and 9.1±3.9%). Again this may reflect the nature of this study population with higher risk pregnancies, as co-morbidities were not excluded. Hull et al (2011) excluded mothers with gestational diabetes, hypertension and pre-eclampsia whereas Au at al (2013) excluded mothers with pre-existing diabetes and babies admitted to neonatal care unit for more than 48hours. Mean fat free mass in this analysis was 2923.1±383.3g which was comparable to that in Au et al (2013) (2947±342g) but marginally higher than that in Hull et al (2011) and Starling et al (2015) (2842±347g and

2851±325g respectively. From this it would seem that FFM is influenced less by presence of co-morbidities than BF%.

Study sites

As expected, the mothers in the UK had the highest weight at first assessment and mean GWG, compared with those in Thailand and South Africa. This is likely to be because the nutrition status of the UK population is bound to be closest to ideal compared with the other two sites. Interestingly, maternal BMI and newborn BF% were highest in the South African population. These newborns also had the lowest FFM. As mentioned before, low FFM has been likened to stunting which can have serious implications for the child (Black et al., 2008; Catalano et al., 1995; Dewey & Begum, 2011). This finding reflects recent public health research regarding the increasing double burden of disease i.e. undernutrition and obesity faced by developing countries (Black et al., 2008). This combination of problems may have very different ramifications in the developing world, where there is a higher prevalence of exposure to infections during childhood, compared with the developed world.

As would be expected, the women in Thailand had lower GWG and BF% compared to the other two groups in this study, however, levels of BF% and FFM were still comparable to other studies where populations were not in environments of high adversity (Au et al., 2013; Starling et al., 2015). The one-way ANOVA showed that the differences in GWG, BF% and FFM amongst the groups were statistically significant, which, given their very different risk profiles and socio-economic situations, is not surprising. The value of comparing their growth against the same international standards for

fetal growth, gestational weight gain and newborn body composition, is that it allows formal measurement of this difference, highlighting women and babies at higher risk of complications.

General observations

The strongest modifiers of the association between GWG z-scores and both BF% and FFM were: maternal obesity, male gender and gestational age at delivery. Smoking only influenced the relationship with BF% while endocrine disorders and age at PEAPOD measurement only influenced FFM. Unlike previous studies (Catalano et al., 2003), GDM did not affect newborn body composition. At a rate 3.7%, GDM was less common in this study than in that of Catalano et al (2003); however in that study women with diabetes were actively recruited, with a smaller proportion of non-diabetic women. Whilst the rate of GDM in this study is representative of the incidence of GDM in the general UK population (National Collaborating Centre for Women's and Children's Health (UK), 2015), this study was likely underpowered to determine whether GDM influences newborn body composition independently of GWG. In addition, the diabetic women in Catalano et al. (2003) study were older, shorter and overweight/obese compared with the non-diabetic women. Those issues are likely to impact results, therefore the population of non-diabetic women and women with GDM were not comparable. I conducted a sensitivity analysis of the distribution of GWG, BF% and FFM was assessed in non-diabetic women only, and no significant differences were found compared with the general study population.

In contrast to Au et al (2013), I did not demonstrate any association between hypertension with BF% (Au et al., 2013). A possible explanation for this is that Au et al (2013) grouped all presentations of hypertension together (pre-existing, pregnancy induced and pre-eclampsia) whereas in this study these were analysed separately. Given the very variable nature of hypertensive disorders of pregnancy, I believe it is more clinically relevant to split the presentation of hypertension up so that specific conclusions and recommendations could be made. It is interesting that no association with hypertension was found in this study as hypertension is traditionally considered to be associated with intrauterine growth restriction and this is routinely monitored for in antenatal practice. Future detailed work into how hypertension affects newborn body composition and whether treated hypertension affects it differently may further improve our understanding.

The literature review did not identify any studies describing a significant association of endocrine disorders (excluding diabetes or GDM) or age at PEAPOD measurement with FFM. The endocrine disorders in this analysis included conditions such as thyroid disease and Addison disease. In previous studies, these conditions were not considered as possible modifiers and were not controlled for. In light of these initial findings, it would be important to tease out which specific conditions were associated with reduced FFM and determine in larger populations whether this observed association is replicable. This will help understand the potential relevance of these conditions in the clinical setting.

Previous studies use GWG either as a continuous variable when examining the association with BF% and FFM (Au et al., 2013; Crozier et al., 2010; Starling et al., 2015; Waters et al., 2012) or compare inadequate or adequate GWG as defined by the IOM recommendations (Davenport et al., 2013; Hull et al., 2011). The IOM GWG recommendations are the most commonly used recommendations since the initial publication in 1970 with updates in 1991 and 2009 (Medicine, Institute of, 2009). Using these guidelines however poses certain disadvantages as discussed in chapter 1.

In addition, the numbers of studies using the PEAPOD® machine to assess newborn body composition and establish the effect of GWG are also limited (Au et al., 2013; Hull et al., 2011; Josefson, Hoffmann, & Metzger, 2013; Starling et al., 2015). Other methods such as total body electricity (Sewell et al., 2006), skin fold thickness (Davenport et al., 2013) and Dual x-ray absorptiometry (Crozier et al., 2010) have been used as alternative methods of assessing newborn body composition. This study adds relevant and valuable information to the growing body of evidence using the current gold-standard method of assessing newborn body composition.

Whilst several groups have documented the association between GWG and BF% (Au et al., 2013; Davenport et al., 2013; Josefson et al., 2013), this study also confirmed an association with FFM. Other groups (Carlsen et al., 2014; Catalano et al., 1995; Starling et al., 2015) observed a similar association. These results imply that although FFM may increase at a constant rate, this velocity may still be influenced by maternal GWG. While previous studies

have focussed on the effect of increased GWG causing increased BF% and the consequences of this, none have focussed on increased FFM and what this may mean for the long-term health of the newborn. FFM is thought to be metabolically active and involved in the body's physiological processes, thus a possible hypothesis to explain my observation is that fetuses exposed to excessive GWG are trying to compensate for the increased BF% by accumulating more FFM.

Comparing the effect of GWG on newborn body composition between this analysis and previous studies is difficult as GWG is often presented in different formats, such as rate of GWG or comparison to IOM guidelines (see Table 1.2). Some authors only present results as percentage variation or variance (Waters et al., 2012; Sewell et al., 2006, Catalano et al., 1995), which has limited interpretation in other populations. Furthermore, this analysis has used GWG z-scores, not raw total GWG in kilograms, limiting the ability to draw comparisons with studies that present results in kg.

Another issue is that some studies reported only fat mass and not BF%, which is a better measure of adiposity (IAEA, 2013). Amongst the studies that did present both regression coefficients and BF%, each kilo of GWG resulted in similar increases in BF% of 0.2% (Carlsen et al., 2014) and 0.13% (Au et al., 2013) in spite of different methods to calculate newborn body composition being used (DEXA vs. ADP respectively). The effect of GWG on FFM is more disparate with one study showing that every kilo increase in GWG is associated with a 13g increase in FFM (Carlsen et al. 2014) and another with

an increase of 30g in FFM (Au et al., 2013). A possible explanation for this difference is the lack of anthropometric rigour and over reliance on self-report methodology in calculating GWG.

I will now discuss noteworthy areas that were highlighted during the assessment and interpretation of data. These concerned the following variables: gestational weight gain z-scores, BMI, gender, gestational age at delivery, smoking and age at PEAPOD® measurement.

GWG z-scores

Studies that assess the effect of total GWG on perinatal outcomes can be problematic, as this measure does not take into account the length of pregnancy. This is important because a premature end to a pregnancy means the mother has less time to gain weight therefore her total GWG will be less than a woman who goes to full term pregnancy. In addition, trying to establish which perinatal outcomes are due to prematurity and which are due to low GWG becomes unclear (Bodnar, Hutcheon, Parisi, Pugh, & Abrams, 2014; Hutcheon et al., 2013). Using the weekly rate of GWG does not improve reliability of results as the rate of weight gain differs from trimester to trimester. This means that a premature delivery could potentially result in a different rate of GWG as opposed to a term delivery (Bodnar et al., 2014).

It is important that gestational age related confounding is excluded in the design of any guidelines, recommendations or studies regarding effects of GWG. Using GWG z-scores in this analysis provided a simple and effective

way to avoid this confounding, as it was independent of gestation length thereby enhancing the reliability of the results. In order to improve our understanding of GWG on perinatal outcomes, it would be beneficial to utilise GWG z-scores in future studies as opposed to GWG in kilograms. This analysis is a step in the direction of understanding the association between GWG and short and long-term newborn, childhood and adulthood health outcomes.

BMI

In this population, both maternal BMI and GWG were independent predictors of NBC. It is thought that raised maternal BMI leads to increased levels of insulin resistance during pregnancy. This results in increased glucose levels crossing the placenta, resulting in fetal hyperglycaemia. This in turn leads to raised fetal insulin levels which in turn cause increased fetal body fat (Au et al., 2013; Ehrenberg, Dierker, Milluzzi, & Mercer, 2003; Institute of Medicine, 2009).

In the study by Carlsen et al (2014), maternal BMI was associated with an increased BF%. BF% was been found to be 11.2% in obese (BMI>30kg/m^2) participants and 8.8% non-obese participants although the birthweights were similar (3332g vs. 3322g). This is an important example of how birthweights are not specific enough and BF% may differ considerably thereby potentially sending the newborn down different paths in terms of future health risks. However, Au et al (2013) demonstrated that in women with an overweight BMI had reduced newborn body fat percentage. The authors attributed this to a

lower mean GWG in that group of women. Yet in the same study women with an obese BMI had neonates with raised body fat in spite of lower GWG, indicating possibly that the role of GWG on modifying BF% is attenuated in women in the highest BMI group (Au et al., 2013). It would be advisable for future studies to include BMI as a covariate in statistical analysis as it may well impact recommendations made from their results. Other studies (Hull et al., 2011; Sewell et al., 2006) did not control for GWG in assessing the effect of BMI on newborn body composition.

