New Preterm Infant Growth Curves

Influence of Gender and Race on Birth Size

A Thesis

Submitted to the Faculty

of

Drexel University

by

Sue A. Groveman

in partial fulfillment of the

requirements for the degree

of

Masters of Science in Human Nutrition

July 2008

DEDICATIONS

I would like to dedicate this project to my parents who instilled a lifelong love of learning and challenge in me. Although they aren't here to see this, I am sure that they would be proud.

ACKNOWLEDGEMENTS

They say it takes a village to raise a child, well, I think it takes a village to complete a thesis project. There were many people who were crucial to this project's completion; I would especially like to thank:

- Dr. Irene Olsen for her guidance, working many long hours with me on this project and providing grant funding.
- Dr. Babette Zemel for contributing her knowledge, guidance, and humor to this project.
- Dr. Reese Clark from Pediatrix Medical Group for providing us with the data set and encouragement when the project began.
- The members of my committee, Dr. Cecilie Goodrich, Dr. Philip Handel, and Dr. Karen Drummond, for support, encouragement, and time.
- Dr. Jason Liao for answering statistics questions at crucial moments.

This project could not have been completed without the support of my friends, family and graduate school colleagues. I would especially like to thank Dr. Jennifer Quinlan and her lab for humoring me and allowing me to stop in for breaks. Finally, I would like to thank my husband who provided many statistics lessons, support, and endured long work hours.

This work was funded, in part, under a grant from the Pennsylvania Department of Health through the Drexel University's Grants for Research Impact at Drexel (#240465).

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

ABSTRACT

New Preterm Infant Growth Curves
Influence of Gender and Race on Birth Size
Sue A. Groveman
Irene E. Olsen, Ph.D., R.D.

Background

Adequate fetal and postnatal growth is crucial to the health of infants and significant to their future adult health. Preterm infant growth curves are used in the assessment of fetal and postnatal growth. Current growth curves, such as the Lubchenco curves, are limited by older homogenous data sets, smaller sample sizes, varying age ranges, combined-gender curves and/or disparate data sources. Evaluation of these curves was needed to determine whether a new set of curves with updated high-risk percentile classification cutoffs for preterm infants was needed. The presence of birth size differences between genders and racial groups also was important to identify and explore.

Methods

In the first specific aim, smoothed growth curves were created for males and females for weight, length and head circumference using a large data set (Pediatrix Medical Group). The original data included 391,681 infants (56% male; 52% white, 23% Hispanic, 15% black, 10% other) ranging from 22 to 42 weeks gestation. Infants of non-singleton pregnancies or with ambiguous gender, congenital anomalies or physiologically improbable growth measurements were excluded. The curves were fit to the data using the LMS method. The goodness-of-fit was assessed using worm plots, z-scores, and visual inspection. The curves were validated through the evaluation of z-scores and

distribution of infants between the high-risk percentile-for-age classifications. The new curves were compared to the Lubchenco curves visually, at selected points, and via the percentages of small-for-gestational-age (SGA) and large-for-gestational-age (LGA) infants.

The second specific aim investigated size differences between males and females and among racial groups. Gender differences in birth size were compared via ANOVAs, overlaid curves, specific points on the new Pediatrix curves, and distribution of infants between the high-risk percentile-for-age classifications. Racial differences in birth size were compared via ANOVAs, z-score comparisons, distribution of infants between the high-risk percentile-for-age classifications, and odds ratios. Several maternal characteristics were examined to explain the differences found in birth size by racial group (logistic regression). The sickness level of the infants was investigated to help explain group differences.

Results

For the first specific aim, the new Pediatrix percentiles were found to have smaller measurements than the Lubchenco curves until about 30 weeks, were somewhat similar between about 30 and 36 weeks, and were larger after 36 weeks. Overall the Lubchenco curves misclassified 15% of males and 10% of females as SGA, appropriate-for-gestational age (AGA), or LGA according to the new Pediatrix curves.

For the second specific aim, males were found to be significantly larger than females in weight, length, and head circumference. A higher percentage of black infants were born earlier and were significantly smaller than Hispanic and white infants. Relative to the medians, depending on gestational age, the z-scores of black infants placed up to

19 percentiles below white infants. At 32-41 weeks, black infants had two to three times the risk of white infants of being born SGA. Controlling for race, preeclampsia/eclampsia was a consistent predictor of SGA for weight, length and head circumference at all age groups and smoking was a predictor for infants born at 32-41 weeks. The prevalence of being sick (APGAR at 1 minute < 3) at birth was greater in black infants and those classified as SGA for weight.

Conclusions

Accurate preterm growth curves are crucial to the assessment of growth status and therefore are vital to the health of infants. This study found strong evidence for the replacement of the Lubchenco growth curves with contemporary gender-specific curves. This study also found significant differences in birth size (weight, length, and head circumference) between male and female infants and among the racial groups that warrant further investigation.

CHAPTER 1. INTRODUCTION

Background/Statement of Problem

Adequate fetal and postnatal growth has been shown to be crucial to the health of infants as well as significant to their future adult health. Being born significantly smaller or larger for one's gestational age puts an infant at greater risk for mortality, disease, complications and/or neurological delays. Reduced postnatal growth can be indicative of neurological delays whereas increased postnatal growth has been associated with the onset of metabolic syndrome as an adult. Therefore, birth size and postnatal growth are key indicators in the postnatal health and later outcomes of preterm infants.

Growth curves are used to assess both birth size and postnatal growth. These curves provide a method for health professionals to compare an infant's size with that of a sample of their peers. Growth curves are used in both the public health and clinical settings. In public health, epidemiologists use growth curves in the evaluation of fetal growth to find at-risk populations and look for opportunities for interventions. In the clinical realm, growth curves are used to evaluate an infant's risk for complications at birth, his/her postnatal growth and to assist in determining daily nutrition care.

For preterm infants, the "gold standard" of growth is fetal (in-utero) growth at the same gestational age. Preterm infant growth curves describe the in-utero growth of infants through a cross-sectional sample of infant birth size measurements. The preterm growth curves currently in use are generally limited by small older homogeneous samples, varying ranges of gestational ages, combined gender curves, disparate data sources, and/or a lack of length, head circumference and body proportionality measurements.

The Lubchenco curves developed in the 1960's contain percentile curves for birth weight, length, head circumference, and ponderal index (a body proportionality index). These curves were based on a data set from a hospital in Colorado and have been well-used over the last four decades. However, it is time to consider whether an update of these curves is needed.

Research has shown that birth size varies by gender and race; so, it is crucial to have a diverse sample of infants when creating growth curves. It is also important to consider the influence of gender and race on the values that are represented in the curves. New contemporary curves are vital to the accurate evaluation of birth size and postnatal growth which in turn is essential to the identification of infants who are at increased health risks and in shaping their daily nutrition care.

Specific Aims

The goal of this study was to perform the initial analysis of a large data set of recent birth size measurements (weight, length and head circumference). Specifically:

1) The birth size of infants in a recent (1998-2006) large sample of preterm infants was investigated through the creation of new growth curves. New preterm intrauterine growth curves were created that describe fetal growth and were compared to the Lubchenco intrauterine growth curves (Lubchenco, Hansman, Dressler et al., 1963). This comparison contributed to an understanding of whether these older growth curves, which contain data from more than four decades ago, should be replaced by new curves.

It was hypothesized that the new growth curves would have statistically significant differences from the Lubchenco curves. In particular, it was expected that there would be statistically significant differences for infants of less than 30 weeks in age in the percentiles of weight, length, and head circumference. Poor agreement was expected when comparing where the infants in the new data set fall in the percentiles of Lubchenco's curves and the new curves. It was expected that Lubchenco's curves would underestimate the number of infants that were small for their gestational age (SGA) and overestimate the number of infants that were large for their gestational age (LGA).

2) The relationship of gender and race of the infants with birth size (weight, length and head circumference), and gestational age was examined. The growth status of the various groups at each gestational age were compared.

It was hypothesized that statistically significant differences would be found among the genders and racial groups in birth size. It was expected that female infants would be lighter by 3-5% depending on gestational age. It was also expected that white and Hispanic infants would have greater birth weights than black infants by approximately 3-5% depending on age.

Limitations

A potential limitation with this study had to do with measurement errors and accuracy. The growth measurements (weight, length and head circumference) were not collected in a controlled research setting. The measurements were not made using standardized techniques or in replicate and were performed in many hospitals by many different nurses. However, the measurements were made in the clinical setting where the nurses generally receive training on measurement procedures. In order to reduce errors in the data set, extreme outliers were removed from the data set.

Another potential limitation with this kind of study was that the gestational age calculation is rarely perfect. Ultrasounds are more accurate than estimates based on the date of the last menstrual period. However, it is unrealistic to expect ultrasounds to be used in a large study of this kind which covers many hospitals. The gestational age in this study was determined by the neonatologist who used his or her best judgment to determine the gestational age based on the information known to him or her. Of course, there could be errors due to the mother's recall of her last period or bleeding which took place after conception. However, by removing outliers by birth size for a gestational age, most of the infants with improbable ages for their size were removed.

A third potential limitation in studying the influence of race on birth size is that the influence of multi-racial infants was not considered. The race/ethnic origin was determined based on the race of the mother. This was consistent with many other studies of race and birth size (Alexander, Kogan, and Himes, 1999; Overpeck, Hediger, Zhang et al., 1999).

Definitions

Appropriate for gestational age: 10^{th} to the 90^{th} percentile for age

Extremely low birth weight: less than 1000 g

Extremely preterm: birth prior to 28 weeks

Fullterm: birth between 37 and 42 weeks

Interquartile range: 25^{th} percentile – 75^{th} percentile

Large for gestational age: greater than the 90^{th} percentile for age

Low birth weight: less than 2500 g

Ponderal index: weight (g) multiplied by 100 divided by $length^3$ (cm^3)

Post-term: birth after 42 weeks

Preterm: birth prior to 37 weeks

Very low birth weight: less than 1500 g

Very preterm: birth prior to 32 weeks

Small for gestational age: less than the 10^{th} percentile for age

Abbreviations

AGA: Appropriate for gestational age

ANOVA: Analysis of Variance

CDC: Centers for Disease Control and Prevention

cm: centimeters

EDF: Equivalent degrees of freedom

ELBW: Extremely low birth weight

g: grams

g/kg/d: grams per kilogram per day

HIV: Human immunodeficiency virus

ICD: International Classification of Diseases

IQR: Interquartile range

IUGR: Intrauterine growth retardation

kg: kilogram

LBW: Low birth weight

LGA: Large for gestational age

m: meters

NICU: Neonatal intensive care unit

SGA: Small for gestational age

VLBW: Very low birth weight

WHO: World Health Organization

CHAPTER 2. LITERATURE REVIEW

This section of the thesis will discuss types of growth curves and their uses in the public health and clinical settings. The significance of birth size and being born preterm will also be reviewed. The known birth size differences by gender and race will be discussed. Lastly, three commonly used growth curves will be examined and the reasons for creating new preterm growth curves will be considered.

Types of growth curves

Growth curves can be thought of as either a standard (prescriptive) curve or a reference (descriptive) curve (Cameron, 1999). A standard curve portrays optimal growth while a reference curve describes the actual growth of a sample population (Cameron, 1999). The data set that is used in creating the curve along with the inclusion and exclusion criteria will determine the type of curve. Both standard and reference curves can be created with cross-sectional or longitudinal data. Cross-sectional curves describe a sample at one point in time whereas longitudinal curves follow a sample over time and thus show growth status over time.

For preterm infants, cross-sectional curves based on birth data represent intrauterine growth while longitudinal curves represent postnatal growth. Intrauterine growth curves, which are defined in this thesis as “preterm growth curves”, describe the in-utero growth of fetuses derived from cross-sectional data of birth sizes of preterm and term infants. In-utero growth is the generally accepted gold standard for the assessment of growth for preterm infants (American Academy of Pediatrics, 1977). Preterm growth curves designed to reflect the best estimations of optimal fetal growth can be used to classify infants' growth status, assess the growth of preterm infants compared to in-utero growth, and be used by epidemiologist's to research the growth status of various populations.

Classifications of newborns based on growth status and gestational age

Growth curves are used to classify newborn infants based on their birth size (weight, length, head circumference) and gestational age. Although all three measurements (weight, length, and head circumference) are important, the healthcare emphasis tends primarily to be on weight and gestational age. These classifications are used to determine nutritional status and guide nutrition care decisions for the infants as well as identify those at risk for complications or future health problems. Levels of risk which predict long-term outcomes for populations can be based on the prevalence of infants in the different classifications. For these reasons, contemporary growth curves are crucial in categorizing newborns and thus vital for individual and public health.

Gestational age, which is generally defined as the number of days since the mother's last normal menstrual period, is used to categorize infants as post-term, fullterm, preterm, very preterm and extremely preterm at birth. Post-term is defined as birth at greater than 42 weeks, fullterm is generally defined as birth between 37 and 42 weeks, and preterm as prior to 37 weeks (Cochran and Lee, 2004). It is also generally accepted that very preterm is birth prior to 32 weeks and extremely preterm is birth before 28 weeks.

Birth weight can be used alone to classify infants. The common categorizations defined by the World Health Organization (WHO) (The World Health Organization,

2007) are low birth weight (LBW), very low birth weight (VLBW) and extremely low birth weight (ELBW). LBW is generally defined as less than 2500 g (5.5 lbs), VLBW as less than 1500 g (3.3 lbs) and ELBW as less than 1000 g (2.2 lbs) (Cochran and Lee, 2004; The World Health Organization, 2007). Approximately two thirds of the individuals classified as LBW are born preterm (Tucker and McGuire, 2004). Often, the birth weight classifications may be used as a research or clinical gauge rather than weight-for-age since the gestational age may be unknown. However, the limitations of using birth weight alone results in the grouping of preterm and small fullterm infants into one category. A preterm infant classified as LBW may actually be small or appropriately sized for their gestational age. LBW fullterm infants are all small for their ages by definition but because they are more mature, they have different outcomes from preterm infants (Schlesinger and Allaway, 1955). When possible it is best to avoid classifying infants on weight alone.

A more accurate picture of growth status is obtained through evaluation of the combined birth weight, length or head circumference for gestational age classification. As mentioned above, weight-for-age is the most commonly used classification. The reasons for this include the fact that weight is the simplest measurement to obtain accurately and that there are numerous weight-for-age growth references (WHO Expert Committee, 1995). The WHO 1995 report comments that recorded birth weight is usually accurate as “mechanical and electronic scales provide reasonably valid and precise readings” (WHO Expert Committee, 1995). Small for gestational age (SGA) is generally accepted to be below the 10^{th} percentile while large for gestational age (LGA) is above the 90^{th} percentile (Battaglia and Lubchenco, 1967). Appropriate for gestational age

(AGA) is the middle range between the 10th and 90th percentiles (Battaglia and Lubchenco, 1967).

Although weight is most commonly used in the assessment of growth status, length and head circumference are also essential parts of the puzzle. Birth length can be used in combination with birth weight to differentiate between shortness or stunting and intrauterine growth retardation (IUGR) (Waterlow, 1972). Unfortunately, birth length is measured less accurately than birth weight due to lack of an electronic mechanism and positioning issues (Kramer, McLean, Olivier et al., 1989). However, length is still an important measurement used in assessing growth status.

Head circumference is also a key measurement in assessing the growth status of infants. With a few exceptions caused by medical conditions, head circumference is a direct measurement of brain growth (Cooke, Lucas, Yudkin et al., 1977). A head circumference that is SGA may be a sign of impaired brain development due to IUGR or a neurological condition (Gibson, 2005). A head circumference that is LGA may indicate a condition, such as hydrocephalus (Nevin-Folino, 2000). Since head circumference is a key method for assessment of neurological development, it is fortunate that studies have shown that head circumference can be measured with reasonable accuracy (Kramer, McLean, Olivier et al., 1989).

Preterm birth

Low birth weight of infants can be related to either IUGR or short gestational periods resulting in preterm birth. The number of preterm births in the United States has been increasing over the last two decades as shown in Figure 1 (B.E. Hamilton, Martin, and Ventura, 2006). More than half of a million children were born preterm in 2005 (B. E. Hamilton, Minino, Martin et al., 2007). The increased usage of assistive reproductive technologies may likely be influencing this trend of increasing numbers of preterm infants (Tucker and McGuire, 2004). This fragile preterm population is certainly large enough to warrant growth curves devoted to them.