In light of the results from this analysis and previous studies, this provides further evidence of the importance of commencing pregnancy in the normal BMI range. Focussed nutritional guidance and support to maintain a healthy lifestyle before and during pregnancy could help reduce neonatal excess adiposity.

Gender

One of the strongest modifiers of the effect that GWG z-scores had on newborn body composition was neonatal gender, with males having 1.879% less body fat than females and 146.89g more FFM. This is in keeping with what other authors have observed (Simon et al., 2013; Villar et al., 2017). Simon et al. (2013) found that the difference between females and males was statistically significant 11.1±3.7% vs. 9.0±3.3%. These results were supported in Villar et al. (2017) who found that girls had a higher BF% than boys (10.7±4.0% vs. 9.6±4.0%) and lower FFM (2739±390g vs. 2965±422g). The Villar et al. population was from the low risk INTERGROWTH 21st population (Villar et al., 2017).

Differences in body composition between females and males is thought to be due to in utero production of testosterone which is assumed to increase the proportion of lean body mass in the fetus (Au et al., 2013; Simon et al., 2013). However, other studies have found an association only between neonate gender and FFM and not BF% (Hull, Dinger, Knehans, Thompson, & Fields, 2008; Waters et al., 2012). This was in contrast to the findings in this analysis where gender was significant for both BF% and FFM.

Further work will be required to understand the relevance of this difference in body composition in the context of future health risks to males and females. The question this raises is what the increased body fat and lower FFM in female's means for their future health. Potential issues to consider are how it influences their reproductive health, cardiovascular risk and risk of morbid obesity.

Gestational age at delivery

Gestational age at delivery is known to significantly influence fat mass and fat free mass in previous studies (Catalano et al., 2003; Sewell et al., 2006; Waters et al., 2012). This correlates with the findings in this study. In contrast to this, an interesting study by Simon et al. (2013) found that preterm infants had increased body fat percentage (Simon et al., 2013). However, the authors concluded that it was unlikely to represent a true reflection of increased body fat mass but simply the result of a much lower relative fat free mass ratio in preterm neonates. Therefore the findings in this analysis of increased BF%

and FFM with advancing gestational age holds true as it seems that the fetus accrues fat free mass in later gestation. This is supported by Villar et al. (2017) who have showed that FFM at 34 weeks gestation is approximately 2kg and at 40 weeks approximately 3kg (Villar et al., 2017).

Smoking

The percentage of smokers (12.5%) in this study was higher on average than that in other studies on newborn body composition with rates ranging from 4.2% to 9.2% (Au et al., 2013; Rodrigues et al., 2010; Starling et al., 2015). It was also higher than the rates in England and Wales (10.6%) Health and Social Care information Centre, 2016) and in countries such as Thailand (<5%) (World Health Organisation, 2010).

In previous studies, smoking was not associated with newborn body fat percentage (Au et al., 2013; Starling et al., 2015). However, in this analysis, smoking was associated with reduced newborn BF%. It is possible that other studies may have been underpowered to detect a significant association.

Smoking has been associated with miscarriage, still birth, and low birth weight and smoking cessation has been associated with a reduction in the incidence of low birthweight and an increase in mean birthweight (World Health Organisation, 2010). The results in this analysis, suggest this smoking related reduction in birthweight is attributable to a decrease in BF%. In light of evidence showing that reduced body fat mimics ‘wasting’ and its serious

consequences (Victora et al., 2015), further work in this area may help towards developing guidance towards smoking cessation in pregnancy.

Age at PEAPOD® measurement

In this analysis all PEAPOD® measurements were taken within the first week (6.93 days) of birth with a mean of 0.16 weeks (1.12 days) and standard deviation of 0.18 weeks (1.26 days). It was demonstrated that the later the PEAPOD examination was performed after birth, FFM significantly reduced.

Failure to account for the timing of the PEAPOD measurement after birth could potentially confound any observation on predictors of newborn fat free mass. In previous studies, the time at which body fat was measured ranged widely from: 1 day (Catalano et al., 2003; Waters et al., 2012) to 10 days (B. Eriksson et al., 2010). In some cases it was not documented (Catalano et al., 2014) .

It is therefore difficult to draw direct comparisons with previous work. From published reports, age at which body composition measurements were taken was not included in the statistical models. As this was identified as a significant potential confounder in my analyses, it would be advisable that future studies consider including this in analyses. At the very least, studies should record the age of the neonate at the time measurements were taken.

Strengths of study

The INTERBIO-21st study provides one of the largest observational datasets on newborn body composition and is the only study to date conducted across international sites. Trained anthropometrists using strict anthropometric protocols collected the data on GWG and PEAPOD measurements used for the analyses. Data, including outcome data, was collected in a prospective, reliable, objective and standardised manner by appropriately trained personnel. This is a unique feature of this study, with other studies containing potential important biases such as self-report of pre-pregnancy weight.

Multiple country sites were included in this study adding information that may be more generalisable to multi-ethnic populations and increasing external validity. As described, there were expected differences in the populations across the three sites and therefore to combine results, study site was adjusted for in all analyses.

An objective and well established device was use to measure body fat percentage and lean body mass. The PEAPOD machine has been validated against existing methods of assessing newborn body composition and been found to be reliable. It is also time saving, acceptable to parents and least distressing to neonates out of all methods to assess newborn body composition (IAEA, 2013; Ma et al., 2004; Yao et al., 2003).

The PEAPOD measurements were performed within a week of birth, reducing potential effects of postnatal influences on BF%. To improve reliability further,

age at PEAPOD measurement was still retained in the regression analysis in this study. This was proven to be important as a significant modifier of the effect of GWG on FFM.

Variables with high levels of collinearity with the exposure and outcome variables, such as weight and height with BMI, were excluded from the multiple regression analyses. Finally, using the standardised measure of GWG z-scores allowed assessment independent of gestation length bringing this study in line with current anthropometric study methodology.

Weaknesses of study

This study has some limitations. In addition, the participants were first assessed between 9-13+6 weeks and so the first weight measurement was in the first trimester rather than pre-pregnancy. However, by providing an actual measure, this analysis avoided the risk of participant recall bias.

Total weight gain as opposed to period specific weight gain was used in this analysis. This means that it is not possible to assess if or how period specific weight gain affects newborn composition. Two recent studies addressed this question. In Starling et al. (2015) increased early and mid pregnancy GWG caused neonatal FM to increase whereas GWG in later pregnancy resulted in an increase in FFM. (Starling et al., 2015). Davenport et al. (2013) found that excessive GWG in the first half of pregnancy was a strong predictor of excessive neonatal fat (Davenport et al., 2013). These findings are interesting

and potentially mean that if considering interventions to modify GWG, it would be more important to implement these in early pregnancy or pre-pregnancy. More work is required to confirm whether the timing of GWG is as important or more important than total GWG when considering the effect on newborn body composition.

GWG z scores for the analysis were calculated using the formulae for calculating mean and SD of GWG for normal weight women. Equivalent equations for underweight, overweight and obese women were not available at the time of writing this thesis. This has the potential to have introduced bias into interpretation of the GWG z scores for these women. Given however that the direction of effect size was consistent with previous observations, it is hoped that any such bias would be small.

Further work

Women in this study gained less weight on average than the low risk women who were included in the INTERGROWTH-21st Standards. It will be important to try and explore why this occurred, and whether it has short or longer-term consequences for the mother or child. This could affect the advice given to mothers with higher risk pregnancies. These initial steps of using standards and GWG z-scores are promising, however, what is now needed is to broaden the study of the effect of GWG z-scores to other women with other BMI groups and determine the association with important perinatal outcomes.

The premise this work is based on is that body composition at birth is important for future health. Longitudinal cohort studies that assess body composition at birth with comprehensive cardiovascular and metabolic risk throughout childhood assessment will be vital. Hopefully this will identify potential modifiable targets, such as GWG, around which suitable clinical trials can be developed.

Conclusion

In this large, well-designed, ethnically diverse study, increased gestational weight gain was shown to be associated with higher body fat percentage and fat free mass in newborns. The landmark report from the IOM (Institute of Medicine, 2009) recommended that future studies focussed on the medium and long-term effects of gestational weight gain. This thesis is a first step in addressing this recommendation to understand the effects of gestational weight gain. BMI and gestational weight gain represent two modifiable factors that influence the intrauterine environment. In turn, this is likely to affect newborn body composition, which may have far-reaching consequences for childhood and adulthood health. While long-term follow up of the newborns assessed in this study is required to determine actual associations with future morbidities, these findings add to a growing body of literature about newborn body composition and will help guide future research.

REFERENCES

References

Abrams, B., Carmichael, S., & Selvin, S. (1995). Factors associated with the pattern of maternal weight gain during pregnancy. *Obstetrics & Gynecology*, *86*(2), 170–176.

Agency for Healthcare Research. (2008). *Outcomes of Maternal Weight Gain* (No. 168) (pp. 1–1038). North Carolina: RTI International-University of California. Retrieved from http://www.ahrq.gov/research/findings/evidence-based-reports/admat-evidence-report.pdf papers2://publication/uuid/EBD7D0A3-D72C-459F-8974-68998D658ACD

Allison, D. B., Paultre, F., Goran, M. I., Poehlman, E. T., & Heymsfield, S. B. (1995). Statistical considerations regarding the use of ratios to adjust data. *International Journal of Obesity and Related Metabolic Disorders: Journal of the International Association for the Study of Obesity*, *19*(9), 644–652.