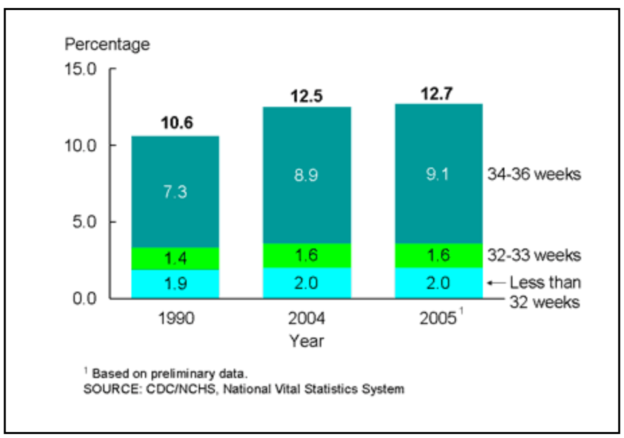


Figure 1: Percentage of preterm births: United States, 1990, 2004, 2005(B.E. Hamilton, Martin, and Ventura, 2006)

The preterm population is at great risk for serious complications or early death. The infant mortality rate in 2003 for very preterm infants (less than 32 weeks of gestation) was 78 times greater than the rate for fullterm infants (Mathews and MacDorman, 2006). Common reasons for preterm birth include smoking, history of preterm birth, socio-economic status, and IUGR (Kramer, Olivier, McLean, Dougherty et al., 1990).

Extended hospital stays of preterm infants can be costly. The length of stay in a neonatal intensive care unit (NICU) is directly related to the degree of prematurity – the younger the infant at birth, the longer the stay. A pattern of growth which is maintained over a period of time and shows satisfactory placement on the growth curve is one of the primary factors in determining when an infant can be discharged from a NICU (American Academy of Pediatrics, 1998). Satisfactory placement on the curve usually refers to being at or above the 10^{th} percentile in weight-for-age (Hovasi Cox and Doorlag, 2000). Thus, good growth may aid in reducing hospital stays and healthcare costs. Schmitt et al. (Schmitt, Sneed, and Phibbs, 2006) studied the health care costs of infants in California born in 2000. The authors found that LBW infants had much longer hospital stays ranging from 6 to 68 days depending on birth weight compared to an average of 2.3 days for children greater than 2500 g (Schmitt, Sneed, and Phibbs, 2006). LBW infants made up only 5.9% of the newborns but accounted for 56.6% of the total hospital costs (Schmitt, Sneed, and Phibbs, 2006). More contemporary growth curves may help to better define growth status through more accurately defined cutoffs potentially resulting in shorter hospital stays.

Use of growth curves

Contemporary accurate preterm growth curves are essential to assess growth status in both the public health and clinical settings. Preterm growth curves are used in the public health arena for epidemiological studies aimed at reducing the number of infants born small. In the clinical setting, they are used for two main purposes: evaluation of infant size at birth and evaluation of postnatal growth. These primary uses of growth curves are discussed below.

A. Public health

Public health practitioners use preterm growth curves for epidemiological studies of fetal growth. The comparison of fetal growth between populations often focuses on the percentage of infants that are born SGA (Alexander, Kogan, and Himes, 1999). Inadequate fetal growth can cause increased rates of mortality and morbidity, increased risks of complications as well as developmental delays (Brenner, Edelman, and Hendricks, 1976; WHO Expert Committee, 1995). Assessment of SGA birth rates can then be used to evaluate needs, look at differences within the population, and develop interventions (Alexander, Kogan, and Himes, 1999). In this way, birth size for age can be used to justify the need for public health interventions. Elevated occurrence of SGA infants can be indicative of poor maternal health and nutrition (Kramer, Olivier, McLean, Dougherty et al., 1990). Studies have shown larger infants in developed countries versus developing nations (Villar and Belizan, 1982).

The weight-for-age classification is the most relevant for this type of epidemiological study (WHO Expert Committee, 1995). Khoury et al. (Khoury, Erickson, Cordero et al., 1988) found that SGA preterm and fullterm infants had almost 2.5 times as many major congenital anomalies when compared to AGA or LGA individuals. Infants who have experienced IUGR have an elevated risk of many serious health conditions including fatality, hypoglycemia, hypocalcemia, polycythemia, and neurocognitive delays (Kramer, Olivier, McLean, Willis et al., 1990). The more severe the level of IUGR, the greater the risk of these insults (Kramer, Olivier, McLean, Willis et al., 1990). In literature, the term IUGR is sometimes used interchangeably with SGA. However, these terms are different as IUGR specifically refers to slowed growth whereas SGA refers to being born small for one's age (less than the 10^{th} percentile for age).

Evaluating infant birth size is also important because birth size has been shown to be an indicator of the future health of the child. Being born small has been indicated as a risk factor for metabolic syndrome. In 1989, Barker et al. first established that low birth weight regardless of gestational age was related to cardiovascular disease or type 2 diabetes in a study of 64 year old men (Barker, Winter, Osmond et al., 1989). Barker put forth the "fetal origins of disease" hypothesis which states that IUGR is associated with increased risk of diseases, such as metabolic syndrome, hypertension, cerebrovascular, coronary heart disease, and type 2 diabetes (Barker, Winter, Osmond et al., 1989).

Theories of origins of disease have led to the idea of "programming" early in life. The term programming refers to the "concept that an insult or stimulus applied at a critical or sensitive period may have long-term or lifetime effects on the structure or function of an organism" (Lucas, 2005). When an infant is born small or experienced IUGR, the infant may have made an adaptation or had an effect of programming which will adversely affect the future health of the individual (Lucas, 2005). Thus, it is crucial to consider the early nutrition of fetuses as well as newborn infants.

Accurate preterm growth curves will be helpful in correctly identifying populations with increased SGA rates. Types of public health programs aimed at lowering SGA rates include smoking reduction campaigns, nutritional education, prenatal health clinics and nutritional supplementation (WHO Expert Committee, 1995). Reduction of SGA rates is crucial to the health of children as is the care of LBW and VLBW infants. The WHO report of 1995 points out that trends in developed countries have shown that infant mortality can be reduced with optimal care (WHO Expert Committee, 1995). It is imperative to monitor rates of SGA and LBW infants as increases can signify new causative factors, such as human immunodeficiency virus (HIV) infection which may require interventions (WHO Expert Committee, 1995).

B. Clinical: Evaluation at birth

Preterm infant growth curves are used in the clinical setting to evaluate the infant's growth status at birth. As mentioned above, both inadequate or excess fetal growth can be indicators of increased risks of early death, complications, and developmental delays (Brenner, Edelman, and Hendricks, 1976; WHO Expert Committee, 1995). The mortality rate bears out the importance of weight as an indicator of risk. For infants born LBW but not VLBW the mortality rate is five times greater than larger infants (Mathews and MacDorman, 2006). However, for VLBW infants – this risk grows to 110 times greater (Mathews and MacDorman, 2006). Use of contemporary growth curves to classify the growth status of preterm infants is necessary to assure that risk categories are assigned appropriately (i.e. SGA/LGA cutoffs are not too high or too low), so that health professionals can be alerted to these increased risks for an SGA or an LGA infant.

When an infant is identified as SGA, measures can be taken to reduce the risks to the child. An SGA infant may be fed an increased calorie and protein diet or in developing countries breast-feeding may be recommended (Ellard, Olsen, and Sun, 2004; WHO Expert Committee, 1995). When an infant is diagnosed as preterm and SGA, the WHO suggests attempting to determine whether the infant suffered from IUGR or is simply preterm (WHO Expert Committee, 1995). Diagnosis of IUGR is crucial as IUGR can result in significant traumas for the infant to which health professionals should be alerted. Issues related to IUGR include infant death, fetal distress, meconium aspiration syndrome, hypoglycemia, polycythemia or hyperviscosity, and hypothermia (Kramer, Olivier, McLean, Dougherty et al., 1990). In order to reduce these risks, the preterm or LBW infant should be monitored for appropriate levels of blood glucose, calcium, and hemoglobin (WHO Expert Committee, 1995).

C. Clinical: Postnatal growth

Preterm infant growth curves are often used to help determine postnatal growth status for these infants which in turn guides nutrition care decisions. The American Academy of Pediatrics has traditionally recommended that nutrition of preterm infants be geared to facilitate in-utero growth rates of 15 g/kg/d weight gain (American Academy of Pediatrics, 1977). To determine if the preterm infant is on-track with in-utero growth, the health professional plots the infant on the preterm growth curve. Placement on the curve and growth velocity are used in assessing nutrition status.

However, these nutrition recommendations have become controversial in recent years. One factor is that catch-up growth may contribute to the occurrence of metabolic syndrome in adults who were born small. Catch-up growth is the process of rapid weight gain of SGA infants after birth in order to catch-up to their appropriately sized peers

(Colle, Schiff, Andrew et al., 1976). Research has begun to show that an increased rate of growth may continue past the catch-up period and that preterm or fullterm subjects who were SGA for weight at birth but obese as children have the highest risk of developing insulin resistance and cardiovascular disease (Eriksson and Forsen, 2002). So, maintaining an optimal postnatal growth velocity which fits with peers born at the same gestational age may be important in prevention of insulin resistance and metabolic syndrome. Contemporary growth curves will assist in identifying when catch-up growth begins and monitoring it over time so that nutrition may be adjusted as appropriate.

Another point of consideration is that when SGA infants catch up in size to their AGA peers – the risk of neurological delays is reduced to the same level as the AGA children (Neu, Hauser, and Douglas-Escobar, 2007). This argues for feeding infants aggressively until they at least reach the 10^{th} percentile. However, the thought that catch up growth may cause adverse adaptations which affect future health would argue for less aggressive feeding. This topic remains controversial and more research is warranted. It may be a situation where both issues need to be balanced in determining proper nutrition. Since there may be a fine line between enough growth and too much growth, this situation shows the importance of contemporary growth charts to accurately define growth status.

According to Neu et al. (Neu, Hauser, and Douglas-Escobar, 2007), the goals of nutrition for a preterm infant are to maintain growth between the 10^{th} and 90^{th} percentiles, sustain lean body mass and bone density, avoid complications, enhance neurodevelopment and foster strong future health. The method used to reach this goal is to attempt to emulate fetal nutritional intake and in turn growth that is achieved in utero

(Thureen and Hay, 2001). Preterm growth curves have been used to track the growth of infants postnatally since these curves have generally been accepted as representing the gold standard of growth for this age group (American Academy of Pediatrics, 1977). Contemporary preterm growth curves based on intrauterine data will help to achieve the goals suggested by Neu et al. (Neu, Hauser, and Douglas-Escobar, 2007) including maintaining growth between the 10^{th} and 90^{th} percentiles.

Methodologies for creating preterm growth curves

Over the years, numerous methodologies have been used in the construction of preterm growth curves. The type of data sets used, exclusions and inclusions, fitting of the curves, as well as the percentiles and cut-offs have all been debated in the literature and will be discussed in this section.

A. Data sets

Data sets for preterm growth curves generally fit into one of three categories: 1) cross-sectional measurements taken at birth, 2) longitudinal measurements taken at birth and at regular intervals thereafter, and 3) ultrasound measurements taken during pregnancy. Each type of data set has advantages and disadvantages both for the practicality in creating the data set and in determining the type of growth curves.

Data sets of cross-sectional sizes at birth are by far the most commonly compiled and used. These data sets are more easily acquired as these measurements are usually taken in hospitals as general practice and are a one-time measurement at birth as opposed to multiple measurements taken over time. However, these data sets do require the review of each participant's medical record; therefore, many studies have instead used birth weights that are recorded on birth certificates. A criticism of preterm growth curves constructed from a cross-sectional sample of infant birth sizes is that these values may not represent the in-utero growth of healthy infants because preterm infants are generally born smaller than healthy fetuses (Lubchenco, Hansman, Dressler et al., 1963). Despite this criticism, size at birth is generally accepted as representative of fetal growth and is currently used in the Lubchenco (Lubchenco, Hansman, Dressler et al., 1963), Fenton (Fenton, 2003) and Babson and Benda (Babson and Benda, 1976) curves. Through the use of inclusions/exclusions and removal of outliers, the data set of cross-sectional sizes at birth can be restricted such that the birth sizes are more representative of optimal fetal growth.

Longitudinal measurements which track preterm infant's growth over time are used in constructing "postnatal growth curves". Postnatal curves track the growth of preterm infants longitudinally and thus show the dip in weight that infants take after birth. These curves show how the growth of an infant compares to that of other preterm infants of the same age. Preterm growth curves may be used in conjunction with postnatal growth curves, such as those created by Ehrenkranz et al. (Ehrenkranz, Younes, Lemons et al., 1999).. A combination of intrauterine growth curves and postnatal curves will provide a more complete picture of growth status than either by itself.

In an attempt to generate a more accurate picture of fetal growth, data sets created from ultrasound measurements have been used. In 1996, Marsal (Marsal, Persson, Larsen et al., 1996) published growth curves based on estimates of fetal weight from ultrasounds. The ultrasound technique is not a direct measurement – multiple measurements are used to infer fetal weight. In Marsal's study, ultrasounds were taken 9 to 11 times over the course of the 89 pregnancies and three ultrasound measurements (biparietal diameter, femur length, and abdominal diameter) were input into a formula to estimate the fetal weight. At close to term these curves were similar to other Swedish curves (Marsal, Persson, Larsen et al., 1996). At less than 34 weeks, the curves were found to have weights of about 100 g larger than other Swedish curves (Marsal, Persson, Larsen et al., 1996). Yet, ultrasounds are not considered to be practical methods for measuring fetus size for a large population (Ehrenkranz, 2007; Hindmarsh, Geary, Rodeck et al., 2002). Also, we are unaware of curves for length and head circumference using ultrasound estimates. So, preterm intrauterine growth curves using infant size at birth are the most practical feasible option at this time for assessing growth status compared to in-utero growth for a large sample.

B. Inclusions and exclusions

Determining the inclusions and exclusions is also an important step in creating growth curves. These determine what type of curve will be presented. If all the subjects are included, then one will be presenting a reference (descriptive) curve which describes the growth of the entire population (Cameron, 1999). If exclusions and inclusions are chosen to limit the data set to presumably healthy children, then the growth curve will be a standard (prescriptive) curve which aims to present optimal growth of the subjects (Cameron, 1999).

C. Outliers

Choosing a method for the removal of outliers in the data set is central to maintaining the integrity of the data. There are generally two primary issues when dealing with outliers in the preterm infant data: measurement errors and improbable gestational ages. Gestational ages may be incorrect due to the mother's recall of her last period or bleeding which took place after conception. Joseph et al. (Joseph, Kramer, Allen et al., 2001) researched the dilemma of how to remove infants with implausible gestational ages from a sample set and found that no method of those tested (four standard deviations, five standard deviations, expert clinical opinion, and Tukey's rule) was superior to the others. All of these methods will also remove the obvious data or measurement errors.

D. Percentiles and cutoffs

Varying percentiles have been displayed on preterm growth curves. The 50^{th} percentile or the median is generally always represented. However, the other percentile curves have varied. Lubchenco includes the 10^{th}, 25^{th}, 50^{th}, 75^{th}, and 90^{th} percentiles (Lubchenco, Hansman, Dressler et al., 1963). The Fenton curves include the 3^{rd}, 10^{th}, 50^{th}, 90^{th}, and 97^{th} percentiles (Fenton, McMillan, and Sauve, 1990).

The high risk percentiles cutoffs of the 10^{th} and 90^{th} percentiles were proposed by Battaglia and Lubchenco (Battaglia and Lubchenco, 1967) in 1967. The initial proposal categorized infants into three gestational age categories: pre-term (<38 weeks), term (38 through 41 weeks), and post-term (≥42 weeks). Each gestational age category was then split into the SGA, AGA, and LGA categories – thereby creating nine subcategories. This proposal was primarily based on the neonatality mortality risk of the infants within these categories. The nine subcategories have not turned out to be used in that form. However, the 10^{th} (SGA), 10^{th}–90^{th}(AGA), and 90^{th} (LGA) percentile-for-age categories have stood the test of time. Given newer technologies, methods of caring for infants in a NICU, as well as maternal nutrition which has influenced the outcomes of infants today – it would make sense to reevaluate the cutoffs for these categories. However, since these are the high-risk categories that have been used in NICU's for four decades, until new research is performed, these categories will likely continue to be used.

E. Methods for smoothing curves

Growth curves are smoothed because, due to variability in the data, charting the empirical or raw data would result in a bumpy curve. Since size generally increases with age, it is not logical to have a bumpy curve which does not represent the average growth of a reference population. In Lubchenco's era, smoothing was done by hand or arithmetically (Lubchenco, Hansman, Dressler et al., 1963). Today, it is more likely that a software program that combines cubic splines is used to fit the curves. The LMS method by Cole and Green (Cole and Green, 1992) is one such technique. This method fits the curve based on the median, variation and skewness and was used by the CDC to create their recent curves (Kuczmarski, Ogden, Grummer-Strawn et al., 2000).

Once curves are fitted to the data, the next step is to choose the curve with the best fit. Methods of testing goodness-of-fit are generally based on the distribution of z-scores or a comparison of the actual and expected observations above and below the percentile curves (Royston and Wright, 2000). Worm plots, a graphical interpretation of the z-score distribution, have been shown by van Buuren and Fredriks (van Buuren and Fredriks, 2001) to be an effective technique for assessing the goodness-of-fit of the curves to the data. In practice, multiple tests are often used together to choose the best curve.

Review of currently used curves

This section will review preterm growth curves that are commonly used in NICUs today. The growth curves that will be discussed are Lubchenco et al. (Lubchenco, Hansman, Dressler et al., 1963), Babson and Benda (Babson and Benda, 1976), and Fenton (Fenton, 2003). Aside from being in common use, these curves are pertinent because they also contain all three measurements (weight, length, and head circumference) of concern. These curves are currently or were previously distributed by formula companies (Abbott Nutrition (previously Ross Pediatrics) and Mead Johnson Nutritionals); so, they have become ubiquitous in the NICU. See Table 1 below for a side-by-side comparison of the curves.

A. Lubchenco

As mentioned above, Lubchenco's curves (Lubchenco, Hansman, Dressler et al., 1963) from 1963 are commonly referenced by health professionals. These curves are distributed by Mead Johnson Nutritionals and the data from these curves is used in software distributed by Abbott Nutrition to classify infants by weight-for-age. Lubchenco's curves contain data from 5,635 infants all with Caucasian mothers of low socioeconomic status. All of the infants were born at Colorado General Hospital at 24 to

42 weeks of gestational age. The children were born from July, 1948 to January, 1961 with the exception of no infants greater than 36 weeks of gestational age admitted after 1955 – due to large numbers of infants at those gestational ages.

Of the original 7,287 infants in Lubchenco's data set, 2,192 were excluded due to incomplete records, non-Caucasian parents (black, Asian, Indian), gestational age less than 24 weeks or greater than 42 weeks, or gross pathological conditions (anencephaly, hydrocephaly, hydrops fetalis, maternal diabetes). Spanish American mothers comprised 30% of the sample and were found to have similar birth weights as the rest of the sample. A small but significant difference of about 100 grams was found between the weights of boys and girls at gestational ages of 38 to 41 weeks.

The Lubchenco curves are used widely because they contain weight, length, head circumference, and ponderal index. They are also easy to use because they contain weekly gridlines and start at 24 weeks of gestational age. However, there are limitations – the major ones being that now the data was two to three generations old and the data was homogeneous – one location and one race. Also now, more sophisticated methods exist for generating growth curves.

Methodology

The percentile curves in Lubchenco's charts were created by grouping the children by completed weeks of gestational age and then rounding the birth weight to the nearest 100 grams. The gestational age was calculated from the mother's last known menstrual period. The data was then graphed and the 10^{th}, 25^{th}, 50^{th}, 75^{th}, and 90^{th} percentiles were read from the graphs. The percentiles were graphed using the mid-point of the week and smoothed arithmetically. The precise method of arithmetical smoothing was not reported in the publication. The birth weights of the Spanish American children were plotted on a scattergram and compared to the rest of the sample. The weights were found to be similar between the groups. This methodology was repeated for length, head circumference and ponderal index.

B. Babson and Benda

Babson and Benda's (Babson and Benda, 1976) growth curves, published in 1976, are recommended for use in NICU's by Krause's Food, Nutrition, and Diet Therapy (Anderson, 2004) primarily because the curves extend from a gestational age of 26 weeks through the ten years of life. The data source for 26 to 40 weeks of age was from a study published by Usher and McLean (Usher and McLean, 1969) in 1969. Usher and McLean (Usher and McLean, 1969) measured 300 newborn singleton infants born in Montreal between 1959 and 1963 at 26 to 40 weeks with "uncomplicated pregnancies in private settings" with Caucasian parents, varying socioeconomic backgrounds, and varying national origins. Infants with major congenital abnormalities, erythroblastosis, or marked fetal malnutrition and infants born to diabetic mothers were excluded.

These curves were used widely because they contained all three measurements of weight, length and head circumference. They also have the advantage of displaying 26 weeks through ten years of life. However, these curves are limited by the very small, homogeneous sample size. All of the infants were of the same race and born in the same city. The data is now more than four decades old and there were a small number of preterm infants. In fact, there were no infants younger than 26 weeks and only 45 infants

30 weeks and younger. The curves also do not contain a measure of proportionality, such as weight for length. The curves do not show the percentiles but instead display the mean plus and minus two standard deviations. Since percentiles are not displayed, the SGA, AGA, and LGA classifications cannot be determined using Babson and Benda's curves. Lastly, the curves lack in precision since they have two week intervals for age and 500 g intervals for weight.

Methodology

The smoothed curve values for the mean and the mean plus and minus two standard deviations were plotted for weight, length, and head circumference (Usher and McLean, 1969). The gestational age was considered to be the number of weeks where the birth was from three days before to three days after a completed week (Usher and McLean, 1969). If the calculated gestational age was not compatible with the gestational age determined by a clinical assessment, the infant was excluded. The method of smoothing is not described by Usher and McLean (Usher and McLean, 1969). Babson and Benda determined that both genders showed nearly parallel rates of growth until ten years of age – so the genders were combined into one graph (Babson and Benda, 1976).

C. Fenton

Noticing the limitations in the popular growth curves due to the lack of recent data, Fenton (Fenton, 2003) created new growth curves in 2003 through the use of more recently published studies with data collected from 1963 to 1996. These curves were created with the intention of updating the Babson and Benda (Babson and Benda, 1976) curves and are now being distributed by Mead Johnson Nutritionals. Multiple data sets were chosen to be included in the study. The Kramer (Kramer, Platt, Wen et al., 2001) 2001 Canadian study was selected to be used for birth weight. For head circumference and length, two studies were used: the Niklasson (Niklasson, Ericson, Fryer et al., 1991) 1991 Swedish study and the Beeby (Beeby, Bhutap, and Taylor, 1996) 1996 Australian study. The Niklasson study contained infants of gestational age only 29 weeks and greater. The Beeby study contained approximately 27,000 infants ranging from 22 to 43 weeks. Beeby and Niklasson were averaged together using a weighted average based on sample size and was primarily Swedish since that data size had a much larger sample size. The Center for Disease Control (Kuczmarski, Ogden, Grummer-Strawn et al., 2000) (CDC) data from 1963-1994 was used for infants greater than 40 weeks of gestational age.

The Fenton curves are an update of Babson and Benda (Babson and Benda, 1976) as they extend to 50 weeks of age and contain more recent data. However, the methodology used in creating these curves has its limitations. The most notable issue is the use of multiple data sets in the creation of one set of curve, which poses numerous issues. The data sets come from different countries with one data set being used for weight and others for length and head circumference. So, clearly, these children of different national origins will likely have different birth sizes – Swedish children may be taller while Canadian children may be lighter. So, an American child may fall in the 50th percentile for weight on the chart but only the 20th percentile for length. The overall growth status of a child may be hard to interpret using Fenton’s set of curves.

Methodology

The published information on percentiles from the four data sources described above were combined to create the curves. Both genders were averaged for the 3rd, 10th, 50th, 90th, and 97th percentiles and placed in one growth chart. A statistically significant difference was found between boys and girls at all ages above 23 weeks for the 50th percentile but according to Fenton it was not a large enough difference to warrant different growth curves. The largest differences between boys and girls were in late gestation and births at greater than 40 weeks. In order to smooth disjunctions between pre and post-term sections (since different data sets), undefined "manual methods" were used for weight between 36 and 46 weeks and for head circumference and length which were smoothed back to 22 weeks.

Fenton compared the new curves to Babson and Benda's curves through a variety of methods. First, the Babson and Benda curves were superimposed on the Fenton graph and compared visually. Using *t*-tests, the means were compared by assuming that the 50th percentile from Babson and Benda was equivalent to the mean. Standard deviations were estimated using a least squares fit of the percentiles to the normal curve. The percentage of new chart values below Babson's 10th percentile was calculated. Z-scores were calculated for 10th percentile and compared to new data. The results showed that the Fenton and Babson and Benda curves were similar on 50th percentile but differed more on the 3rd and 97th percentiles. The head circumference and length were significantly different after term. However, Babson and Benda had a small sample after term. The largest difference in birth weight was at 36 weeks.

Table 1: Comparison of the Lubchenco, Babson and Benda, and Fenton curves

	Lubchenco (Lubchenco, Hansman, Dressler et al., 1963)	**Babson & Benda** (Babson and Benda, 1976)	**Fenton** (Fenton, 2003)
Sample size (n)	5,635	300	26,973 – 82 million*
Data source	Colorado General Hospital	Usher and McLean	Kramer, Niklasson, Beeby, CDC
Gestational age range (weeks)	Weight: 24 – 42 Head, Length: 26 – 42	26 – 40	22 – 50
Years data collected	1948 – 1961	1959 – 1963	1963 – 1996
Demographics	Caucasian with low socioeconomic status from Colorado	Caucasian with varying socioeconomic backgrounds from Montreal	Weight: Canada Head, length: Sweden & Australia Wt, Head, Length: United States
Inclusion criteria	Caucasian	Uncomplicated pregnancy in private setting	Each data source used different criteria
Exclusion criteria	Gross pathological conditions, age < 24 or > 42, maternal diabetes, incomplete records	Major congenital anomalies, marked fetal malnutrition, maternal diabetes, no last menstrual period listed	Each data source used different criteria
Outliers removed	Infants (26-35 weeks) whose weights were far above the 90^{th} percentile	Gestational age not compatible with clinical assessment of gestational age	Each data source used different criteria
Methodology	Percentiles taken from midpoint of week and smoothed mathematically	Published percentiles plotted and joined to other data for older children	Published percentile data combined and smoothed disjunction at 40 weeks
Limitations	Old data All Caucasians All from Colorado All low socioeconomic status	Small sample size (especially at <30 wk) Old data All Caucasians All from Montreal Biweekly grid lines	Mixing of data sources

* Fenton's sample size differed based on measurement and gestational age

Why new preterm growth curves are needed

There are a number of reasons why new preterm growth curves may be needed to replace the existing growth curves. As discussed above, limitations of current growth curves include small older homogeneous data sets, varying ranges of gestational ages, combined gender curves, and lack of length, head circumference, and body proportionality measurements obtained on the same sample. Some older curves which are in current use, such as Lubchenco (Lubchenco, Hansman, Dressler et al., 1963) and Babson and Benda (Babson and Benda, 1976) contain all three measurements of weight, length and head circumference but consist of data gathered from at least two generations ago and from small homogeneous data sets. Most recent growth curves only contain information regarding birth weight (Alexander, Himes, Kaufman et al., 1996; Oken, Kleinman, Rich-Edwards et al., 2003) since that measurement is easily obtained from birth certificates.

According to Fenton (Fenton, 2003) in 2003, a survey of 118 neonatal health professionals showed that 50% still used curves by Babson and Benda (Babson and Benda, 1976), 42% Lubchenco et al. (Lubchenco, Hansman, Dressler et al., 1963) and 18% Dancis (Dancis, O'Connell, and Holt, 1948). In a 1995 report, the WHO Committee recommended using the growth curves of Williams et al. (Williams, Creasy, Cunningham et al., 1982) because a large multiracial sample was used and presented the risk of neonatal mortality based on weight-for-age (WHO Expert Committee, 1995). However, Williams' curves include only weight-for-age measurements and therefore are not used widely in the clinical setting.

The lack of heterogeneity in the data sources for Lubchenco et al. (Lubchenco, Hansman, Dressler et al., 1963) and Babson and Benda (Babson and Benda, 1976) is an important limitation. These curves were each derived from data gathered from a limited geographical area and contained infants from primarily one racial/ethnic group. Birth sizes do vary by locale and racial/ethnic group and limiting data sets by geographical region could bias the data significantly (B. E. Hamilton, Minino, Martin et al., 2007). For example, the LBW rate in 2004 for Colorado was 9.0% while Mississippi had an 11.6% rate (B. E. Hamilton, Minino, Martin et al., 2007). The WHO 1995 report discusses the pros and cons of using population-based data versus single hospital sourced data in the creation of curves. Single location data generally has the advantage of consistent measurement techniques and data collection (WHO Expert Committee, 1995). However, as mentioned above single source data will lack diversity – possibly in socioeconomic status, racial and ethnic groups, and geographic locations – which is needed to generalize for use in NICUs across the United States.

Since growth curves plot size by gestational age, other factors for consideration are the sample size at each gestational age as well as the definition of gestational age. Currently used growth curves, such as Babson and Benda, sometimes lack data for a large range of gestational ages. Assessment of the size of the youngest of infants on the growth curve is essential in the NICU. The way gestational age has been defined in growth curves has varied between completed and partial weeks. The World Health Organization's recommendation that gestational age be measured as completed weeks was only released in 1995 (WHO Expert Committee, 1995). A standard definition of gestational age is needed for growth curves as almost a week of growth can be misrepresented.

Given that the percentile cutoffs for SGA, AGA, and LGA are obtained from growth curves, it is essential that growth curves be as accurate as possible. Goldenberg et al. (Goldenberg, Cutter, Hoffman et al., 1989) compared the 10th percentile for birth weight from various studies and found differences ranging by gestational age from 180 g to greater than 500 g. These large differences would greatly affect the number of children that would be classified as SGA. The cutoffs for SGA, AGA, and LGA affect which infants are included in epidemiological studies, are labeled "at risk", and the NICU goals for an infant's growth. New contemporary curves created from a large heterogeneous data set would provide more accurate percentiles for determining cutoffs for SGA, AGA and LGA determinations.

Prematurity and birth size: Influence of gender and race

Differences in preterm birth rates and birth sizes have been reported by gender and in various ethnic and racial groups. The rationale for studying these sub-populations is similar to studying larger populations. Since significant differences or increases have been found, research into the origins of the differences is warranted and may lead to public health interventions. A second reason is that large differences may signify the need for separate growth curves for the various groups. For example, one curve for boys and one for girls may be justified.

A. Gender differences

Risk of prematurity has been found to be significantly greater in male infants than female infants (Astolfi and Zonta, 1999). In a meta-analysis of preterm births that focused on male infants, it was found that the percentage of infants born at term was approximately 51% male while preterm births were 54% male (Zeitlin, Saurel-Cubizolles, De Mouzon et al., 2002). The reason for the increased risk of preterm delivery in male infants is not fully known. However, hypotheses for the increased risk include the greater fetal size of male infants and increased vulnerability to pregnancy complications (Zeitlin, Saurel-Cubizolles, De Mouzon et al., 2002).

Birth size has also been shown to vary by gender. Male infants tend to be larger than females at birth (Kramer, Olivier, McLean, Dougherty et al., 1990; Thomas, Peabody, Turnier et al., 2000). Kramer et al. (Kramer, Olivier, McLean, Dougherty et al., 1990) found that male infants weighed 4% more than females. A study by Thomas et al. (Thomas, Peabody, Turnier et al., 2000) found that males were larger by an average of 95 g with a range depending on gestational age (23 to 41 weeks) from approximately 40 g to 180 g. Hindmarsh et al. (Hindmarsh, Geary, Rodeck et al., 2002) observed that males between the gestational ages of 20 to 30 weeks were significantly larger in birth weight, length, and head circumference. These differences support the need for separate growth curves for males and females.

B. Racial group differences

Preterm birth rates and birth weights have also been found to vary by race and ethnic group. The "Annual Summary of Vital Statistics: 2005" which gathers data from birth certificates and fetal death reports in the United States, reports that the rate of preterm birth varies depending on the maternal race/ethnic group (B. E. Hamilton, Minino, Martin et al., 2007). In all populations, the percentages of children (Table 2) that were born preterm increased from 10.6% in 1990 to 13.5% in 2005 (B. E. Hamilton, Minino, Martin et al., 2007). However, as shown below in Table 2, black women have significantly higher preterm birth rates than white and Hispanic women.