Au, C. P., Raynes-Greenow, C. H., Turner, R. M., Carberry, A. E., & Jeffery, H. (2013). Fetal and maternal factors associated with neonatal adiposity as measured by air displacement plethysmography: a large cross-sectional study. *Early Human Development*, *89*(10), 839–843. http://doi.org/10.1016/j.earlhumdev.2013.07.028

Barker, D. J. P., Osmond, C., Forsen, T. J., Kajantie, E., & Eriksson, J. G. (2005). Trajectories of growth among children who have coronary events as adults. *The New England Journal of Medicine*, *353*(17), 1802–1809. http://doi.org/10.1056/NEJMoa044160

Bettiol, H. (2003). Neonatal anthropometry and neonatal outcome. *São Paulo Medical Journal = Revista Paulista De Medicina*, *121*(4), 147–148.

Black, R. E., Allen, L. H., Bhutta, Z. A., Caulfield, L. E., de Onis, M., Ezzati, M., et al. (2008). Maternal and child undernutrition: global and regional exposures and health consequences. *The Lancet*, *371*(9608), 243–260. http://doi.org/10.1016/S0140-6736(07)61690-0

Black, R., Victora, C., Walker, S., Bhutta, Z., Christian, P., de Onis, M. et al. (2013). Series Maternal and Child Nutrition 1 Maternal and child undernutrition and overweight in low-income and middle-income countries. *The Lancet*, *382*(9890), 427–451. http://doi.org/10.1016/S0140-6736(13)60937-X

Bodnar, L. M., Catov, J. M., Klebanoff, M. A., Ness, R. B., & Roberts, J. M. (2007). Pre-pregnancy body mass index and the occurrence of severe hypertensive disorders of pregnancy. *Epidemiology*, *18*(2), 234–239. http://doi.org/10.1097/01.ede.0000254119.99660.e7

Bodnar, L. M., Hutcheon, J. A., Parisi, S. M., Pugh, S. J., & Abrams, B. (2014). Comparison of Gestational Weight Gain z-Scores and Traditional Weight Gain Measures in Relation to Perinatal Outcomes. *Paediatric and Perinatal Epidemiology*, *29*(1), 11–21. http://doi.org/10.1111/ppe.12168

Boodnha, G. (2014, December 5). Children's body mass index, overweight and obesity. Retrieved October 5, 2017, from http://content.digital.nhs.uk/catalogue/PUB16076/HSE2013-Ch11-Child-BMI.pdf

Calvo, E. B., López, L. B., Balmaceda, Y. D. V., Poy, M. S., González, C., Quintana, L., et al. (2009). Reference charts for weight gain and body mass index during pregnancy obtained from a healthy cohort. *Journal of Maternal-Fetal and Neonatal Medicine*, *22*(1), 36–42.

http://doi.org/10.1080/14767050802464502
Carlsen, E. M., Renault, K. M., Nørgaard, K., Nilas, L., Jensen, J., Hyldstrup, L., et al. (2014). Newborn regional body composition is influenced by maternal obesity, gestational weight gain and the birthweight standard score. *Acta Paediatrica*, *103*(9), 939–945. http://doi.org/10.1111/apa.12713
Carmichael, S. S., Abrams, B. B., & Selvin, S. S. (1997). The pattern of maternal weight gain in women with good pregnancy outcomes. *Am J Public Health*, *87*(12), 1984–1988. http://doi.org/10.2105/AJPH.87.12.1984
Catalano, P. M., Drago, N. M., & Amini, S. B. (1995). Factors affecting fetal growth and body composition. *Ymob*, *172*(5), 1459–1463.
Catalano, P. M., Farrell, K., Thomas, A., Huston-Presley, L., Mencin, P., de Mouzon, S. H., & Amini, S. B. (2009). Perinatal risk factors for childhood obesity and metabolic dysregulation. *The American Journal of Clinical Nutrition*, *90*(5), 1303–1313. http://doi.org/10.3945/ajcn.2008.27416
Catalano, P. M., Thomas, A., Huston-Presley, L., & Amini, S. B. (2003). Increased fetal adiposity: A very sensitive marker of abnormal in utero development. *American Journal of Obstetrics and Gynecology*, *189*(6), 1698–1704. http://doi.org/10.1016/S0002-9378(03)00828-7
Catalano, P., Mele, L., Landon, M., Ramin, S., Reddy, U., Casey, B., et al. Inadequate weight gain in overweight and obese pregnant women: what is the effect on fetal growth? *Ymob*, *211*(2), 137.e1–137.e7. http://doi.org/10.1016/j.ajog.2014.02.004
Cheikh Ismail, L., Bishop, D. C., Pang, R., Ohuma, E. O., Kac, G., Abrams, B., et al. (2016). Gestational weight gain standards based on women enrolled in the Fetal Growth Longitudinal Study of the INTERGROWTH-21 stProject: a prospective longitudinal cohort study. *Bmj*, i555–10. http://doi.org/10.1136/bmj.i555
Cheikh Ismail, L., Knight, H. E., Bhutta, Z., Chumlea, W. C., for the International Fetal and Newborn Growth Consortium for the 21st Century (INTERGROWTH-21st). (2013a). Anthropometric protocols for the construction of new international fetal and newborn growth standards: the INTERGROWTH-21 stProject. *BJOG: an International Journal of Obstetrics & Gynaecology*, *120*, 42–47. http://doi.org/10.1111/1471-0528.12125
Cheikh Ismail, L., Knight, H. E., Ohuma, E. O., Hoch, L., Chumlea, W. C., for the International Fetal and Newborn Growth Consortium for the 21st Century (INTERGROWTH-21st). (2013b). Anthropometric standardisation and quality control protocols for the construction of new, international, fetal and newborn growth standards: the INTERGROWTH-21 stProject. *BJOG: an International Journal of Obstetrics & Gynaecology*, *120*, 48–55. http://doi.org/10.1111/1471-0528.12127
Crane, J. M. G., White, J., Murphy, P., Burrage, L., & Hutchens, D. (2009). The effect of gestational weight gain by body mass index on maternal and neonatal outcomes. *Journal of Obstetrics and Gynaecology Canada : JOGC = Journal D'obstetrique Et Gynecologie Du Canada : JOGC*, *31*(1), 28–35.
Crozier, S. R., Inskip, H. M., Godfrey, K. M., Cooper, C., Harvey, N. C., Cole, Z. A., et al. (2010). Weight gain in pregnancy and childhood body

composition: findings from the Southampton Women's Survey. *The American Journal of Clinical Nutrition*, *91*(6), 1745–1751. http://doi.org/10.3945/ajcn.2009.29128
Davenport, M. H., Ruchat, S.-M., Giroux, I., Sopper, M. M., & Mottola, M. F. (2013). Timing of Excessive Pregnancy-Related Weight Gain and Offspring Adiposity at Birth. *Obstetrics & Gynecology*, *122*(2, PART 1), 255–261. http://doi.org/10.1097/AOG.0b013e31829a3b86
Dawes, M. G., & Grudzinskas, J. G. (1991). Patterns of maternal weight gain in pregnancy. *British Journal of Obstetrics and Gynaecology*, *98*(2), 195–201.
de Onis, M., & Lobstein, T. (2010). Defining obesity risk status in the general childhood population: which cut-offs should we use? *International Journal of Pediatric Obesity : IJPO : an Official Journal of the International Association for the Study of Obesity*, *5*(6), 458–460. http://doi.org/10.3109/17477161003615583
de Onis, M., Onyango, A. W., Van den Broeck, J., Chumlea, W. C., & Martorell, R. (2004). Measurement and standardization protocols for anthropometry used in the construction of a new international growth reference. *Food and Nutrition Bulletin*, *25*(1 Suppl), S27–S36.
Dewey, K. G., & Begum, K. (2011). Long-term consequences of stunting in early life. *Maternal and Child Nutrition*, *7*(1), 5–18. http://doi.org/10.1111/j.1740-8709.2011.00349.x
Diesel, J. C., Eckhardt, C. L., Day, N. L., Brooks, M. M., Arslanian, S. A., & Bodnar, L. M. (2015). Gestational weight gain and the risk of offspring obesity at 10 and 16 years: a prospective cohort study in low-income women. *BJOG: an International Journal of Obstetrics & Gynaecology*, *122*(10), 1395–1402. http://doi.org/10.1111/1471-0528.13448
Edwards, L. E., Hellerstedt, W. L., Alton, I. R., story, M., & Himes, J. H. (1996). Pregnancy complications and birth outcomes in obese and normal-weight women: effects of gestational weight change. *Obstetrics & Gynecology*, *87*(3), 389–394.
Ehrenberg, H. M., Dierker, L., Milluzzi, C., & Mercer, B. M. (2003). Low maternal weight, failure to thrive in pregnancy, and adverse pregnancy outcomes. *American Journal of Obstetrics and Gynecology*, *189*(6), 1726–1730. http://doi.org/10.1016/S0002-9378(03)00860-3
Ellis, K. J., Yao, M., Shypailo, R. J., Urlando, A., Wong, W. W., & Heird, W. C. (2007). Body-composition assessment in infancy: air-displacement plethysmography compared with a reference 4-compartment model. *The American Journal of Clinical Nutrition*, *85*(1), 90–95.
Eriksson, B., Löf, M., & Forsum, E. (2010). Body composition in full-term healthy infants measured with air displacement plethysmography at 1 and 12 weeks of age. *Acta Paediatrica*, *99*(4), 563–568. http://doi.org/10.1111/j.1651-2227.2009.01665.x
Fattah, C., Farah, N., Barry, S. C., O'Connor, N., Stuart, B., & Turner, M. J. (2010). Maternal weight and body composition in the first trimester of pregnancy. *Acta Obstetricia Et Gynecologica Scandinavica*, *89*(7), 952–955. http://doi.org/10.3109/00016341003801706
Fraser, A., Tilling, K., Macdonald-Wallis, C., Hughes, R., Sattar, N., Nelson, S. M., & Lawlor, D. A. (2011). Associations of gestational weight gain with maternal body mass index, waist circumference, and blood pressure

measured 16 y after pregnancy: the Avon Longitudinal Study of Parents and Children (ALSPAC). *American Journal of Clinical Nutrition*, *93*(6), 1285–1292. http://doi.org/10.3945/ajcn.110.008326

Gluckman, P. D., Hanson, M. A., Cooper, C., & Thornburg, K. L. (2008). Effect of in utero and early-life conditions on adult health and disease. *The New England Journal of Medicine*, *359*(1), 61–73. http://doi.org/10.1056/NEJMra0708473

Goldstein, R. F., Abell, S. K., Ranasinha, S., Misso, M., Boyle, J. A., Black, M. H., et al. (2017). Association of Gestational Weight Gain With Maternal and Infant Outcomes. *Jama*, *317*(21), 2207–19. http://doi.org/10.1001/jama.2017.3635

Grandi, C., & Luchtenberg, G. (2007). Evaluacion nutricional durante el embarazo. Nuevo estandar. *Medicina*, *67*, 677–684.

Guelinckx, I., Beckers, K., Vansant, G., & Devlieger, R. (2010). Construction of Weight Gain Charts in a Low-Risk Obstetric Belgian Population. *Gynecologic and Obstetric Investigation*, *69*(1), 57–61. http://doi.org/10.1159/000255762

Hawkes, C. P., Hourihane, J. O., Kenny, L. C., Irvine, A. D., Kiely, M., & Murray, D. M. (2011). Gender- and Gestational Age-Specific Body Fat Percentage at Birth. *Pediatrics*, 1–9. http://doi.org/10.1542/peds.2010-3856

HEALTH AND SOCIAL CARE INFORMATION CETRE. (2016, March 17). Statistics on Women's Smoking at the Time of Delivery: England. Retrieved October 5, 2017, from http://content.digital.nhs.uk/catalogue/PUB20219/stat-wome-smok-time-deli-eng-q3-15-16-rep.pdf

Hull, H. R., Dinger, M. K., Knehans, A. W., Thompson, D. M., & Fields, D. A. (2008). Impact of maternal body mass index on neonate birthweight and body composition. *American Journal of Obstetrics and Gynecology*, *198*(4), 416.e1–416.e6. http://doi.org/10.1016/j.ajog.2007.10.796

Hull, H. R., Thornton, J. C., Ji, Y., Paley, C., Rosenn, B., Mathews, P., et al. (2011). Higher infant body fat with excessive gestational weight gain in overweight women. *Ymob*, *205*(3), 211.e1–211.e7. http://doi.org/10.1016/j.ajog.2011.04.004

Hutcheon, J. A., Platt, R. W., Abrams, B., Himes, K. P., Simhan, H. N., & Bodnar, L. M. (2013). A weight-gain-for-gestational-age z score chart for the assessment of maternal weight gain in pregnancy. *American Journal of Clinical Nutrition*, *97*(5), 1062–1067. http://doi.org/10.3945/ajcn.112.051706

IAEA. (2013). *Body Composition Assessment from Birth to Two Years of age* (p. 62). International Atomic Energy Agency.

Ioannou, C., Talbot, K., Ohuma, E., Sarris, I., Villar, J., Conde-Agudelo, A., & Papageorghiou, A. T. (2012). Systematic review of methodology used in ultrasound studies aimed at creating charts of fetal size. *British Journal of Obstetrics and Gynaecology*, *119*(12), 1425–1439. http://doi.org/10.1111/j.1471-0528.2012.03451.x

Interbio 21st Consortium Study protocols. (2012, October 19). The INTERBIO-21st Study. Retrieved October 5, 2017, from http://www.medscinet.net/Interbio/Uploads/ProtocolDocs/INTERBIO-21st%20Study%20Protocol%20October%202012%20(SK%20revised)[4].

pdf
Johansson, K., Hutcheon, J. A., Stephansson, O., & Cnattingius, S. (2016). Pregnancy weight gain by gestational age and BMI in Sweden: a population-based cohort study. *The American Journal of Clinical Nutrition*, *103*(5), 1278–1284. http://doi.org/10.3945/ajcn.115.110197
Josefson, J. L., Hoffmann, J. A., & Metzger, B. E. (2013). Excessive weight gain in women with a normal pre-pregnancy BMI is associated with increased neonatal adiposity. *Pediatric Obesity*, *8*(2), e33–6. http://doi.org/10.1111/j.2047-6310.2012.00132.x
Joseph, K. (2003). Changes in maternal characteristics and obstetric practice and recent increases in primary cesarean delivery. *Obstetrics & Gynecology*, *102*(4), 791–800. http://doi.org/10.1016/S0029-7844(03)00620-3
Kabiru, W., & Denise Raynor, B. (2004). Obstetric outcomes associated with increase in BMI category during pregnancy. *American Journal of Obstetrics and Gynecology*, *191*(3), 928–932. http://doi.org/10.1016/j.ajog.2004.06.051
Kirkwood, B., & Sterne, J. (2003). Essential Medical Statistics. Wiley-Blackwell.
Lee, W., Balasubramaniam, M., Deter, R. L., Hassan, S. S., Gotsch, F., Kusanovic, J. P., et al. (2009). Fetal growth parameters and birth weight: their relationship to neonatal body composition. *Ultrasound in Obstetrics and Gynecology*, *33*(4), 441–446. http://doi.org/10.1002/uog.6317
Luke, B., Hediger, M. L., & Scholl, T. O. (1996). Point of diminishing returns: when does gestational weight gain cease benefiting birthweight and begin adding to maternal obesity? *The Journal of Maternal-Fetal Medicine*, *5*(4), 168–173. http://doi.org/10.1002/(SICI)1520-6661(199607/08)5:4<168::AID-MFM2>3.0.CO;2-I
Ma, G., Yao, M., Liu, Y., Lin, A., Zou, H., Urlando, A., et al. (2004). Validation of a new pediatric air-displacement plethysmograph for assessing body composition in infants. *The American Journal of Clinical Nutrition*, *79*(4), 653–660.
Mardonnes, F., & Rosso, P. (1997). Design of a weight gain chart for pregnant women. *Revista Médica De Chile*, *125*(12), 1437–1448.
Medicine, Institute of. (2009). *WEIGHT GAIN DURING PREGNANCY. IOM PREGNANCY WEIGHT GUIDELINES* (pp. 1–868). Washington DC: National academies press.
Moher, D., Liberati, A., Tetzlaff, J., Altman, D. G., The PRISMA Group. (2009). Preferred Reporting Items for Systematic Reviews and Meta-Analyses: The PRISMA Statement. *PLoS Medicine*, *6*(7), e1000097. http://doi.org/10.1371/journal.pmed.1000097.s002
Müller, M. J., Braun, W., Pourhassan, M., Geisler, C., & Bosy-Westphal, A. (2016). Application of standards and models in body composition analysis. *The Proceedings of the Nutrition Society*, *75*(2), 181–187. http://doi.org/10.1017/S0029665115004206
Napolitano, R., Dhami, J., Ohuma, E. O., Ioannou, C., Conde-Agudelo, A., Kennedy, S. H., et al. (2014). Pregnancy dating by fetal crown-rump length: a systematic review of charts. *British Journal of Obstetrics and Gynaecology*, 556–565. http://doi.org/10.1111/1471-0528.12478
National Collaborating Centre for Women's and Children's Health (UK).