Table 2: Preterm and LBW rates by racial group

Race of Mother	**Preterm Birth Rate**		**LBW Rate**	
	1990 (%)	**2005 (%)**	**1990 (%)**	**2005 (%)**
All races	10.6	13.5	7.0	8.2
White non-Hispanic	8.5	12.5	5.6	7.3
Hispanic	11.0	12.9	6.1	6.9
Non-Hispanic black	18.9	19.4	13.3	13.9

Source: (B. E. Hamilton, Minino, Martin et al., 2007)

There also appears to be greater correlation between poverty and preterm births in black mothers than other groups. A recent study by Reagan and Salsberry (Reagan and Salsberry, 2005) found that there is a significant correlation between neighborhood poverty and housing vacancy with very preterm births for blacks but not for Hispanics or whites. So, the causes of increased preterm births and low birth weights of black infants may be affected more strongly by the socioeconomic status hurdles of this population.

The same trends seen in preterm birth rates are also seen in LBW rates with an increase in the national LBW rate from 7.0% in 1990 to 8.2% in 2005 (B. E. Hamilton, Minino, Martin et al., 2007). However, as shown in Table 2, black women have significantly higher rates of LBW births. This difference in size was confirmed by Thomas et al. (Thomas, Peabody, Turnier et al., 2000) who found that white and Hispanic infants were larger than black infants by an average of 90 g with a range depending on gestational age from 30 g to 220 g.

The reasons for differences in birth weights may have various explanations ranging from genetics to medical issues to socioeconomic status. In 1999, Alexander et al. (Alexander, Kogan, and Himes, 1999) published a study which compared the birth weights of infants from white and black women who were classified as “low-risk” (absence of medical conditions, smoking, or alcohol use, are married, educated, and age 20-34) to those who were “not low-risk”. The authors found that in both cases of “low-risk” and “not low-risk”, the black women had infants with smaller birth weights. This was also the case with the “low-risk” group compared to the “not low-risk” group in both racial groups. However, white women were 2.5 more times likely to be classified as “low-risk” than black women. So, it appears that genetics as well as socioeconomic status may both play a role in the lower birth weights of black infants.

CHAPTER 3. PRELIMINARY STUDIES

"A new look at intrauterine growth and the impact of race, altitude and gender" (Thomas, Peabody, Turnier et al., 2000)

In this preliminary study performed by Thomas et al. (Thomas, Peabody, Turnier et al., 2000) using an earlier and smaller data set from the same source (the Pediatrix Medical Group) as the data in this study, the effect of race, gender and altitude on birth size was investigated. The data set included 27,229 subjects born from 1996 to 1998 between 23 weeks of gestational age to 41 weeks. The data was collected at 85 NICUs throughout the United States. The authors compared their data set to other curves and data sets.

It was found that gender and race had significant effects on birth weights. Females were found to be significantly smaller than males by an average of 95 g with a range depending on gestational age (23 to 41 weeks) of approximately 40 g to 180 g. White and Hispanic infants were found to be larger than black infants by an average of 90 g with a range depending on gestational age from 30 g to 220 g. It has been speculated that high altitude is correlated to decrease in birth size. However, in this study altitude was not found to have a significant effect on birth size and the authors concluded that high-altitude growth curves were not warranted.

The authors plotted the empirical percentiles to create basic growth curves. No infants were excluded in this study. These curves were compared to Babson and Benda's (Babson and Benda, 1976) and Lubchenco's curves (Lubchenco, Hansman, Dressler et al., 1963). The methods used in comparison included a graphical overlay (Figure 2) of the various curves as well as a comparison of the percentage of infants classified as SGA, AGA, and LGA with the various curves. At less than 30 weeks, the Pediatrix data was found to have less variance, lower weights, lengths and head circumferences on average than the Lubchenco and Babson and Benda curves. From 30 to 36 weeks, the data were found to be similar to Lubchenco and Babson and Benda. For greater than 36 weeks, the infants were found to be of similar size to Babson and Benda but larger than those in Lubchenco's curves. The authors concluded that the Lubchenco and Babson and Benda curves overestimate the number of infants classified as LGA and underestimate the number of infants classified as SGA.

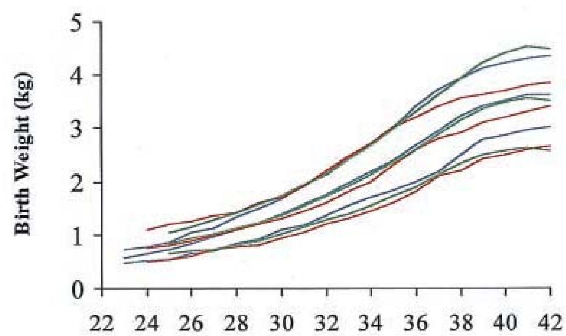


Figure 2: Comparison of growth curves; blue lines from Pediatrix data; green lines from Babson and Benda; red lines from Lubchenco (Thomas, Peabody, Turnier et al., 2000)

The birth weight data was verified against other sources in order to assure that the data was not biased. Since the infants had all been admitted to a NICU, it was possible that for near-term infants (greater than 33 weeks) more SGA or LGA children may be represented in the sample. The birth weight data were compared to a large data set from Neo Knowledge data systems and the curves were found to be very similar. The mean values for birth weight were also found to be similar to these other large national studies

– Brenner (Brenner, Edelman, and Hendricks, 1976), Arbuckle (Arbuckle, Wilkins, and Sherman, 1993), and Alexander (Alexander, Himes, Kaufman et al., 1996).

CHAPTER 4. RESEARCH DESIGN AND METHODS

A large heterogeneous data set gathered from 1998 to 2006 was used for this study. The de-identified data set was collected by the Pediatrix Medical Group at 248 NICUs in 33 states across the United States. As a next step from the earlier study by Thomas et al. (Thomas, Peabody, Turnier et al., 2000), this more recent and larger data set was used to create new “smoothed” growth curves by gender that depict percentiles for birth size measurements (weight, length, head circumference) by week for infants of gestational ages 22 to 42 weeks. The percentile curves were smoothed in order to illustrate the natural tendency of these measurements to increase with age (Cole and Green, 1992). As Figure 2 shows, curves that have not been smoothed (empirical percentiles) can have bumps due to variations in the data that are not representative of the average growth of a reference population.

Six sets of growth curves were created with the proposed intent that they will be used in the NICU for clinical purposes and will update the Lubchenco curves. The generally accepted gold standard of growth for preterm infants is fetal growth at the same gestational age (American Academy of Pediatrics, 1977). With this in mind, the curves were constructed to depict growth that is as close to optimal intrauterine growth as possible with this birth data set. This type of growth chart is considered prescriptive since it reflects desirable growth (Cameron, 1999). The inclusion and exclusion criteria for the curves were constructed to achieve this goal.

The data set was analyzed for the influence of gender and race on birth size. In preliminary studies (Thomas, Peabody, Turnier et al., 2000), differences in birth size have been shown between gender and race and further analysis into these differences were performed. The differences that were found were explored and explanations were attempted by controlling for the available variables in the Pediatrix data set.

SAS version 9.1 (SAS Institute, Cary NC) was used as the primary analysis tool and lmsChartmaker Pro version 2.7 (Medical Research Council, UK) is the tool that was used in creating the curves. S-PLUS version 8.0 (Insightful Corporation, Seattle WA) software was used in creating worm plots to analyze the fit of the curves to the data.

The data set does not include personal identifying information. The data set includes the following information, shown in Table 3, below.

Table 3: Data set fields

Site Characteristics	
Facility_ID	NICU identifier code
Facility_State	State where the NICU is located
Infant Characteristics	
PatientSeqID	Unique patient identifier
GestAge	Gestational age in completed weeks
GestAgeDays	Days since the completed weeks (0-6)
BirthWeight	Weight at birth in kilograms
BirthLength	Length at birth in centimeters
BirthHC	Head circumference at birth in centimeters
Sex	Gender of the infant
BirthNumber	Number of fetuses per pregnancy
Anomaly	Indicator of a congenital anomaly (Y,N)
MinOfAnomaly	First anomaly, when listed alphabetically
MaxOfAnomaly	Last anomaly, when listed alphabetically
AdmitGroup	Inborn (birth in facility) or outborn (transferred into facility)
Admit_DSB	Number of days since birth when child was admitted or readmitted to the NICU
BirthYear	Year of infant's birth
Delivery	Type of delivery (cesarean section, forceps extraction, vacuum extraction, vaginal delivery)
APGAR1	Apgar score at 1 minute (0-10)
APGAR5	Apgar score at 5 minutes (0-10)
DischType	Type of discharge (acute transfer, convalescent transfer, died, discharged home, discharged/transfer, transfer of service)
Maternal Characteristics	
Race	Race/ethnic origin of mother
MatAgePreg	Age of mother in years
AntenatalSteriods	Mother received steroids (Y,N)
RptOfSmoking	Mother reported smoking (Y,N)
RPT_of_IUGR	Intrauterine growth retardation (IUGR) reported (Y,N)
DiabetesRPT	Maternal diabetes (Y,N)
InsulinRPT	Mother received insulin (Y,N)
PreOREclampsiaRpt	Maternal pre-eclampsia (Y,N)

Measurement methods

The anthropometric measurements in the data set were performed by NICU nurses on the infants' admission to the unit as is standard practice. Pediatrix did not provide training on standardized measuring techniques. However, it is general practice for new nurses to be trained on measurement techniques during their orientation to the unit.

As described in the preliminary study by Thomas et al. (Thomas, Peabody, Turnier et al., 2000), birth weights were measured to the nearest gram and entered into the research data system at admission into the NICU. Weight was measured using an electronic balance. Length and head circumference were measured at birth either to the nearest millimeter or 0.5 cm and recorded in the database (Thomas, Peabody, Turnier et al., 2000). Length was measured crown to heel with a tape measure. Head circumference was measured using a tape measure at the largest circumference around the forehead.

The gestational age was estimated to the closest week and days based on the following information: the neonatologist's best approximation of gestational age, obstetrical history which would include the last known menstrual period, obstetrical examinations, prenatal ultrasounds, and physical examination after birth (Thomas, Peabody, Turnier et al., 2000). As recommended by the WHO, the completed gestational weeks were used in the curves produced in this study (WHO Expert Committee, 1995).

Data set description

The total number of infants in the original de-identified data set was 391,681. The summary statistics for the original data set are shown in Table 4. The demographics of the original data set are shown in Table 5. Since the aim was to create growth curves that can be used in NICUs throughout the United States, the demographics of the data set was compared to the general population of the United States.

All of the information in the data set was gathered from infants that were admitted to a NICU. Each infant was admitted to a NICU either because he/she was born preterm or was born fullterm with a medical condition which required more care. The most common reasons that fullterm infants were admitted to the NICU are hyperbilirubinemia or to rule out sepsis. These conditions generally do not affect the birth size of the infant and are often not serious medical issues. As shown in the preliminary study (Thomas, Peabody, Turnier et al., 2000), the birth weights of fullterm infants in Pediatrix NICUs were comparable to near term (greater than 33 weeks) infants in the United States who were not admitted to a NICU.

Table 4: Summary statistics of original data set

	Number of infants	Mean	Standard deviation	Median	Interquartile range	Minimum	Maximum
Gestational age (weeks)	391,681	34.83	4.07	35	33 – 38	22	42
Weight (kg)	389,596	2.470	0.940	2.449	1.785 – 3.180	0.267	8.400
Length (cm)	364,192	45.5	5.8	46.0	42.4 – 49.5	20.0	61.0
Head circumference (cm)	368,617	31.6	3.7	32.0	29.8 – 34.0	1.0	53.5

Table 5: Demographic data of original data set

Gender	44% Female, 56% Male
Race	52% white, 15% black, 23% Hispanic, 3% Asian, 7 % other
States represented	21.8% TX, 10.8% FL, 8.7% CA, 5.4% WA, 5.2% CO 0.1 - 4.7% from AK, AR, AZ, GA, IA, ID, IL, IN, KS, KY, LA, MD, MI, MO, NC, NJ, NM, NV, NY, OH, OK, PA, PR, SC, TN, UT, VA, WV
Congenital anomalies	10.6%
Discharge types	79.0% home, 18.0% transfer, 2.6% died
Inborn (born in facility) **or Outborn** (transferred to facility)	82.0% inborn, 15.5% outborn
Birth number (number of fetuses per pregnancy)	83.9% one, 14.1 two, 2.0% other

<u>Preliminary analysis</u>

Step 1: Data cleaning

The first step in the methodology was to clean the data. This process helped to identify and exclude data records that contained data entry or measurement errors or missing data. First, all records that had a null value for weight, length or head circumference were removed. Second, all records where the gender was recorded as ambiguous or unknown were removed from the data set.

Step 2: Inclusion and exclusion criteria

The next step was to restrict the data set based on inclusion and exclusion criteria. The inclusion criteria were all singleton infants born between 22 and 42 weeks from the available data. Multiple births were not included because multiple fetuses can influence the birth size of the infant. The exclusion criteria were factors that have a known or suspected impact on intrauterine growth. Infants that were excluded were those individuals that died during their NICU stay and those with one or more congenital anomalies with a known or suspected impact on growth. These congenital anomalies were defined as those with an ICD-9 code for anomalies in the International Classification of Diseases published by the World Health Organization. Examples of these types of anomalies included such conditions as congenital leukemia, aortic stenosis, and polycystic kidney.

Step 3: Removal of outliers

In order to identify and exclude data entry or measurement errors, Tukey's (Tukey, 1977) method used to create growth curves by Arbuckle (Arbuckle, Wilkins, and Sherman, 1993) and Beeby (Beeby, Bhutap, and Taylor, 1996) was emulated. In this method, the outliers that were two times the interquartile range below the first quartile and above the third quartile in birth weight, birth length, or head circumference for the infant's gestational age were removed (Arbuckle, Wilkins, and Sherman, 1993). Removal of the outliers was performed on the male and female infants separately. Tukey's method should have removed infants that have an improbable gestational age for their birth size and other potential data errors. Joseph et al. (Joseph, Kramer, Allen et al., 2001) investigated the dilemma of how to remove infants with implausible gestational ages from a sample set and found that no method of those tested (four standard deviations, five standard deviations, expert clinical opinion, and Tukey's rule) was superior to the others.

Step 4: Creation of random samples

The data set was cleaned and restricted via the inclusion and exclusion criteria using SAS. Since preliminary studies (Thomas, Peabody, Turnier et al., 2000) showed significant differences by gender, the data was split into two samples based on gender. Due to a limitation in the lmsChartMaker Pro software used for curve creation, the two samples were then each apportioned into two random samples with total observations of less than 100,000 each. The random samples were stratified by gestational age, race and facility state. Creating two random samples allowed each growth curve to be created with one set and verified with the second set. Summary statistics for each data set were calculated and analyzed as a whole and by gestational age. Each measurement (weight, length, head circumference) along with gestational ages was exported from SAS as a tab-delimited file for lmsChartMaker Pro to read.

Analysis for Specific Aim I:

The birth size of infants in a recent (1998-2006) large sample of preterm infants was investigated through the creation of new growth curves. New preterm intrauterine growth curves were created that describe fetal growth and were compared to the Lubchenco (Lubchenco, Hansman, Dressler et al., 1963) intrauterine growth curves.

A. Percentile calculation and growth curve creation

After applying the inclusion and exclusion criteria, the next step was to create the curves and calculate the percentiles. As shown in Figure 3, percentile curves depict selected percentiles over a range of ages. Percentile curves as opposed to a simple range of values are more desirable when a measurement is highly correlated with a variable, such as age (Cole and Green, 1992). Having the visual curves as a reference is also an easier method for pinpointing the infant's growth and allows the plotting of their growth over time. As age changes, birth size is known to increase, so the curves should also increase and in a smooth way (Cole and Green, 1992).

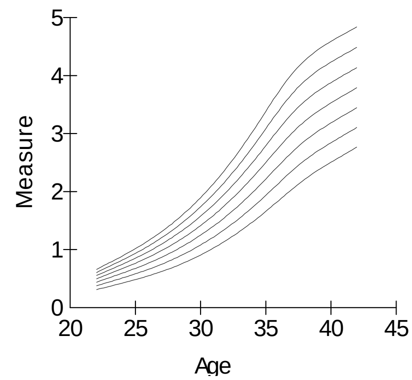


Figure 3: Percentile curves example (2nd, 10th, 25th, 50th, 75th, 90th, and 98th percentiles)

The LMS method, developed by Cole and Green (Cole and Green, 1992), and the lmsChartMaker Pro software was used to create and smooth the growth curves and calculate percentiles for weight, length, and head circumference. The lmsChartMaker Pro software was created specifically to fit percentile curves to growth data by implementing the LMS method. As previously noted, empirical data tends to be “bumpy” and thus the curves were smoothed in order to best represent the data with reduced variability (Flegal, 1999). There are five stages in using the lmsChartMaker Pro software: data entry, model fitting, graphical display, model checking, and model saving. The major effort in creating the curves and percentiles is the model fitting which primarily involves smoothing the distribution between the various gestational ages and within age at the same time (van Buuren and Fredriks, 2001). To fit the model, the LMS method uses the

L curve (skewness), M curve (median), and S curve (coefficient of variation) to perform the percentile calculations via an equation which is based on the L, M, and S values at each gestational age (Cole, Freeman, and Preece, 1998). The L curve is expressed as a Box-Cox power transformation which reduces the skewness of the distribution (Cole, Freeman, and Preece, 1998). The three curves are fitted via cubic splines by non-linear regression (Cole, Freeman, and Preece, 1998).

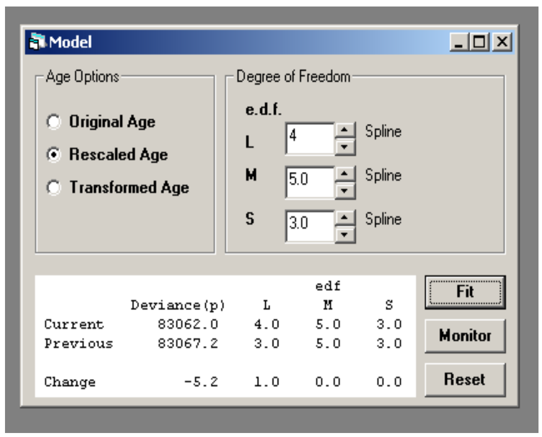


Figure 4: EDF Modeling

In order to find smooth curves that best fit the data, the equivalent degrees of freedom (EDF) were adjusted in lmsChartMaker Pro. The EDF (Figure 4) represent the complexity of each of the three curves and are each adjusted independently to modify the amount of smoothing needed for the data set (Cole, Freeman, and Preece, 1998). The EDF values range from two upwards, where two represents a straight line and the larger values represent increasingly less straight splines (Cole, Freeman, and Preece, 1998). For the chosen EDF parameters, lmsChartMaker Pro maximized the penalized likelihood (van Buuren and Fredriks, 2001). The penalized likelihood is a calculation which rewards smoother curves and penalizes rougher curves which is then maximized by the software to find the smoothest curve.

The LMS method does not have a straightforward procedure for testing the goodness of fit of the curves to the data. Therefore, worm plots were used to test the goodness of fit as they are a tool for comparing two distributions. S-PLUS software was used to draw the worm plots. The lmsChartMaker Pro software generated the z-score (number of standard deviations from the median) for each infant in the data set which was then imported by S-PLUS in order to create the worm plots. As shown in Figure 5, a worm plot is a set of detrended quantile-quantile plots (Q-Q plots) of the z-scores each for a different age group (van Buuren and Fredriks, 2001). A Q-Q plot is a visual way of comparing the shape of an empirical distribution to a theoretical distribution (van Buuren and Fredriks, 2001). In detrending the Q-Q plots, the difference between the observed quantile and the matching theoretical quantile is calculated and plotted (van Buuren and Fredriks, 2001). "The vertical axis of the worm plot displays for each observation, the difference between its location in the theoretical and empirical distributions"(van Buuren and Fredriks, 2001). When the worm is close to the line of origin (or 0.0) as in the outlined quadrants of Figure 5, a good fit is indicated.

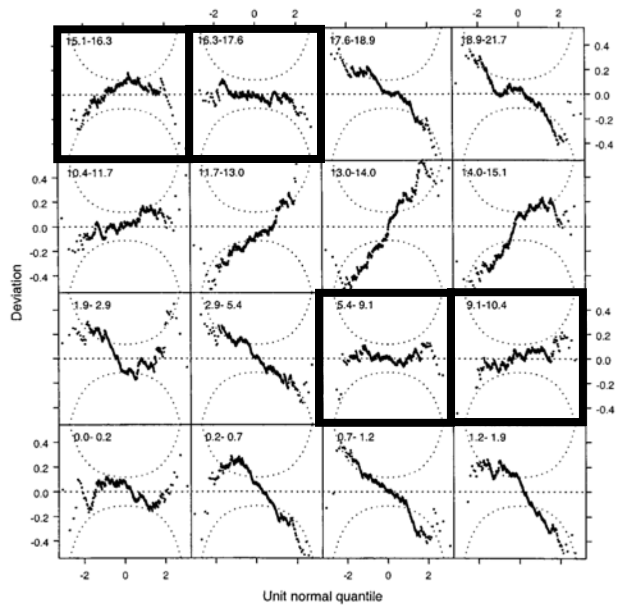


Figure 5: Example of a worm plot; outlined quadrants are good fits (van Buuren and Fredriks, 2001)

In total, six sets of growth curves were constructed. For both males and females, curves were created by gestational age for weight, length, and head circumference. The curves were calculated for the 3rd, 10th, 25th, 50th, 75th, 90th, and 97th percentiles. The first random samples were used to create a set of curves. The process of finding a good fit for each set of percentile curves was iterative involving the following steps:

- As shown in Figure 4, EDF values in lmsChartMaker Pro were adjusted one by one in the order suggested by van Buuren and Frediks (van Buuren and Fredriks, 2001). Once the change in deviance (-2 * penalized log likelihood) was below 30, worm plots were generated (via S-PLUS) with each change in the EDF values.

- The EDF value for the Median curve was adjusted first. The EDF value was adjusted until the worm plots went through the origin of the plot.

- Second, the EDF value for the S curve (variance) was adjusted until the worm plots had a slope that was close to zero.

- The EDF value for the L curve (skewness) was adjusted until any “U-shapes” were removed.

- After these adjustments, when the "worm" was close to flat and went through the center of the graph, a good fit was indicated.

- The charts (Figure 6) in lmsChartMaker Pro were also examined to determine if the curves fit the plotted data well.

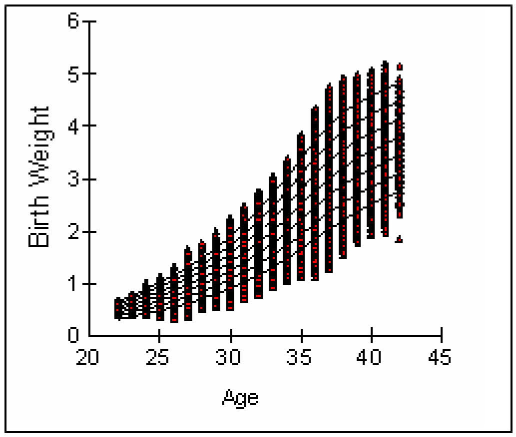


Figure 6: Plotted data with percentiles

- Surrounding LMS models were also calculated and worm plots were also analyzed, in order to check for optimal fit of the model.

- Smoothed curves from the LMS models with the best worm plots were overlaid graphically to check for any noticeable differences.

- The empirical curves were plotted on top of the smoothed curves to check for goodness of fit.

- Z-scores of the infants' measurements from the curve samples were calculated. The means and standard deviation were checked to determine if they were close to one and zero respectively.

B. Validation of growth curves

The second random sample data set was used to validate the curves that were created using the first random sample. The smoothed curves were used to calculate the growth measurement z-scores for each subject in the validation data set. The z-scores were then plotted and analyzed to confirm that the mean was close to zero and standard deviation was close to one as expected. *T*-tests were used to determine if the mean z-scores were significantly different from zero. The z-scores were also reviewed to assure that they did not vary by gestational age to test for bias. The number of infants in the validation data set that were within the SGA, AGA, and LGA percentile classifications were calculated to check that the percentages were close to the expected percentages of 10% (<10th percentile), 80% (10th to 90th percentile), and 10% (>90th percentile).

C. Comparison to the Lubchenco curves

After the growth curves were created and validated, they were compared to the Lubchenco curves. These particular curves were chosen to compare against because they also contain weight, length, and head circumference measurements for preterm infants and are commonly used in the NICU setting. Another important reason was that the Lubchenco published studies contain the actual percentile values for use in comparison.

A number of methods were used to compare the curves. As performed in the preliminary studies (Thomas, Peabody, Turnier et al., 2000), the curves were first compared graphically by overlaying the Lubchenco percentiles against the new Pediatrix percentiles. Differences in selected points on the graphs were calculated and presented in tables. Next, each infant in the Pediatrix data set was classified as SGA and LGA based on the Lubchenco curves. This information was analyzed by gestational age and gender to determine if the percentages were in line with the expected 10% for SGA and 10% for LGA. As a summary, the number of infants that were misclassified as SGA, AGA and

LGA on the Lubchenco curves compared to the new Pediatrix curves was calculated for each measurement and gender.

<u>Analysis for Specific Aim II:</u>

The relationship of gender and race of the infants with birth size (weight, length and head circumference), and gestational age was examined. The growth status of the various groups at each gestational age were compared.

Comparisons were made between the racial/ethnic groups represented in the data set in significant number (white, Hispanic and black). The z-scores, means, medians and standard deviation of the birth size measurements by gestational age were calculated by gender and race. As mentioned earlier, the racial designation was determined based on the race/ethnic group of the mother.

A. Gender analysis

The differences between the new male and female Pediatrix percentile curves were analyzed in order to find out whether gender-specific curves were warranted. The curves for males and females were overlaid graphically to visually show the differences in the curves. Selected points on the curves were compared in tables in order to quantify the differences.

B. Racial group analysis

The differences in the racial groups of birth size and gestational age were investigated for males and females separately. Four age groups (23-26, 27-31, 32-36, and 37-41) were used in all of the analyses by age groups. Gestational ages of 22 and 42 weeks were omitted from this analysis due to the small sample size at 22 weeks and the drop-off in size at 42 weeks. The racial groups analyzed were the black, Hispanic and white groups since these each had a significant number of infants.

The percentage of infants in the racial groups at the various age groups were calculated to comment on the differences in gestational age by racial group. The mean birth sizes were then compared. The mean weight, length, and head circumference were each calculated by age and racial group. Analysis of variance (ANOVA) was performed to look for statistical significance between the groups. The Bonferroni (Dunn) *t*-tests were also performed in order to show pairwise comparisons among the birth sizes of the various groups.

In order to also consider the distributions within the racial groups of the validation sample, the z-scores by age and racial group were calculated and displayed as bar-charts. In general terms, z-scores are the number of standard deviations from the mean. Z-scores provide a unitless quantity for easy comparison and since they are a function of the standard deviation provide information regarding the distribution of the sample. The z-scores in this study were calculated using the L, M, and S values from the new Pediatrix curves. Therefore, the z-score values are relative to the new curves but provide information regarding the distribution of the population and should be considered in conjunction with the means as they offer more information than the means alone. The z-scores were converted to percentiles to estimate clinical significance. To further assess clinical significance, the prevalence of SGA and LGA for the birth size measurements (weight, length, and head circumference) was computed by gender and racial group.

Logistic regression analysis was used to determine if the racial groups were predictors of SGA for all three birth size measurements for both genders. Hispanic and black infants were compared to white infants and the odds ratios were presented for this analysis. Logistic regression techniques were also used to investigate the differences that were found in the odds of SGA measurements to determine if they could be explained based on the maternal characteristics in the data set (smoking, preeclampsia/eclampsia, maternal age, insulin use, antenatal steroids, and diabetes). For the characteristics that were found to be predictors, the prevalence within the racial groups by age group of these characteristics was calculated. The significant differences in prevalence among the racial groups was computed by the use of the chi-square calculation (alpha=0.0125).

APGAR scores at 1 and 5 minutes of life are an indication of the sickness level of the infants. Therefore, the APGAR scores were analyzed to determine if there were differences in prevalence of low scores. APGAR at 1 minute of less than 3 and APGAR at 5 minutes of less than 5 were used to define “sick” infants and these were analyzed for differences between infants classified as SGA, AGA, and LGA. These scores were also analyzed by racial group via chi-squared to determine if there were significant differences in the scores by racial group.

CHAPTER 5. RESULTS AND DISCUSSION

Results of Data Preparation

In preparing the data for analysis and creation of the growth curves, application of the exclusions, inclusions, and data cleaning procedures removed a total of 34.2% of the infants from the original data set. The details of these results are outlined in Table 6. In the first step, 8% of the records were removed from the sample due to missing measurements (birth weight, birth length, or head circumference) or unknown/ambiguous gender. Another 24.5% of the original data set was removed due to characteristics (death in the NICU, multiple births, or congenital anomalies) which were likely to negatively impact fetal growth. Congenital anomalies were defined as those with an ICD-9 code for anomalies in the International Classification of Diseases published by the World Health Organization.

The remaining data set was then separated by gender and Tukey's method (Tukey, 1977) was used to remove outliers from each of the data sets. Those records with measurements that were two times the interquartile range below the first quartile and above the third quartile for the infant's gestational age were removed. In total, Tukey's method removed 1.6% of the infants from the original data set. Tukey's method has been used primarily in birth weight studies and the number of infants removed as outliers in weight in this study was consistent with the Arbuckle (Arbuckle, Wilkins, and Sherman, 1993), Beeby (Beeby, Bhutap, and Taylor, 1996), and Bonelie (Bonellie, Chalmers, Gray et al., 2008) studies.

Table 6: Results of data preparation

	# Infants remaining	# Infants removed	% of original data set
Original data set	391,681		
Unknown birth weight	389,596	2,085	0.53
Unknown birth length	364,117	25,479	6.51
Unknown head circumference	360,362	3,755	0.96
Ambiguous or unknown gender	360,147	215	0.05
Multiple births	301,390	58,757	15.00
Died in the NICU	294,817	6,573	1.68
Congenital anomalies	264,185	30,632	7.82
Females:	113,042		
Weight < 1st quartile – 2*IQR	112,867	175	0.04
Weight > 3rd quartile + 2*IQR	112,000	867	0.22
Length < 1st quartile - 2*IQR	110,920	1,080	0.28
Length > 3rd quartile + 2*IQR	110,820	100	0.03
Head circumference < 1st quartile - 2*IQR	110,515	305	0.08
Head circumference > 3rd quartile + 2*IQR	110,290	225	0.06
Males:	151,143		
Weight < 1st quartile – 2*IQR	150,896	247	0.06
Weight > 3rd quartile + 2*IQR	149,878	1,018	0.26
Length < 1st quartile - 2*IQR	148,416	1,462	0.37
Length > 3rd quartile + 2*IQR	148,231	185	0.05
Head circumference < 1st quartile - 2*IQR	147,852	379	0.10
Head circumference > 3rd quartile + 2*IQR	147,565	287	0.07
Males & females	257,855	133,826	34.17

IQR refers to interquartile range

The data was then randomly apportioned into data sets for curve creation and validation. The gender-specific data were stratified by race, gestational age, and state of facility. Those infants in the data set with a race of “missing”, ”unknown data”, “other”, “Pacific Islander”, “Asian” or “American/Alaska Native” were grouped together into an “other” category. The SAS SurveySelect procedure with the “simple random sampling” option was used to create the two random samples with the specified stratifications for each gender. The simple random sampling method selects observations independently within each stratum and with equal probability. The number of infants within each stratum for each sample was determined by dividing the number of infants within a strata in half. If there were an odd number of infants within a stratum, the number of infants assigned to the curve sample was rounded up. For example, if there were seven infants that were male, black, 26 weeks of age, and from New Jersey - four were assigned to the curve sample and three were assigned to the validation sample. This resulted in the curve samples having a slightly greater number of infants than the validation samples. Table 7 displays the number of infants in each sample. The samples used to create the growth curves are referred to as the "curve" samples while the samples used to validate the curves and used in other analysis are referred to as the "validation" samples.

Table 7: Final random sample data sets

Data samples	# Infants
Female curve sample	55,721
Female validation sample	54,569
Male curve sample	74,390
Male validation sample	73,175

The racial distribution of the curve sample was compared to the demographics of the births in 2005 in the United States in Table 8. The curve sample was very similar in racial demographics to recent births across the United States; thus the growth curves should be fairly representative of the racial demographics of the nation.

Table 8: Racial demographic of the curve samples compared to United States births in 2005

	Black (%)	Hispanic (%)	Other (%)	White (%)
U.S. 2005 (*National Vital Statistics System*, 2008)	14.11	23.81	6.99	55.09
Pediatrix curve samples	15.72	24.42	9.32	50.54

The curve samples were analyzed for distribution of infants at various gestational ages by race. It was found (Table 9) that there were a larger percentage of black infants in the younger age ranges. This finding points out that more black infants were born earlier in pregnancies. This distribution may influence the curves in that at the earlier ages, there were more black infants. However, it would not be appropriate to create race-specific growth curves because the reasons for differences in fetal growth of racial groups were unknown (Alexander, Kogan, and Himes, 1999). The differences may be caused partially by socio-economic status and should not be accepted as the norm (Alexander, Kogan, and Himes, 1999; Reagan and Salsberry, 2005).