(2015). Diabetes in Pregnancy: Management of Diabetes and Its Complications from Preconception to the Postnatal Period.
Nohr, E. A., Vaeth, M., Baker, J. L., Sorensen, T. I., Olsen, J., & Rasmussen, K. M. (2009). Pregnancy outcomes related to gestational weight gain in women defined by their body mass index, parity, height, and smoking status. *American Journal of Clinical Nutrition*, *90*(5), 1288–1294. http://doi.org/10.3945/ajcn.2009.27919
Ochsenbein-Kölble, N., Roos, M., Gasser, T., & Zimmermann, R. (2007). Cross-sectional study of weight gain and increase in BMI throughout pregnancy. *European Journal of Obstetrics & Gynecology and Reproductive Biology*, *130*(2), 180–186. http://doi.org/10.1016/j.ejogrb.2006.03.024
Ohadike, C. O., Cheikh Ismail, L., Ohuma, E. O., Giuliani, F., Bishop, D., Kac, G., et al. (2016). Systematic Review of the Methodological Quality of Studies Aimed at Creating Gestational Weight Gain Charts. *Advances in Nutrition: an International Review Journal*, *7*(2), 313–322. http://doi.org/10.3945/an.115.010413
Oken, E., & Gillman, M. W. (2003). Fetal origins of obesity. *Obesity Research*, *11*(4), 496–506. http://doi.org/10.1038/oby.2003.69
Oken, E., Taveras, E. M., Kleinman, K. P., Rich-Edwards, J. W., & Gillman, M. W. (2007). Gestational weight gain and child adiposity at age 3 years. *American Journal of Obstetrics and Gynecology*, *196*(4), 322.e1–322.e8. http://doi.org/10.1016/j.ajog.2006.11.027
Osmond, C., & Barker, D. J. (2000). Fetal, infant, and childhood growth are predictors of coronary heart disease, diabetes, and hypertension in adult men and women. *Environmental Health Perspectives*, *108 Suppl 3*(Suppl 3), 545–553.
Poston, L. (2012). Gestational weight gain: influences on the long-term health of the child. *Current Opinion in Clinical Nutrition and Metabolic Care*, *15*(3), 252–257. http://doi.org/10.1097/MCO.0b013e3283527cf2
Public Health England, (2016, November 1). National Child Measurement Programme, England, 2015/16 school year. Retrieved October 5, 2017, from http://www.content.digital.nhs.uk/catalogue/PUB22269/nati-chil-meas-prog-eng-2015-2016-rep.pdf
Rasmussen, K. M., Abrams, B., Bodnar, L. M., Butte, N. F., Catalano, P. M., & Maria Siega-Riz, A. (2010). Recommendations for weight gain during pregnancy in the context of the obesity epidemic. *Obstetrics & Gynecology*, *116*(5), 1191–1195. http://doi.org/10.1097/AOG.0b013e3181f60da7
Roggero, P., Giannì, M. L., Amato, O., Orsi, A., Piemontese, P., Morlacchi, L., & Mosca, F. (2009). Is term newborn body composition being achieved postnatally in preterm infants? *Early Human Development*, *85*(6), 349–352. http://doi.org/10.1016/j.earlhumdev.2008.12.011
Rosenberg, T. J., Garbers, S., Lipkind, H., & Chiasson, M. A. (2005). Maternal obesity and diabetes as risk factors for adverse pregnancy outcomes: differences among 4 racial/ethnic groups. *Am J Public Health*, *95*(9), 1545–1551. http://doi.org/10.2105/AJPH.2005.065680
Rosso, P. (1985). A new chart to monitor weight gain during pregnancy. *American Journal of Clinical Nutrition*, (41), 644–652. Retrieved from http://ezproxy.ouls.ox.ac.uk:2727/content/41/3/644.long

Russell, A., Gillespie, S., Satya, S., & Gaudet, L. M. (2013). Assessing the accuracy of pregnant women in recalling pre-pregnancy weight and gestational weight gain. *Journal of Obstetrics and Gynaecology Canada : JOGC = Journal D'obstetrique Et Gynecologie Du Canada : JOGC*, *35*(9), 802–809.

Russo, Dello, M., Ahrens, W., De Vriendt, T., Marild, S., Molnar, D., Moreno, L. A., et al. (2013). Gestational weight gain and adiposity, fat distribution, metabolic profile, and blood pressure in offspring: the IDEFICS project, *37*(7), 914–919. http://doi.org/10.1038/ijo.2013.35

Saldana, T. M., Siega-Riz, A. M., Adair, L. S., & Suchindran, C. (2006). The relationship between pregnancy weight gain and glucose tolerance status among black and white women in central North Carolina. *American Journal of Obstetrics and Gynecology*, *195*(6), 1629–1635. http://doi.org/10.1016/j.ajog.2006.05.017

Sayer, A. A., Syddall, H. E., Dennison, E. M., Gilbody, H. J., Duggleby, S. L., Cooper, C., et al. (2004). Birth weight, weight at 1 y of age, and body composition in older men: findings from the Hertfordshire Cohort Study. *The American Journal of Clinical Nutrition*, *80*(1), 199–203.

Rodrigues, P. de Olivieira, I., Brito, A., Kac, G. (2010). Determinant factors of insufficient and excessive gestational weight gain and maternal-child adverse outcomes. *Nutrition*, *26*(6), 617–623. http://doi.org/10.1016/j.nut.2009.06.025

Schack-Nielsen, L., Michaelsen, K. F., Gamborg, M., Mortensen, E. L., & Rensen, T. I. A. S. O. (2009). Gestational weight gain in relation to offspring body mass index and obesity from infancy through adulthood. *International Journal of Obesity*, *34*(1), 67–74. http://doi.org/10.1038/ijo.2009.206

Scholl, T. O., Hediger, M. L., & Schall, J. I. (1995). Gestational weight gain, pregnancy outcome and postpartum weight retention. *Obstetrics & Gynecology*, *86*(3), 423–427.

Sewell, M. F., Huston-Presley, L., Super, D. M., & Catalano, P. (2006). Increased neonatal fat mass, not lean body mass, is associated with maternal obesity. *American Journal of Obstetrics and Gynecology*, *195*(4), 1100–1103. http://doi.org/10.1016/j.ajog.2006.06.014

Sherrard, A., Platt, R. W., Vallerand, D., Usher, R. H., Zhang, X., & Kramer, M. S. (2007). Maternal anthropometric risk factors for caesarean delivery before or after onset of labour. *British Journal of Obstetrics and Gynaecology*, *114*(9), 1088–1096. http://doi.org/10.1111/j.1471-0528.2007.01275.x

Sibai, B. M., Ewell, M., Levine, R. J., Klebanoff, M. A., Esterlitz, J., Catalano, P. M., et al. (1997). Risk factors associated with preeclampsia in healthy nulliparous women. *American Journal of Obstetrics and Gynecology*, *177*(5), 1003–1010. http://doi.org/10.1016/S0002-9378(97)70004-8

Simon, L., Borrego, P., Darmaun, D., Legrand, A., Rozé, J.-C., & Chauty-Frondas, A. (2013). Effect of sex and gestational age on neonatal body composition. *British Journal of Nutrition*, *109*(6), 1105–1108. http://doi.org/10.1017/S0007114512002991

Starling, A. P., Brinton, J. T., Glueck, D. H., Shapiro, A. L., Harrod, C. S., Lynch, A. M., et al. (2015). Associations of maternal BMI and gestational weight gain with neonatal adiposity in the Healthy Start study. *The*

American Journal of Clinical Nutrition, *101*(2), 302–309. http://doi.org/10.3945/ajcn.114.094946
Straube, S., Voigt, M., Briese, V., Schneider, K. T. M., & Voigt, M. (2008). Weight gain in pregnancy according to maternal height and weight. *Journal of Perinatal Medicine*, *36*(5), 405–412. http://doi.org/10.1515/JPM.2008.073
Thadhani, R., Stampfer, M. J., Hunter, D. J., Manson, J. E., Solomon, C. G., & Curhan, G. C. (1999). High body mass index and hypercholesterolemia: risk of hypertensive disorders of pregnancy. *Obstetrics & Gynecology*, *94*(4), 543–550. http://doi.org/10.1016/S0029-7844(99)00400-7
Theron, G. B., & Thompson, M. L. (1990). A centile chart for weight gain in pregnancy for the urban population of the Western Cape, South Africa. *International Journal of Gynecology and Obstetrics*, *33*(2), 127–133. http://doi.org/10.1016/0020-7292(90)90585-9
Urlando, A., Dempster, P., & Aitkens, S. (2003). A New Air Displacement Plethysmograph for the Measurement of Body Composition in Infants. *Pediatric Research*, *53*(3), 486–492. http://doi.org/10.1203/01.PDR.0000049669.74793.E3
Victora, C. G., Villar, J., Barros, F. C., Ismail, L. C., Chumlea, C., Papageorghiou, A. T., et al. (2015). Anthropometric Characterization of Impaired Fetal Growth. *JAMA Pediatrics*, *169*(7), e151431–10. http://doi.org/10.1001/jamapediatrics.2015.1431
Vila-Candel, R., & Hevilla-Cucarella, E. (2009). Curva de ganancia ponderal de la gestante de bajo riesgo y su relación con el peso del recién nacido. *Matronas Profesión*, *10*(4), 6–13.
Villar, J., Altman, D. G., Purwar, M., Noble, J. A., Knight, H. E., Ruyan, P., et al. (2013a). The objectives, design and implementation of the INTERGROWTH-21 stProject. *British Journal of Obstetrics and Gynaecology*, *120*, 9–26. http://doi.org/10.1111/1471-0528.12047
Villar, J., Puglia, F. A., Fenton, T. R., Ismail, L. C., Staines-Urias, E., Giuliani, F., et al. (2017). Body composition at birth and its relationship with neonatal anthropometric ratios: the newborn body composition study of the INTERGROWTH-21st project. *Nature Publishing Group*, *82*(2), 305–316. http://doi.org/10.1038/pr.2017.52
Waters, T. P., Huston-Presley, L., & Catalano, P. M. (2012). Neonatal Body Composition According to the Revised Institute of Medicine Recommendations for Maternal Weight Gain. *The Journal of Clinical Endocrinology & Metabolism*, *97*(10), 3648–3654. http://doi.org/10.1210/jc.2012-1781
Widen, E. M., Whyatt, R. M., Hoepner, L. A., Mueller, N. T., Ramirez-Carvey, J., Oberfield, S. E., et al. (2015). Gestational weight gain and obesity, adiposity and body size in African-American and Dominican children in the Bronx and Northern Manhattan. *Maternal and Child Nutrition*, *12*(4), 918–928. http://doi.org/10.1111/mcn.12174
Withers, R. T., LaForgia, J., Pillans, R. K., Shipp, N. J., Chatterton, B. E., Schultz, C. G., & Leaney, F. (1998). Comparisons of two-, three-, and four-compartment models of body composition analysis in men and women. *Journal of Applied Physiology (Bethesda, Md. : 1985)*, *85*(1), 238–245.
Witter, F., Caulfield, L. E., & Stoltzfus, R. J. (1995). Influence of maternal

anthropometric status and birth weight on the risk of cesarean delivery. *Obstetrics & Gynecology*, *85*(6), 947–951.
Wong, W., Tang, N. L. S., Lau, T. K., & Wong, T. W. (2000). A New Recommendation for Maternal Weight Gain in Chinese Women. *Journal of the American Dietetic Association*, *100*(7), 791–796. http://doi.org/10.1016/S0002-8223(00)00230-3
World health Organisation, (1995). *Physical status: the use and interpretation of anthropometry.* (No. 854). *American Journal of Human Biology*. Geneva: WHO Library Cataloguing in Publication Data.
World Health Organisation. (2010). Gender, Women and the Tobacco Epidemic. Retrieved October 5, 2017, from http://apps.who.int/iris/bitstream/10665/44342/1/9789241599511_eng.pdf
Xu, J., Luntamo, M., Kulmala, T., Ashorn, P., & Cheung, Y. B. (2014). A longitudinal study of weight gain in pregnancy in Malawi: unconditional and conditional standards. *American Journal of Clinical Nutrition*, *99*(2), 296–301. http://doi.org/10.3945/ajcn.113.074120
Yao, M., Nommsen-Rivers, L., Dewey, K., & Urlando, A. (2003). Preliminary evaluation of a new pediatric air displacement plethysmograph for body composition assessment in infants. *Acta Diabetologica*, *40*(0), s55–s58. http://doi.org/10.1007/s00592-003-0027-9
Zilko, C. E. M., Rehkopf, D., & Abrams, B. (2010). Association of maternal gestational weight gain with short- and long-term maternal and child health outcomes. *AJOG*, *202*(6), 574.e1–574.e8. http://doi.org/10.1016/j.ajog.2009.12.007

APPENDICES

Appendices

Appendix 1

The PEAPOD® device is a mobile apparatus that comprises of a test chamber, electronic scale, reference chamber, calibration volume, electronic components, monitor, printer and CPU. There is a calibration valve between the test chamber and calibration chamber. In addition, there is an oscillating diaphragm and pressure transducers between the reference and test chambers (Urlando, Dempster, & Aitkens, 2003).

The test chamber has a moveable tray and a door held closed by an electromagnet. The relationships between pressure and volume expressed by Boyles Law and Poissons Law are the foundation principles on which the PEAPOD is based. Boyles Law describes the behaviour of air compressed under isothermal conditions i.e. at constant temperature $P_1/P_2=V_2/V_1$. When air is allowed to change in temperature in response to pressure changes Poissons Law describes its behaviour: $P_1/P_2= (V_2/V_1)^{\gamma}$ (for air, $\gamma = 1.4$). The reason the distinction between isothermal and adiabatic air is important is because of the adjustments that PEAPOD has to make to calculate body volume (Urlando, Dempster, & Aitkens, 2003). For a fixed reference chamber volume, the pressure changes caused by an oscillating diaphragm between the two chambers have a linear relationship with the test chamber volume. This is assuming adiabatic conditions. However, there are isothermal areas of air near the infants skin and within the lungs that need to be corrected for otherwise volumes would be overestimated by 40%. Once the volumes are obtained, whole body density, FFM and BF% can be calculated using

specified equations for infants (Urlando, 2003). The oscillating diaphragm causes sinusoidal volume changes that are equal in size but opposite in direction. The pressure changes are used to calculate the test chamber volumes. The subject volume is then calculated by subtracting the test chamber volume when empty and with the subject inside it. In order to ensure that the air behaves isothermally, the infant must be nude and hair flattened against the head using baby oil (Yao et al., 2003). The pressure changes are not detectable by the infant and amount to the equivalent of climbing up one flight of stairs. The potential risk of build up of carbon dioxide is avoided by not allowing the chamber to be sealed more than a few minutes and an alarm that sounds if carbon dioxide levels are high (IAEA, 2013).

The recommended protocol for measuring ADP is as follows:

1) Infant data: gender, length and age is recorded
2) Weight is measured on the PEAPOD scale
3) Automatic volume calibration is performed in the empty test volume chamber
4) After the 2 minute calibration test, the test chamber automatically opens
5) The infant is placed on the test chamber tray and moved into the chamber
6) Volume measurements are then performed 3 times
7) After these measurements the test chamber automatically opens. The infant is removed from the tray and the whole process is repeated with the chamber empty

8) Infant volume is calculated as the average of the three volume measurements (Yao et al., 2003)

Figure (i): PEAPOD® machine

Figure (ii): Neonate being weighed on PEAPOD® machine

Figure (iii): Neonate about to be put into PEAPOD® chamber

Appendix 2

Measurement protocols

Measurement of newborn body composition

The procedure for measurement of each neonate was as follows:

1) general information was entered including gender, date of birth, body length
2) Each neonate was weighed on the PEAPOD in-built electronic scales;
3) Each neonate was placed headfirst in a supine position into the test chamber and the operator then shut the chamber door
4) The PEAPOD then automatically started the measuring protocol.

Measurement of maternal weight and BMI

Women were weighed on a Seca 877 scale, and their height measured using a Seca stadiometer 242 or 264 models. The equipment was calibrated at least twice a week with strict calibration protocols. The Seca BMI calculator 491 was used to calculate the BMI of women at the screening phase.

The following steps were taken when measuring participants' height:

1) Shoes and hair ornaments were removed; Thick braids and corn/rows were measured separately
2) The participant stood with the back to the measuring rod, feet slightly apart, trunk balanced over the waist, knees straight and shoulders relaxed
3) The head was positioned
4) The participant was advised to take a deep breath and stand tall
5) The head board was brought in contact with the top of the head
6) The height on the display was recorded

The following steps were taken when measuring participants' weight:

1) The participant was wearing light and minimal clothing
2) Shoes and heavy objects were removed
3) The scale was zeroed after checking the surface was flat
4) The participant was advised to stand centrally on the scale with the feet in the footmarks

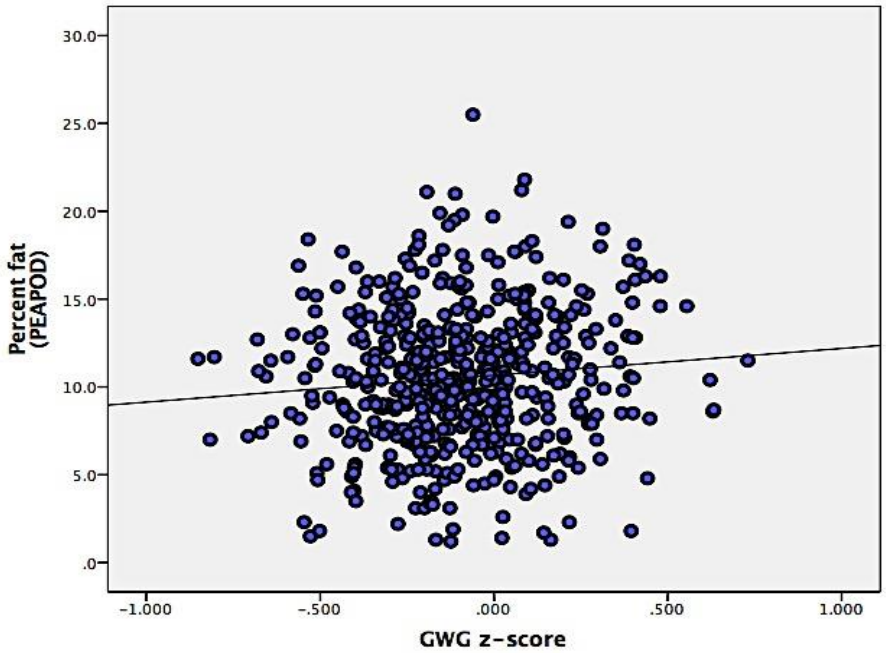
   The weight was recorded

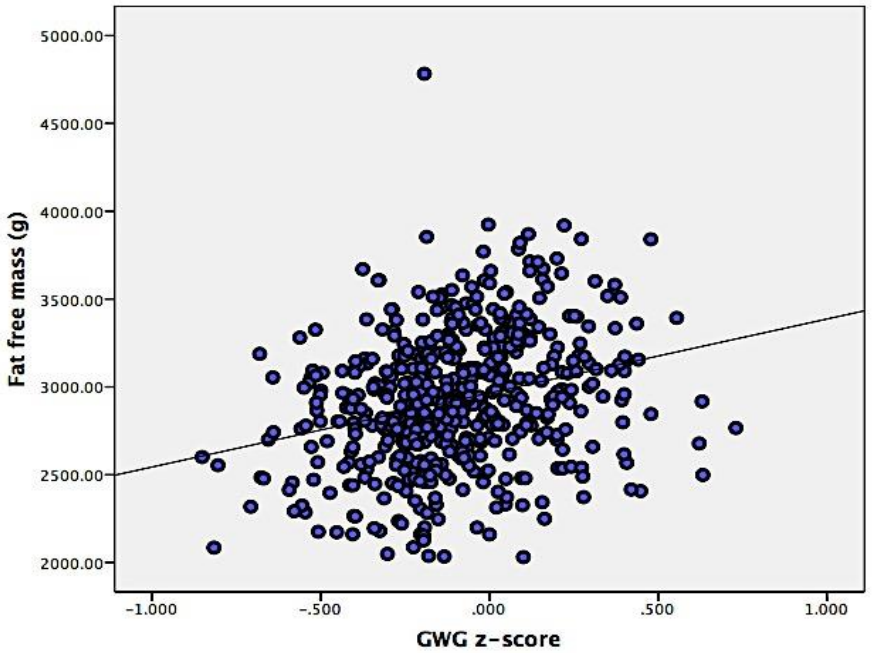
Appendix 3 – Search terms and Citations

Search terms	Citations obtained
MEDLINE; EMBASE	
Pregnancy kw.+ body weight af.	2666
Maternal weight kw.	142
Maternal weight af. + chart af.	63
Pregnancy kw. + gestational weight gain kw.	281
Pregnancy kw. + weight gain af. +chart af.	34
Pregnancy kw. +weight gain af. + curve af.	36
Pregn* af. + weight af. + chart* af.	1994
Pregnancy kw. + gestational weight gain af.	757
Pregnancy af. + gestational af. + body weight change af.	18
Pregnancy kw. + body weight af. + reference af.	984
Pregn* af. or gest* af. + reference af. + weight af. + trend af.	181
Pregnancy kw. + range af. + weight af. + body af.	310
WEB OF SCIENCE	
Weight gain + pregn* +chart	249
Gest* weight gain + women+ chart	143
INCLUDED AFTER SCREENING TITLES	3176