Table 9: Distribution of the curve samples by racial group

Age group (weeks)	Females				Males			
	Black (%)	Hispanic (%)	Other (%)	White (%)	Black (%)	Hispanic (%)	Other (%)	White (%)
23-26	31.18	19.93	7.81	41.09	28.00	21.93	9.11	40.96
27-31	24.32	20.52	8.19	46.97	21.07	22.17	8.28	48.48
32-36	16.84	22.27	8.56	52.33	13.99	23.32	8.94	53.76
37-41	13.79	27.66	10.29	48.27	12.96	26.28	10.06	50.70

<u>Results for Specific Aim I:</u>

A. Smoothing of curves

The birth size measurements were analyzed to create gender-specific birth weight, head circumference, and length growth curves for preterm infants. As discussed in the methods, the LMS method (lmsChartMaker Pro, version 2.7; Medical Research Council, UK) was used to smooth the curves. Worm plots, z-scores and comparison to empirical curves were used to choose the curve with the best fit. The individual sets of curves are discussed below.

The curves were fit using the LMS method with the technique recommended by van Buuren and Fredriks (van Buuren and Fredriks, 2001). The EDF value for the M curve was increased until the worm plots passed through the origin. The EDF value for the S curve was increased until the worm plots had slopes close to zero. Lastly, the EDF for the L curve was increased until any quadratic shapes (U shapes) in the worm plots were decreased. An example of the progression of worm plots for the various LMS models tested for female birth weight percentiles can be seen in Appendix 2, Figures 55-56.

Once a model was found using the van Buuren technique described above, the various EDF values were raised and lowered to double-check the surrounding models for optimal fit. In total, 51 LMS models (Appendix 3, Table 54) were created and completely analyzed. It was found that significantly raising the EDF value for the M curve improved the fit of the smoothed curve to the empirical data. However, by increasing the EDF value for the M curve to the point where a difference in the growth curves was observed, the growth curve then seemed to overfit the data as unnecessary ripples were introduced to the smoothed curve to make it fit closer to the empirical data (van Buuren and Fredriks, 2001). These ripples were small aberrations in the curve that caused it to be less smooth. The ripples were undesirable because then curves may have slight downturns which are due to variations in the data and are not true representations of average growth.

The main difference between the models with the lower and higher EDF for the M curve was found at the tails of curve (22 weeks and 42 weeks of gestational age). At 22 weeks, there were a small number of infants (n< 20); so, this finding may indicate that the sample size was simply too small at this gestational age. The empirical data for 42 week old infants showed smaller sizes than the 41 week old infants which may indicate that 42 week old infants in a NICU have health issues that affect their weight. Weeks 22 through 42 are presented in all of the growth curves below, however, when presented for practical use, these two weeks should be omitted from the graphs because they may not be accurate representations of the optimal birth sizes of infants at those gestational ages.

The LMS method uses a coding system to identify the models of the curves. The coding represents the EDF values of the L, M, and S curves followed by the age scale (original “o”, rescaled “r” or transformed “t”). So, for a model where L=5, M=12, and S=7 and the rescaled option was used, the model would be coded as 051207r. This coding is used below to identify the specific models of the smoothed curves.

<u>Female curves:</u>

Birth weight

When fitting the curves, the best curves as determined by the worm plots for female birth weight had the LMS codes of 031309r and 032009r. These two curves were very similar when overlaid graphically. The primary difference was at the tails of the curves. The percentiles of the 031309r model compared to the 032009r model had larger weights at 22 weeks, on average, of approximately 0.002 kg and at 42 weeks of 0.035 kg.

The z-scores of the curve sample (Table 10) were calculated for each potential curve. The 032009r model had z-scores with means closer to zero which indicates a close fit to the empirical data. However, the 032009r model actually overfit the data resulting in ripples in the smoothed curve. Based on this information, 031309r was determined to be the best fit overall.

The final smoothed curves for female birth weights (Figure 7) are displayed below and the percentile values are available in Appendix 4, Table 55. The associated worm plot is displayed in Figure 8. The fit of the curves was confirmed by overlaying the empirical curves on top of the smoothed curves (Figure 9). The empirical and smoothed curves were very similar with the main differences at the highest and lowest gestational ages and in the 3rd and 97th percentiles.

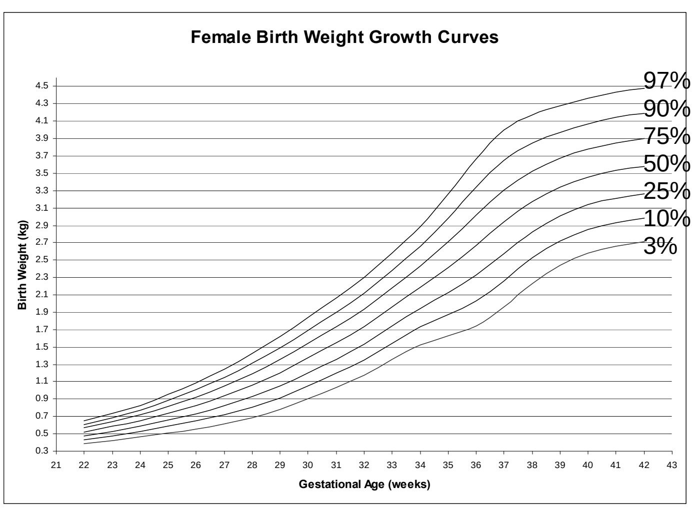


Figure 7: Female birth weight curves (LMS=031309r)

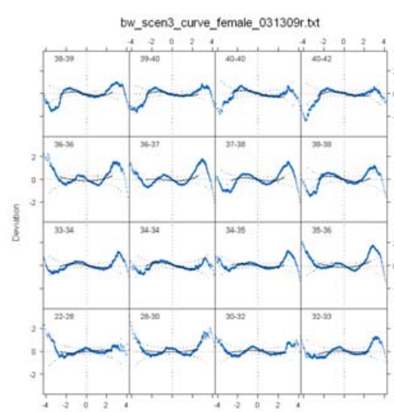


Figure 8: Worm plot for female birth weight curves (LMS=031309r)

Table 10: Female birth weight: Z-scores of curve sample calculated based on two smoothed curves

Gestational Age (weeks)	LMS = 031309r		LMS = 032009r	
	Z-scores		Z-scores	
	Mean	Std Dev	Mean	Std Dev
22	-0.01	0.99	0.01	1.00
23	0.05	0.97	0.01	0.97
24	-0.01	0.92	0.03	0.93
25	0.02	1.03	-0.01	1.02
26	-0.02	1.02	-0.01	1.03
27	0.00	1.00	0.00	1.00
28	-0.01	1.03	-0.01	1.03
29	-0.02	1.01	0.00	1.01
30	0.01	0.98	0.01	0.98
31	0.01	0.99	0.00	0.99
32	-0.01	1.01	0.00	1.01
33	0.00	1.01	0.00	1.01
34	0.01	0.98	0.00	0.98
35	0.00	1.01	0.00	1.01
36	0.00	1.00	0.00	1.00
37	0.00	1.01	0.00	1.01
38	0.01	0.99	0.00	0.99
39	-0.01	1.00	0.00	1.00
40	0.00	0.98	0.00	0.98
41	0.02	1.01	0.01	1.02
42	-0.18	1.13	-0.10	1.14

Shaded cells indicate z-score means above 0.05 or below -0.05

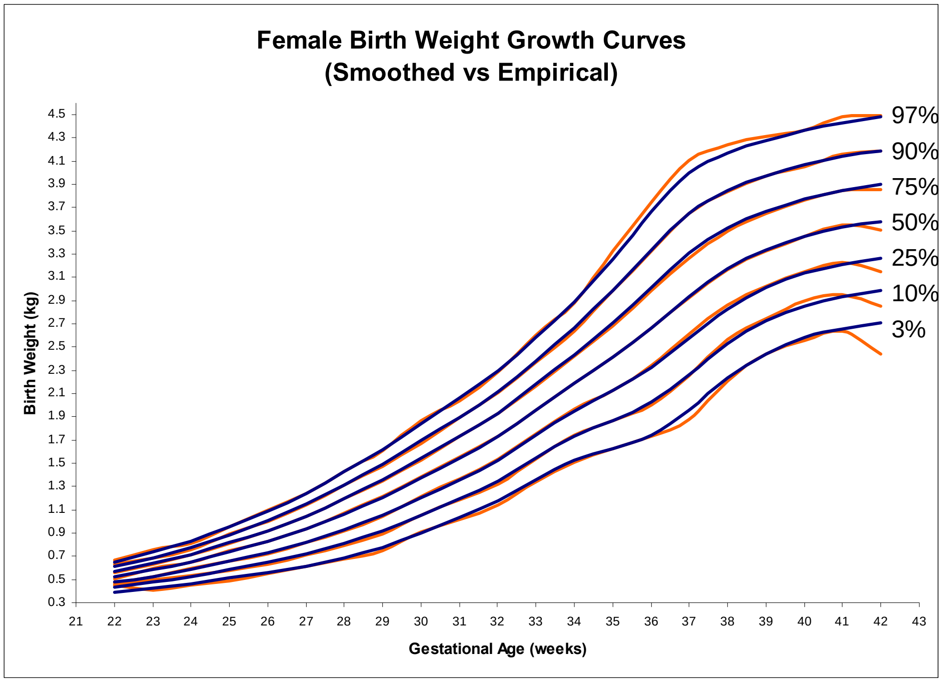


Figure 9: Female birth weight: Smoothed (blue) versus empirical (orange) growth curves

Birth length

When fitting the curves, the best curves as determined by the worm plots for female birth length had the LMS codes of 070707r and 072007r. These two curves were very similar when overlaid graphically. The primary difference was at the tails of the curves. The percentiles of the 070707r model compared to the 072007r model had smaller lengths at 22 weeks, on average, of approximately 0.3 cm and at 42 weeks were greater by an average of 0.5 cm.

The z-scores of the curve sample (Table 11) were calculated for each potential curve. The 072007r model had z-scores with means closer to zero which indicates a close fit to the empirical data. However, the 072007r model actually overfit the data resulting in ripples in the smoothed curve. Based on this information, 070707r was determined to be the best fit overall.

The final smoothed curves for female birth lengths (Figure 10) are displayed below and the percentile values are available in Appendix 4, Table 56. The associated worm plot is displayed in Figure 11. The fit of the curves was confirmed by overlaying the empirical curves on top of the smoothed curves (Figure 12). The empirical and smoothed curves were very similar with the main differences at the highest and lowest gestational ages and in the 3rd and 97th percentiles.

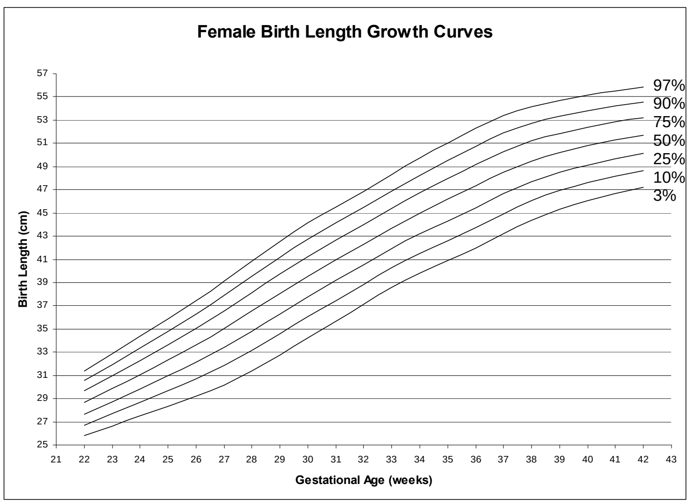


Figure 10: Female birth length curves (LMS=070707r)

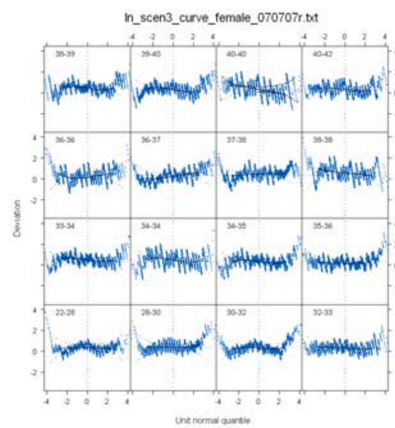


Figure 11: Worm plot for female birth length curves (LMS=070707r)

Table 11: Female birth length: Z-scores of curve sample calculated based on two smoothed curves

Gestational Age (weeks)	LMS = 070707r		LMS = 072007r	
	Z-scores		Z-scores	
	Mean	Std Dev	Mean	Std Dev
22	0.23	0.83	0.03	0.81
23	0.03	1.08	-0.01	1.07
24	0.00	0.93	0.01	0.93
25	0.02	1.02	0.00	1.02
26	-0.06	1.01	-0.01	1.02
27	0.03	0.99	0.01	0.98
28	-0.01	1.01	0.00	1.01
29	-0.03	1.02	-0.01	1.03
30	0.05	1.00	0.00	0.99
31	0.00	1.00	0.00	1.01
32	-0.04	1.00	0.00	1.00
33	0.02	0.99	0.00	0.99
34	0.02	0.99	0.00	0.99
35	-0.03	1.00	0.00	1.00
36	-0.02	1.01	0.00	1.01
37	-0.01	1.04	-0.01	1.04
38	0.03	0.98	0.00	0.99
39	-0.01	1.01	0.00	1.00
40	0.00	0.97	0.00	0.97
41	0.04	1.02	0.02	1.02
42	-0.33	1.10	-0.11	1.12

Shaded cells indicate z-score means above 0.05 or below -0.05

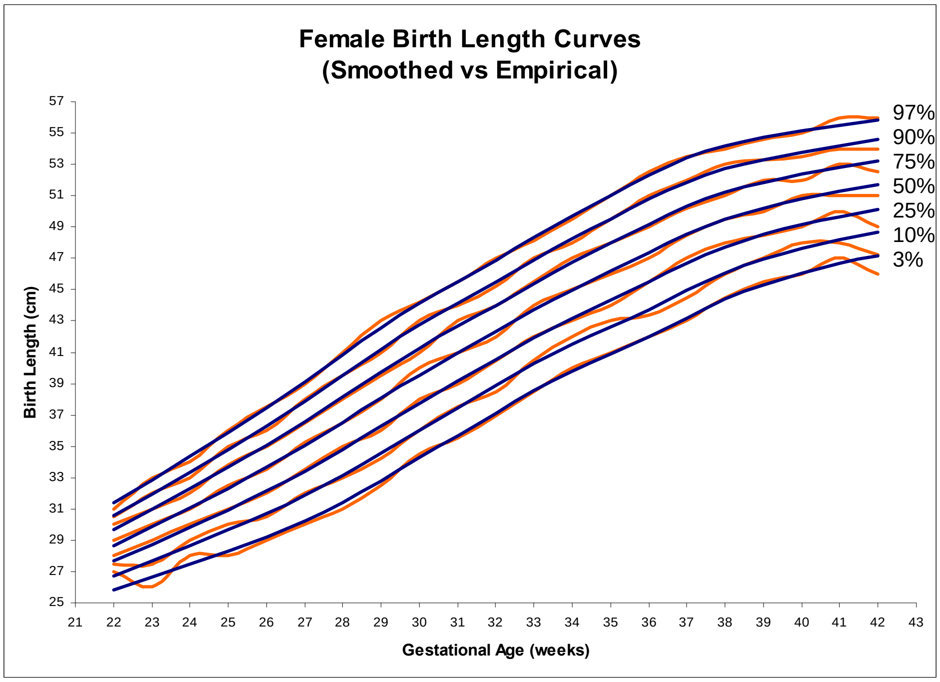


Figure 12: Female birth length: Smoothed (blue) versus empirical (orange) growth curves

Head circumference

When fitting the curves, the best curves as determined by the worm plots for female head circumference had the LMS codes of 040808r and 042008r. These two curves were very similar when overlaid graphically. The primary difference was at the tails of the curves. The percentiles of the 040808r model compared to the 042008r model had smaller head circumferences at 22 weeks by an average of 0.1 cm and at 42 weeks were greater ranging from 0.04 – 0.1 cm depending on percentile.