Appendix 4

Scatterplots showing association between GWG z-score and newborn body composition





Stepwise regression analyses

Model Summary[b]

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate	Durbin-Watson
1	.187[a]	.035	.030	3.8929	1.974

a. Predictors: (Constant), Site2, zscore, Site1
b. Dependent Variable: Percent fat

(PEAPOD)

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.	95.0% Confidence Interval for B	
		B	Std. Error	Beta			Lower Bound	Upper Bound
1	(Constant)	10.812	.199		54.364	.000	10.421	11.202
	zscore	1.240	.698	.078	1.776	.076	-.132	2.612
	Site1	-1.484	.463	-.141	-3.208	.001	-2.393	-.575
	Site2	1.001	.667	.065	1.502	.134	-.309	2.311

Model Summary[b]

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate	Durbin-Watson
1	.221[a]	.049	.038	3.8759	2.002

a. Predictors: (Constant), obese, Underweight, Site2, zscore, overweight, Site1
b. Dependent Variable: Percent fat(PEAPOD)

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.	95.0% Confidence Interval for B	
		B	Std. Error	Beta			Lower Bound	Upper Bound
1	(Constant)	10.548	.246		42.911	.000	10.066	11.031
	zscore	1.411	.704	.088	2.006	.045	.029	2.794
	Site1	-1.101	.483	-.104	-2.277	.023	-2.050	-.151
	Site2	.838	.669	.055	1.253	.211	-.476	2.153
	Underweight	-1.221	.986	-.054	-1.238	.216	-3.158	.716
	overweight	.488	.401	.054	1.218	.224	-.299	1.276
	obese	1.316	.576	.102	2.285	.023	.185	2.448

Model Summary[b]

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate	Durbin-Watson
1	.329[a]	.108	.096	3.7565	2.000

a. Predictors: (Constant), Newborn sex, obese, Underweight, Site2, zscore, overweight, Site1

b. Dependent Variable: Percent fat(PEAPOD)

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.	95.0% Confidence Interval for B	
		B	Std. Error	Beta			Lower Bound	Upper Bound
1	(Constant)	11.561	.293		39.426	.000	10.985	12.137
	zscore	1.823	.685	.114	2.659	.008	.476	3.169
	Site1	-.996	.469	-.094	-2.125	.034	-1.917	-.075
	Site2	1.008	.649	.066	1.554	.121	-.267	2.283
	Underweight	-1.435	.956	-.064	-1.500	.134	-3.314	.444
	overweight	.520	.389	.058	1.337	.182	-.244	1.283
	obese	1.363	.558	.106	2.441	.015	.266	2.460
	Newborn sex	-1.938	.327	-.245	-5.923	.000	-2.581	-1.295

Model Summary[b]

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate	Durbin-Watson
1	.354[a]	.125	.112	3.7243	1.994

a. Predictors: (Constant), GA at delivery, Site1, Newborn sex, obese, Site2, Underweight, zscore, overweight

b. Dependent Variable: Percent fat(PEAPOD)

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.	95.0% Confidence Interval for B	
		B	Std. Error	Beta			Lower Bound	Upper Bound
1	(Constant)	-3.173	4.634		-.685	.494	-12.275	5.930
	zscore	1.676	.681	.105	2.461	.014	.338	3.014
	Site1	-1.015	.465	-.096	-2.182	.030	-1.928	-.101
	Site2	1.168	.645	.076	1.809	.071	-.100	2.435
	Underweight	-1.205	.951	-.054	-1.267	.206	-3.073	.663
	overweight	.583	.386	.065	1.513	.131	-.174	1.341
	obese	1.388	.554	.108	2.508	.012	.301	2.476
	Newborn sex	-1.889	.325	-.239	-5.815	.000	-2.527	-1.251
	GA at delivery	.053	.017	.131	3.186	.002	.020	.085

Model Summary[b]

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate	Durbin-Watson
1	.366[a]	.134	.119	3.7088	1.984

a. Predictors: (Constant), History of diabetes, Underweight, Newborn sex, Site2, obese, GA at delivery, zscore, overweight, Site1

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.	95.0% Confidence Interval for B	
		B	Std. Error	Beta			Lower Bound	Upper Bound
1	(Constant)	-3.874	4.624		-.838	.403	-12.958	5.210
	zscore	1.864	.683	.117	2.729	.007	.522	3.206
	Site1	-.975	.463	-.092	-2.104	.036	-1.885	-.064
	Site2	1.255	.644	.082	1.950	.052	-.009	2.520
	Underweight	-1.193	.947	-.053	-1.260	.208	-3.054	.667
	overweight	.508	.386	.057	1.317	.188	-.250	1.265
	obese	1.319	.552	.102	2.388	.017	.234	2.403
	Newborn sex	-1.874	.324	-.237	-5.793	.000	-2.510	-1.238
	GA at delivery	.055	.017	.138	3.344	.001	.023	.088
	History of diabetes	5.081	2.182	.096	2.329	.020	.795	9.367

Model Summary[b]

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate	Durbin-Watson
1	.375[a]	.141	.125	3.6978	1.999

a. Predictors: (Constant), Gestational diabetes, Site1, Newborn sex, GA at delivery, History of diabetes, Site2, overweight, Underweight, zscore, obese
b. Dependent Variable: Percent fat(PEAPOD)

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.	95.0% Confidence Interval for B	
		B	Std. Error	Beta			Lower Bound	Upper Bound
1	(Constant)	-4.718	4.629		-1.019	.309	-13.812	4.376
	zscore	1.882	.681	.118	2.763	.006	.544	3.220
	Site1	-1.000	.462	-.095	-2.164	.031	-1.907	-.092
	Site2	1.310	.642	.085	2.039	.042	.048	2.572
	Underweight	-1.135	.945	-.050	-1.201	.230	-2.990	.721
	overweight	.455	.385	.051	1.181	.238	-.302	1.212
	obese	1.135	.558	.088	2.034	.042	.039	2.231
	Newborn sex	-1.881	.323	-.238	-5.832	.000	-2.515	-1.248
	GA at delivery	.058	.017	.145	3.519	.000	.026	.091
	History of diabetes	4.657	2.185	.088	2.131	.034	.364	8.950
	Gestational diabetes	1.756	.865	.084	2.030	.043	.057	3.456

Model Summary[b]

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate	Durbin-Watson
1	.377[a]	.142	.124	3.6994	1.996

a. Predictors: (Constant), Diabetes,thyroid disease,other endo, Site1, GAat delivery, Newborn sex, Site2, obese, Underweight, History of diabetes, zscore, Gestational diabetes, overweight

b. Dependent Variable: Percent fat(PEAPOD)

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.	95.0% Confidence Interval for B	
		B	Std. Error	Beta			Lower Bound	Upper Bound
1	(Constant)	-4.609	4.633		-.995	.320	-13.711	4.494
	zscore	1.891	.681	.119	2.775	.006	.553	3.230
	Site1	-1.010	.462	-.096	-2.184	.029	-1.919	-.102
	Site2	1.339	.644	.087	2.080	.038	.074	2.604
	Underweight	-1.124	.945	-.050	-1.190	.235	-2.981	.732
	overweight	.424	.388	.047	1.095	.274	-.337	1.186
	obese	1.101	.560	.085	1.965	.050	.001	2.201
	Newborn sex	-1.886	.323	-.239	-5.843	.000	-2.520	-1.252
	GA at delivery	.058	.017	.144	3.492	.001	.025	.090
	History of diabetes	4.270	2.247	.081	1.900	.058	-.144	8.684
	Gestational diabetes	1.586	.895	.076	1.772	.077	-.173	3.344
	Diabetes, thyroid disease, other endo	.719	.965	.033	.746	.456	-1.176	2.614

Model Summary[b]

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate	Durbin-Watson
1	.385[a]	.148	.128	3.6894	1.998

a. Predictors: (Constant), Mat smoking, obese, GA at delivery Newborn sex, Underweight, History of diabetes, Site2, Gestational diabetes, zscore, overweight, Site1, Diabetes,thyroid disease, other endo

b. Dependent Variable: Percent fat(PEAPOD)

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.	95.0% Confidence Interval for B	
		B	Std. Error	Beta			Lower Bound	Upper Bound
1	(Constant)	-4.167	4.626		-.901	.368	-13.256	4.922
	zscore	1.816	.681	.114	2.668	.008	.479	3.154
	Site1	-.991	.461	-.094	-2.149	.032	-1.898	-.085
	Site2	1.318	.642	.086	2.052	.041	.056	2.580
	Underweight	-1.131	.943	-.050	-1.200	.231	-2.983	.720
	overweight	.514	.389	.057	1.321	.187	-.250	1.279
	obese	1.133	.559	.088	2.027	.043	.035	2.230
	Newborn sex	-1.920	.322	-.243	-5.957	.000	-2.554	-1.287
	GA at delivery	.057	.017	.141	3.426	.001	.024	.089
	History of diabetes	4.052	2.244	.077	1.806	.071	-.356	8.460
	Gestational diabetes	1.478	.894	.071	1.652	.099	-.279	3.235
	Diabetes, thyroid disease, other endo	.783	.963	.036	.814	.416	-1.108	2.674
	Mat smoking	-.956	.488	-.080	-1.957	.051	-1.915	.003