The z-scores of the curve sample (Table 12) were calculated for each potential curve. The 042008r model had z-scores with means closer to zero which indicates a close fit to the empirical data. However, the 040808r model actually overfit the data resulting in ripples in the smoothed curve. Based on this information, 040808r was determined to be the best fit overall.

The final smoothed curves for female head circumference (Figure 13) are displayed below and the percentile values are available in Appendix 4, Table 57. The associated worm plot is displayed in Figure 14. The fit of the curves was confirmed by overlaying the empirical curves on top of the smoothed curves (Figure 15). The empirical and smoothed curves were very similar with the main differences at the highest and lowest gestational ages and in the 3rd and 97th percentiles.

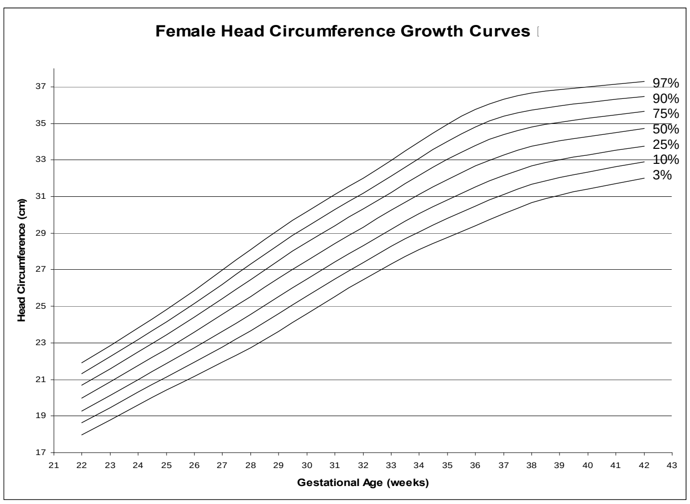


Figure 13: Female head circumference curves (LMS=040808r)

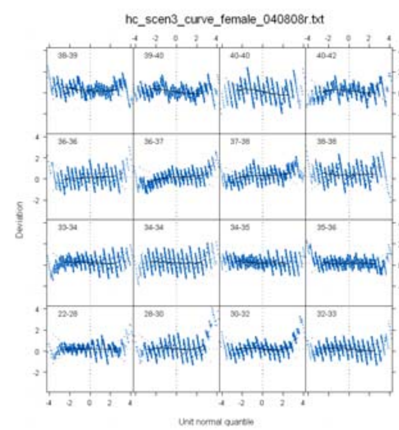


Figure 14: Worm plot for female head circumference curves (LMS=040808r)

Table 12: Female head circumference: Z-scores of curve sample calculated based on two smoothed curves

Gestational Age (weeks)	LMS = 040808r		LMS = 042008r	
	Z-scores		Z-scores	
	Mean	Std Dev	Mean	Std Dev
22	0.12	1.00	0.01	1.00
23	-0.01	1.11	-0.02	1.11
24	-0.02	0.96	0.01	0.96
25	0.04	1.01	0.00	1.01
26	-0.02	0.97	0.00	0.97
27	0.00	1.00	0.00	1.00
28	0.00	1.03	0.00	1.03
29	-0.02	0.99	0.00	0.99
30	0.02	1.01	0.00	1.01
31	0.00	1.00	0.00	1.00
32	-0.01	1.00	0.00	1.00
33	0.00	0.99	0.00	0.99
34	0.00	1.00	0.00	1.00
35	0.00	0.99	0.00	0.99
36	-0.01	1.01	0.00	1.01
37	0.00	1.03	0.00	1.03
38	0.03	1.00	0.00	1.01
39	-0.01	0.98	0.00	0.98
40	-0.01	0.99	0.00	0.98
41	0.03	1.01	0.00	1.01
42	-0.07	1.09	-0.02	1.08

Shaded cells indicate z-score means above 0.05 or below -0.05

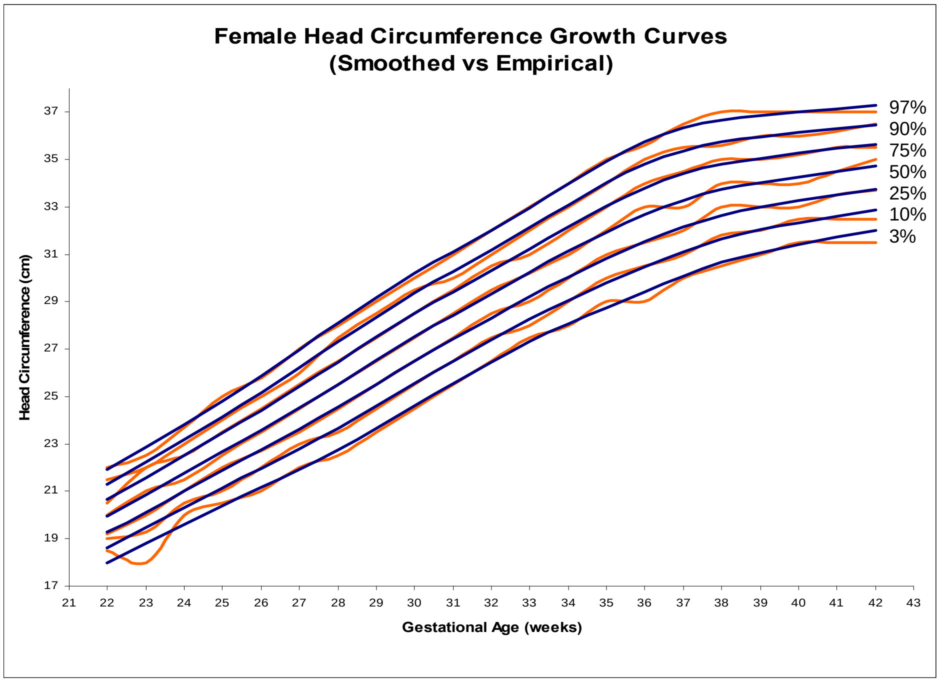


Figure 15: Female head circumference: Smoothed (blue) versus empirical (orange) growth curves

<u>Male curves:</u>

Birth weight

When fitting the curves, the best curves as determined by the worm plots for male birth weight had the LMS codes of 031208r and 032008r. These two curves were very similar when overlaid graphically. The primary difference was at the tails of the curves. The percentiles of the 031208r model compared to the 032008r model had smaller birth weights at 22 weeks ranging from 0.007 – 0.015 kg depending on percentile and at 42 weeks were larger by an average of 0.048 kg.

The z-scores of the curve sample (Table 13) were calculated for each potential curve. The 032008r model had z-scores with means closer to zero which indicates a close fit to the empirical data. However, the 031208r model actually overfit the data resulting in ripples in the smoothed curve. Based on this information, 031208r was determined to be the best fit overall.

The final smoothed curves for male birth weights (Figure 16) are displayed below and the percentile values are available in Appendix 4, Table 58. The associated worm plot is displayed in Figure 17. The fit of the curves was confirmed by overlaying the empirical curves on top of the smoothed curves (Figure 18). The empirical and smoothed curves were very similar with the main differences at the highest and lowest gestational ages and in the 3rd and 97th percentiles.

Figure 16: Male birth weight curves (LMS=031208r)

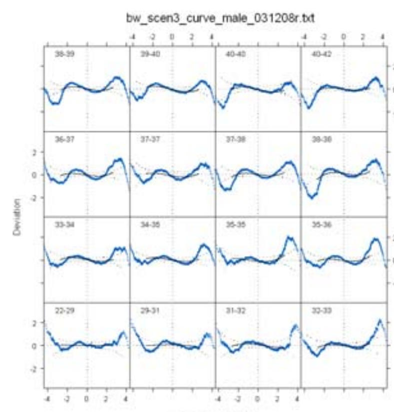


Figure 17: Worm plot for male birth weight curves (LMS=031208r)

Table 13: Male birth weight: Z-scores of curve sample calculated based on two smoothed curves

Gestational Age (weeks)	LMS = 031208r		LMS = 032008r	
	Z-scores		Z-scores	
	Mean	Std Dev	Mean	Std Dev
22	0.14	1.12	-0.02	1.08
23	0.03	0.87	0.04	0.87
24	0.01	0.98	0.01	0.98
25	0.00	0.99	0.00	0.99
26	0.01	1.03	-0.01	1.02
27	-0.02	1.01	-0.01	1.01
28	0.01	1.03	-0.01	1.03
29	0.00	0.97	0.01	0.97
30	-0.02	1.03	-0.01	1.03
31	0.02	0.98	0.01	0.98
32	-0.01	1.00	0.00	1.00
33	0.00	1.01	0.00	1.01
34	0.01	0.98	0.01	0.98
35	0.00	1.01	0.00	1.01
36	0.00	1.01	0.00	1.01
37	-0.01	1.00	0.00	0.99
38	0.01	1.01	0.00	1.02
39	0.00	0.98	0.00	0.98
40	-0.01	0.99	0.00	0.99
41	0.03	1.04	0.00	1.04
42	-0.13	1.01	-0.03	1.01

Shaded cells indicate z-score means above 0.05 or below -0.05

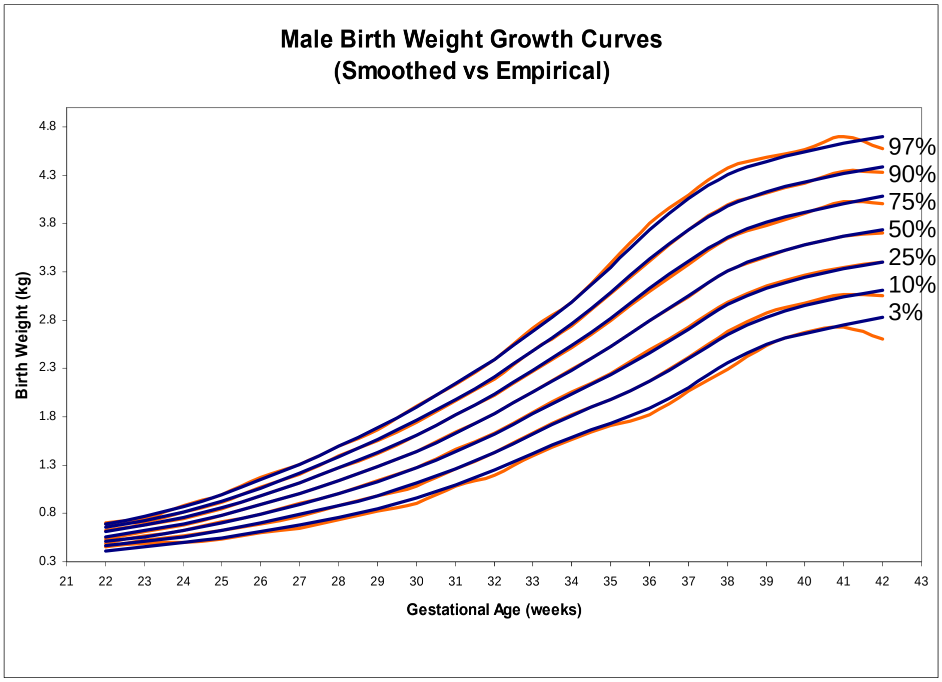


Figure 18: Male birth weight: Smoothed (blue) versus empirical (orange) growth curves

Birth length

When fitting the curves, the best curves as determined by the worm plots for male birth length had the LMS codes of 030807r and 032007r. These two curves were very similar when overlaid graphically. The primary difference was at the tails of the curves. The percentiles of the 030807r model compared to the 032007r model had longer lengths at 22 weeks of 0.08 cm on average and at 42 weeks were shorter by 0.25 cm on average.

The z-scores of the curve sample (Table 14) were calculated for each potential curve. The 032007r model had z-scores with means closer to zero which indicates a close fit to the empirical data. However, the 032007r model actually overfit the data resulting in ripples in the smoothed curve. Based on this information, 030807r was determined to be the best fit overall.

The final smoothed curves for male birth lengths (Figure 19) are displayed below and the percentile values are available in Appendix 4, Table 59. The associated worm plot is displayed in Figure 20. The fit of the curves was confirmed by overlaying the empirical curves on top of the smoothed curves (Figure 21). The empirical and smoothed curves were very similar with the main differences at the highest and lowest gestational ages and in the 3rd and 97th percentiles.

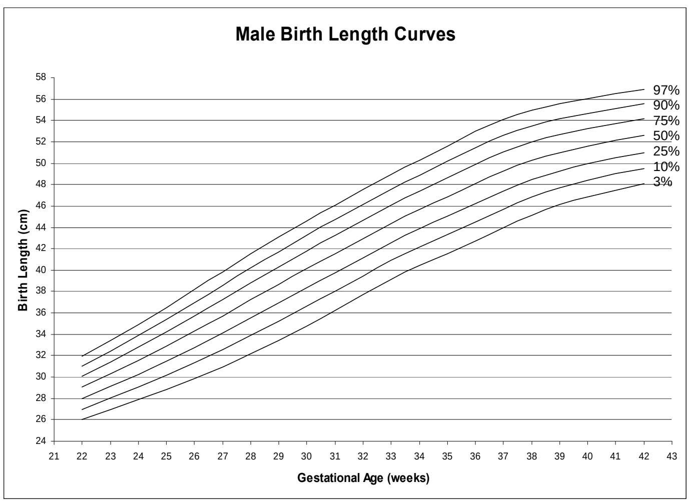


Figure 19: Male birth length curves (LMS=030807r)

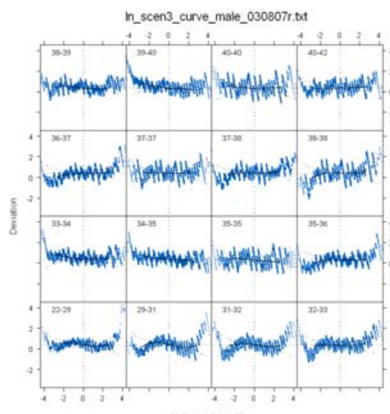


Figure 20: Worm plot for male birth length curves (LMS=030807r)

Table 14: Male birth length: Z-scores of curve sample calculated based on two smoothed curves

Gestational Age (weeks)	LMS = 030807r		LMS = 032007r	
	Z-scores		Z-scores	
	Mean	Std Dev	Mean	Std Dev
22	-0.10	1.22	-0.04	1.22
23	0.12	0.95	0.01	0.94
24	-0.02	0.96	0.01	0.96
25	-0.03	1.02	-0.01	1.02
26	0.01	1.00	0.00	1.00
27	-0.03	1.02	-0.01	1.02
28	0.04	1.01	0.00	1.01
29	0.00	0.96	0.01	0.96
30	-0.03	1.04	-0.01	1.05
31	0.03	0.97	0.01	0.97
32	-0.02	1.01	0.00	1.01
33	0.01	1.00	0.00	1.01
34	0.01	0.98	0.01	0.98
35	-0.01	0.99	0.00	0.99
36	-0.01	1.03	-0.01	1.03
37	0.01	1.00	0.00	1.00
38	0.00	1.02	0.00	1.02
39	0.00	0.97	0.01	0.97
40	0.00	1.00	0.00	1.00
41	0.01	1.02	0.00	1.02
42	-0.12	1.05	-0.01	1.04

Shaded cells indicate z-score means above 0.05 or below -0.05

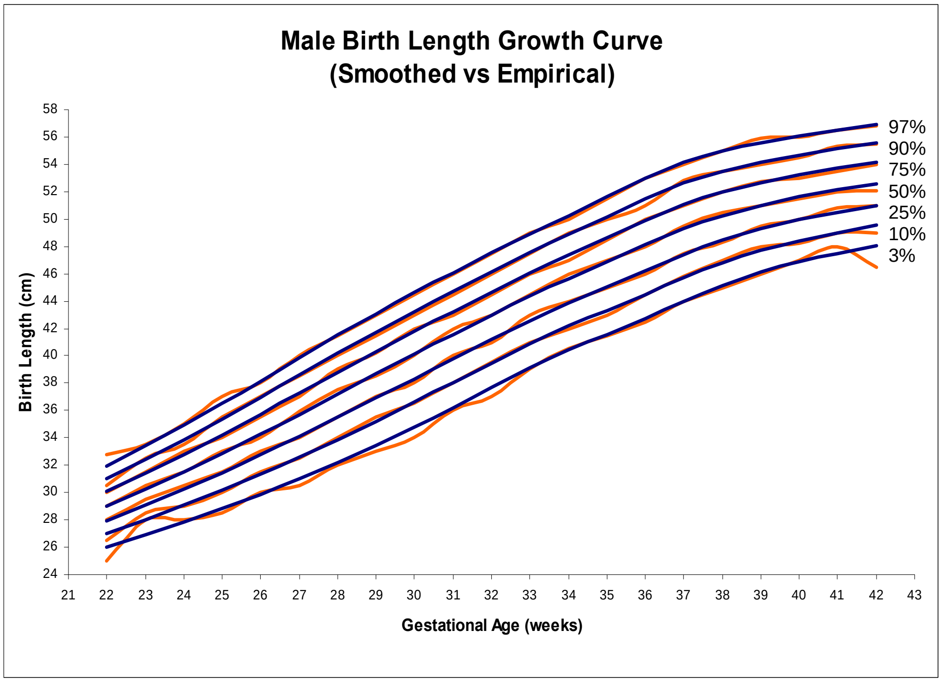


Figure 21: Male birth length: Smoothed (blue) versus empirical (orange) growth curves

Head circumference

When fitting the curves, the best curves as determined by the worm plots for male head circumference had the LMS codes of 030908r and 032011r. These two curves were very similar when overlaid graphically. The primary difference was at the tails of the curves. The percentiles of the 030908r model compared to the 032011r model had smaller head circumferences at 22 weeks ranging from 0.04 – 0.11 cm depending on percentile and at 42 weeks were larger by 0.02 cm on average.