Model Summary[b]

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate	Durbin-Watson
1	.348[a]	.121	.116	360.34624	1.955

a. Predictors: (Constant), Site2, zscore, Site1
b. Dependent Variable: truefatfreemassg

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.	95.0% Confidence Interval for B	
		B	Std. Error	Beta			Lower Bound	Upper Bound
1	(Constant)	3003.542	18.409		163.157	.000	2967.379	3039.705
	zscore	346.501	64.634	.224	5.361	.000	219.531	473.472
	Site1	-157.779	42.818	-.154	-3.685	.000	-241.892	-73.666
	Site2	-269.332	61.722	-.181	-4.364	.000	-390.582	-148.083

Model Summary[b]

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	.376[a]	.142	.132	357.14503

a. Predictors: (Constant), obese, Underweight, Site2, zscore, overweight, Site1
b. Dependent Variable: truefatfreemassg

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.	95.0% Confidence Interval for B	
		B	Std. Error	Beta			Lower Bound	Upper Bound
1	(Constant)	2987.619	22.651		131.897	.000	2943.121	3032.117
	zscore	365.096	64.829	.236	5.632	.000	237.741	492.450
	Site1	-113.363	44.545	-.111	-2.545	.011	-200.871	-25.855
	Site2	-281.995	61.642	-.189	-4.575	.000	-403.089	-160.901
	Underweight	-231.523	90.859	-.106	-2.548	.011	-410.012	-53.033
	overweight	21.478	36.937	.025	.581	.561	-51.082	94.039
	obese	123.360	53.079	.099	2.324	.020	19.088	227.631

Model Summary[b]

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	.415[a]	.172	.161	351.03553

a. Predictors: (Constant), Newborn sex, obese, Underweight, Site2, zscore, overweight, Site1
b. Dependent Variable: truefatfreemassg

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.	95.0% Confidence Interval for B	
		B	Std. Error	Beta			Lower Bound	Upper Bound
1	(Constant)	2917.011	27.401		106.455	.000	2863.182	2970.840
	zscore	336.415	64.050	.217	5.252	.000	210.591	462.239
	Site1	-120.645	43.814	-.118	-2.754	.006	-206.717	-34.573
	Site2	-293.847	60.647	-.197	-4.845	.000	-412.986	-174.708
	Underweight	-216.595	89.369	-.099	-2.424	.016	-392.158	-41.033
	overweight	19.290	36.308	.022	.531	.595	-52.036	90.616
	obese	120.103	52.176	.096	2.302	.022	17.604	222.601
	Newborn sex	135.173	30.580	.176	4.420	.000	75.099	195.247

Model Summary[b]

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	.644[a]	.414	.406	295.54952

a. Predictors: (Constant), GAat delivery, Site1, Newborn sex, obese, Site2, Underweight, zscore, overweight
b. Dependent Variable: truefatfreemassg

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.	95.0% Confidence Interval for B	
		B	Std. Error	Beta			Lower Bound	Upper Bound
1	(Constant)	-2494.641	367.710		-6.784	.000	-3217.002	-1772.280
	zscore	282.492	54.049	.183	5.227	.000	176.312	388.671
	Site1	-127.401	36.892	-.124	-3.453	.001	-199.874	-54.928
	Site2	-235.358	51.215	-.158	-4.596	.000	-335.968	-134.747
	Underweight	-132.005	75.461	-.060	-1.749	.081	-280.247	16.237
	overweight	42.777	30.611	.049	1.397	.163	-17.357	102.911
	obese	129.417	43.934	.103	2.946	.003	43.110	215.723
	Newborn sex	153.343	25.776	.200	5.949	.000	102.706	203.980
	GA at delivery	19.377	1.314	.498	14.746	.000	16.796	21.959

Model Summary[b]

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	.647[a]	.419	.409	294.74133

a. Predictors: (Constant), Mat smoking, obese, GA at delivery, Newborn sex, Underweight, Site2, zscore, overweight, Site1

b. Dependent Variable: truefatfreemassg

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.	95.0% Confidence Interval for B	
		B	Std. Error	Beta			Lower Bound	Upper Bound
1	(Constant)	-2465.834	366.996		-6.719	.000	-3186.794	-1744.873
	zscore	277.094	53.971	.179	5.134	.000	171.069	383.120
	Site1	-125.794	36.800	-.123	-3.418	.001	-198.087	-53.501
	Site2	-236.769	51.080	-.159	-4.635	.000	-337.114	-136.423
	Underweight	-132.342	75.255	-.061	-1.759	.079	-280.180	15.495
	overweight	49.758	30.731	.057	1.619	.106	-10.614	110.130
	obese	131.207	43.823	.105	2.994	.003	45.118	217.297
	Newborn sex	150.643	25.742	.197	5.852	.000	100.073	201.213
	GA at delivery	19.303	1.311	.496	14.724	.000	16.728	21.879
	Mat smoking	-76.686	38.888	-.066	-1.972	.049	-153.082	-.290

Model Summary[b]

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	.651[a]	.424	.413	293.68069

a. Predictors: (Constant), Age at PEAPOD (wks), zscore, Underweight, Mat smoking, Newborn sex, obese, GA at delivery, overweight, Site1, Site2

b. Dependent Variable: truefatfreemassg

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.	95.0% Confidence Interval for B	
		B	Std. Error	Beta			Lower Bound	Upper Bound
1	(Constant)	-2377.208	367.906		-6.461	.000	-3099.960	-1654.456
	zscore	289.070	54.054	.187	5.348	.000	182.880	395.259
	Site1	-116.015	36.938	-.113	-3.141	.002	-188.580	-43.450
	Site2	-167.640	59.885	-.112	-2.799	.005	-285.283	-49.997
	Underweight	-126.714	75.028	-.058	-1.689	.092	-274.107	20.679
	overweight	50.692	30.624	.058	1.655	.098	-9.468	110.853
	obese	129.539	43.672	.104	2.966	.003	43.746	215.333
	Newborn sex	149.081	25.659	.194	5.810	.000	98.673	199.489
	GA at delivery	19.072	1.311	.490	14.553	.000	16.497	21.646
	Mat smoking	-79.415	38.769	-.069	-2.048	.041	-155.575	-3.254
	Age at PEAPOD (wks)	-185.044	84.469	-.086	-2.191	.029	-350.984	-19.103

Model Summary[b]

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	.652[a]	.425	.413	293.70165

a. Predictors: (Constant), Mat recreational drugs, Age at PEAPOD (wks), obese, Newborn sex, Underweight, GAat delivery, Mat smoking, zscore, overweight, Site1, Site2
b. Dependent Variable: truefatfreemassg

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.	95.0% Confidence Interval for B	
		B	Std. Error	Beta			Lower Bound	Upper Bound
1	(Constant)	-2363.541	368.207		-6.419	.000	-3086.886	-1640.195
	zscore	287.264	54.090	.186	5.311	.000	181.003	393.526
	Site1	-118.262	37.015	-.116	-3.195	.001	-190.977	-45.546
	Site2	-164.938	59.955	-.111	-2.751	.006	-282.720	-47.157
	Underweight	-128.145	75.048	-.059	-1.707	.088	-275.578	19.288
	overweight	47.714	30.782	.055	1.550	.122	-12.758	108.186
	obese	126.513	43.788	.101	2.889	.004	40.491	212.535
	Newborn sex	147.776	25.697	.193	5.751	.000	97.294	198.258
	GA at delivery	19.032	1.311	.489	14.514	.000	16.456	21.608
	Mat smoking	-73.959	39.184	-.064	-1.887	.060	-150.937	3.018
	Age at PEAPOD (wks)	-185.701	84.478	-.087	-2.198	.028	-351.659	-19.743
	Mat recreational drugs	-129.592	134.728	-.033	-.962	.337	-394.266	135.082

Model Summary[b]

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	.656[a]	.430	.417	292.69519

a. Predictors: (Constant), Diabetes,thyroid disease, other endo, Site1, GAat delivery, Mat smoking, Newborn sex, Age at PEAPOD (wks), Mat recreational drugs, obese, Underweight, zscore, overweight, Site2

b. Dependent Variable: truefatfreemassg

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.	95.0% Confidence Interval for B	
		B	Std. Error	Beta			Lower Bound	Upper Bound
1	(Constant)	-2372.705	366.970		-6.466	.000	-3093.624	-1651.786
	zscore	292.713	53.965	.189	5.424	.000	186.698	398.728
	Site1	-120.546	36.903	-.118	-3.267	.001	-193.043	-48.049
	Site2	-157.980	59.837	-.106	-2.640	.009	-275.531	-40.429
	Underweight	-124.727	74.808	-.057	-1.667	.096	-271.688	22.235
	overweight	38.733	30.961	.045	1.251	.211	-22.090	99.557
	obese	114.242	44.011	.091	2.596	.010	27.781	200.703
	Newborn sex	146.893	25.612	.192	5.735	.000	96.577	197.209
	GA at delivery	19.059	1.307	.489	14.584	.000	16.492	21.627
	Mat smoking	-74.424	39.050	-.064	-1.906	.057	-151.139	2.291
	Age at PEAPOD (wks)	-180.515	84.223	-.084	-2.143	.033	-345.973	-15.057
	Mat recreational drugs	-128.694	134.267	-.032	-.958	.338	-392.464	135.075
	Diabetes, thyroid disease, other endo	153.476	71.536	.072	2.145	.032	12.943	294.009