The z-scores of the curve sample (Table 15) were calculated for each potential curve. The 032011r model had z-scores with means closer to zero which indicates a close fit to the empirical data. However, the 032011r model actually overfit the data resulting in ripples in the smoothed curve. Based on this information, 030908r was determined to be the best fit overall.

The final smoothed curves for male head circumference (Figure 22) are displayed below and the percentile values are available in Appendix 4, Table 60. The associated worm plot is displayed in Figure 23. The fit of the curves was confirmed by overlaying the empirical curves on top of the smoothed curves (Figure 24). The empirical and smoothed curves were very similar with the main differences at the highest and lowest gestational ages and in the 3rd and 97th percentiles.

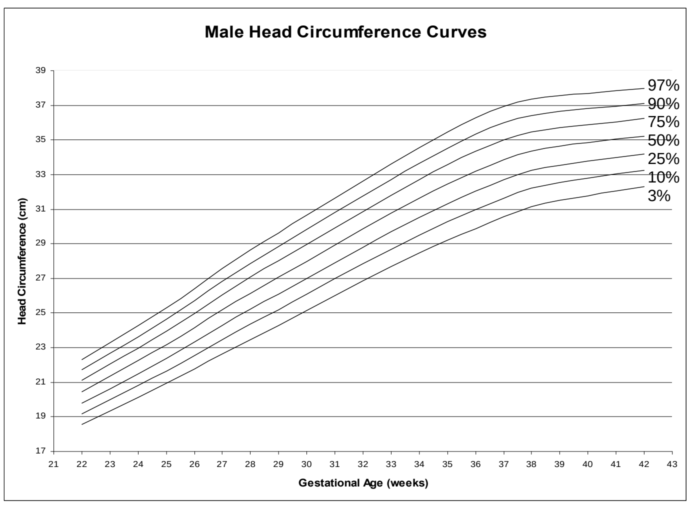


Figure 22: Male head circumference curves (LMS=030908r)

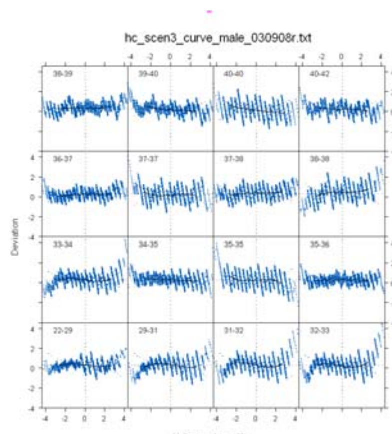


Figure 23: Worm plot for male head circumference curves (LMS=030908r)

Table 15: Male head circumference: Z-scores of curve sample calculated based on two smoothed curves

Gestational Age (weeks)	LMS = 030908r		LMS = 032011r	
	Z-scores		Z-scores	
	Mean	Std Dev	Mean	Std Dev
22	0.04	1.32	-0.04	1.29
23	0.01	0.94	0.01	0.94
24	0.01	1.04	-0.01	1.04
25	-0.02	0.96	0.00	0.97
26	-0.01	1.02	0.00	1.02
27	0.02	0.99	0.00	0.99
28	0.01	1.02	0.00	1.02
29	0.00	0.97	0.00	0.98
30	-0.02	1.02	0.00	1.02
31	0.02	0.99	0.00	0.99
32	0.00	1.00	0.00	1.00
33	0.00	1.01	0.00	1.00
34	0.00	1.00	0.00	1.00
35	0.00	0.99	0.00	0.99
36	0.00	1.01	0.00	1.01
37	-0.01	1.00	0.00	0.99
38	0.02	1.02	0.00	1.02
39	0.00	0.99	0.00	0.99
40	-0.02	0.98	0.00	0.99
41	0.03	1.00	0.00	1.01
42	-0.02	1.07	-0.01	1.07

Shaded cells indicate z-score means above 0.05 or below -0.05

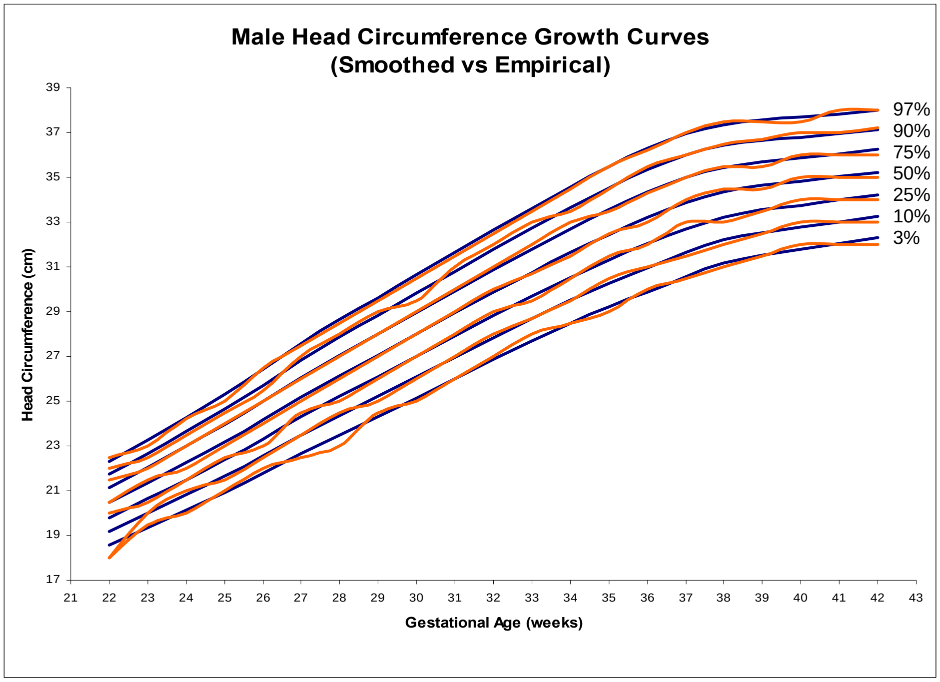


Figure 24: Male head circumference: Smoothed (blue) versus empirical (orange) growth curves

B. Validation steps

The six sets of smoothed growth curves were validated to assure that they represented the Pediatrix sample population. The z-scores and standard deviations (Tables 16-17) of the validation sample were calculated using the final growth curve L, M, and S parameters (Appendix 5, Tables 61-66). With the exception of gestational age of 22 and 42 weeks, it was found that the z-scores were close to zero and the standard deviation was close to one as expected. At 22 weeks, the sample size was small (n=11) which explains the deviations seen at this gestational age.

Table 16: Females: Z-scores of the validation set

Gestational age (weeks)	Birth weight		Birth length		Head circumference	
	Mean z-score	Standard deviation	Mean z-score	Standard deviation	Mean z-score	Standard deviation
22	0.33	1.42	0.69	1.33	0.59	1.11
23	-0.03	0.82	0.15	0.88	0.09	1.08
24	0.00	0.90	0.03	1.00	-0.03	0.96
25	-0.09	0.96	-0.03	0.98	-0.07	1.01
26	0.09	1.04	0.10	1.07	0.10	0.99
27	0.05	1.03	0.02	1.01	0.05	1.00
28	0.06	0.98	0.07	0.97	0.08	1.00
29	0.00	0.98	-0.02	0.98	0.04	1.01
30	0.00	0.99	-0.01	0.96	0.02	0.98
31	-0.03	1.00	-0.03	1.02	-0.04	1.02
32	0.00	1.00	-0.05	1.00	-0.02	1.02
33	0.03	1.01	0.01	1.00	0.01	1.03
34	-0.03	0.98	-0.02	0.98	-0.02	0.99
35	-0.04	0.97	-0.02	0.98	-0.02	0.98
36	-0.04	0.99	-0.04	1.02	-0.03	1.01
37	-0.01	1.02	0.00	1.03	0.00	1.02
38	0.03	1.00	0.04	0.98	0.02	1.00
39	0.00	0.99	0.00	1.00	0.00	0.97
40	-0.01	0.99	0.01	0.97	-0.03	0.98
41	0.02	1.03	-0.01	1.01	-0.01	1.03
42	-0.10	1.08	-0.18	0.99	-0.03	1.03
All	-0.01	1.00	0.00	1.00	-0.01	1.00

Z-scores calculated from the final smoothed curves L, M, and S values (Appendix 5)

Table 17: Males: Z-scores of the validation set

Gestational age (weeks)	Birth weight		Birth length		Head circumference	
	Mean z-score	Standard deviation	Mean z-score	Standard deviation	Mean z-score	Standard deviation
22	0.02	0.84	0.14	1.69	0.04	0.71
23	-0.04	0.72	-0.09	0.95	0.07	0.88
24	0.00	1.00	-0.01	1.02	-0.03	1.01
25	-0.03	0.99	0.00	1.03	-0.05	0.99
26	0.01	1.02	0.08	0.95	0.00	1.02
27	-0.03	1.06	0.00	1.02	-0.05	1.06
28	-0.04	1.00	-0.04	1.02	-0.04	1.05
29	-0.04	0.97	-0.02	0.95	-0.01	1.00
30	0.02	1.00	0.00	1.03	0.03	1.02
31	-0.03	0.98	0.00	0.98	-0.02	0.98
32	-0.01	1.03	0.00	1.03	-0.03	1.01
33	-0.01	1.04	-0.01	1.00	0.00	1.01
34	0.02	0.99	-0.01	0.97	0.02	0.99
35	0.02	1.02	0.01	1.01	0.04	1.02
36	0.02	1.01	0.01	1.01	0.03	0.99
37	0.02	1.01	0.00	1.00	0.02	1.01
38	-0.02	1.01	-0.01	1.01	0.00	1.02
39	0.00	0.98	-0.01	0.98	0.00	1.00
40	-0.01	0.99	-0.02	1.01	0.02	0.97
41	-0.01	1.03	0.00	1.02	0.03	1.01
42	-0.15	1.12	-0.21	1.01	-0.06	1.05
All	0.00	1.01	0.00	1.00	0.01	1.00

Z-scores calculated from the final smoothed curves L, M, and S values (Appendix 5)

T-tests were used to check that the mean z-score by age group was not significantly different from zero. An alpha of 0.0125 was used when checking the p-values for significance because there were four age groups. In Table 18, all but four p-values were not significantly differently from zero. The values (shaded cells) that were different were in the age group of 32-36 weeks. However, the mean difference was 0.02 z-score which clinically is a change of less than a one percentile difference. So, this confirms that the new sets of smoothed growth curves were representative of the population.

Table 18: P-values from *t*-test comparing the z-scores of the validation set to zero

Age group (weeks)	Female			Male		
	Birth weight p-value	Birth length p-value	Head circumference p-value	Birth weight p-value	Birth length p-value	Head circumference p-value
23-26	0.697	0.0637	0.4388	0.6613	0.2031	0.4189
27-31	0.460	0.9018	0.0898	0.0666	0.4341	0.1861
32-36	0.002	0.0004	0.0075	0.0959	0.762	0.0016
37-41	0.646	0.1014	0.6544	0.2954	0.1736	0.0386

Shaded cells are significantly different (*t*-test) from zero (alpha=0.0125).

The percentages of infants in the validation sample that were SGA (<10th percentile), AGA (between 10th and 90th percentiles) and LGA (<90th percentile) were calculated. These values (Table 19) were found to be close to the expected values which also confirms the goodness of fit of the curves.

Table 19: Percentages of infants in the validation set within the classifications

	Female			Male		
	SGA (%)	AGA (%)	LGA (%)	SGA (%)	AGA (%)	LGA (%)
Expected Value	*10.00*	*80.00*	*10.00*	*10.00*	*80.00*	*10.00*
Birth Weight	9.88	80.64	9.48	9.91	80.29	9.80
Birth Length	9.80	80.44	9.76	10.13	80.75	9.12
Head Circumference	10.19	80.86	8.95	9.93	79.42	10.65

C. Comparison to the Lubchenco curves

In order to compare the new Pediatrix curves with the Lubchenco curves, the following were analyzed: graphical overlays of the curves, specific points on the curves, and the SGA and LGA percentages. Figures 25-27 are the new Pediatrix female curves overlaid with the Lubchenco combined-gender curves and Figures 28-30 compare the new Pediatrix male curves to the Lubchenco combined-gender curves. It was found that generally the Lubchenco percentiles were larger until gestational ages 31-36 and then after 36 weeks, the Lubchenco percentiles were smaller. In each of the graphs (Figures 25-30) below, most of Lubchenco percentile curves (dashed lines) cross the new Pediatrix percentile curves (solid lines).

Included along with each set of curves, a table (Tables 20-25) is presented which shows the actual differences between the new Pediatrix percentiles and the Lubchenco curves at selected gestational ages of 26, 30, 36 and 40 weeks. The Pediatrix birth weight percentiles were different for females ranging from 26.2% smaller to 8.5% larger and for males from 21.7% smaller to 12.2% larger. The Pediatrix birth length percentiles were different for females ranging from 9.1% smaller to 4.0% larger and for males from 7.4% smaller to 5.7% larger. The Pediatrix head circumference percentiles were different for females ranging from 11.8% smaller to 1.7% larger and for males from 9.8% smaller to 3.1% larger.

<u>Females</u>:

Weight

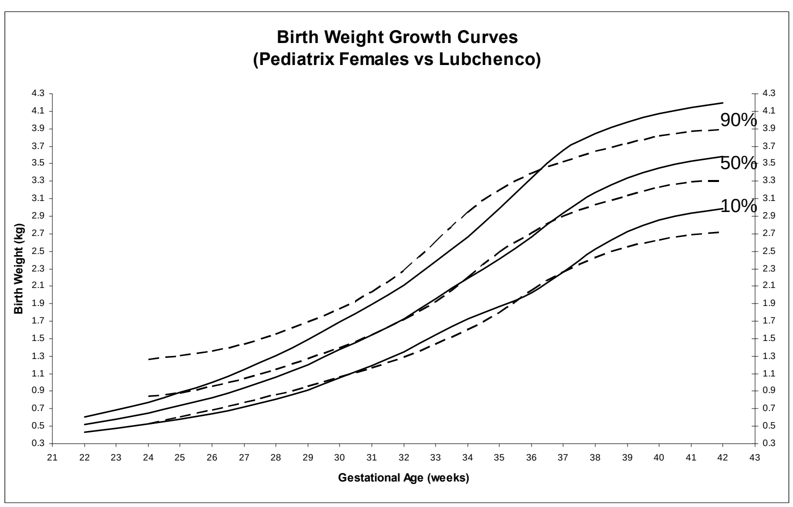


Figure 25: Female birth weight: Pediatrix (solid) compared to Lubchenco combined-gender curves (dashed, start at 24 weeks)

Table 20: Female birth weight: Comparison between Pediatrix curves and Lubchenco curves at selected percentiles and gestational ages

Gestational age (weeks)	10th percentile			50th percentile			90th percentile		
	Lub (kg)	Ped (kg)	% Ped greater	Lub (kg)	Ped (kg)	% Ped greater	Lub (kg)	Ped (kg)	% Ped greater
26	0.69	0.65	-5.81	0.96	0.83	-13.36	1.36	1.00	-26.17
30	1.06	1.05	-0.74	1.40	1.37	-1.56	1.84	1.69	-7.99
36	2.05	2.03	-1.09	2.71	2.66	-1.68	3.39	3.34	-1.49
40	2.63	2.85	8.54	3.23	3.45	6.92	3.82	4.07	6.68

Lub refers to the Lubchenco curves; Ped refers to the new curves with Pediatrix data

Females:

Length

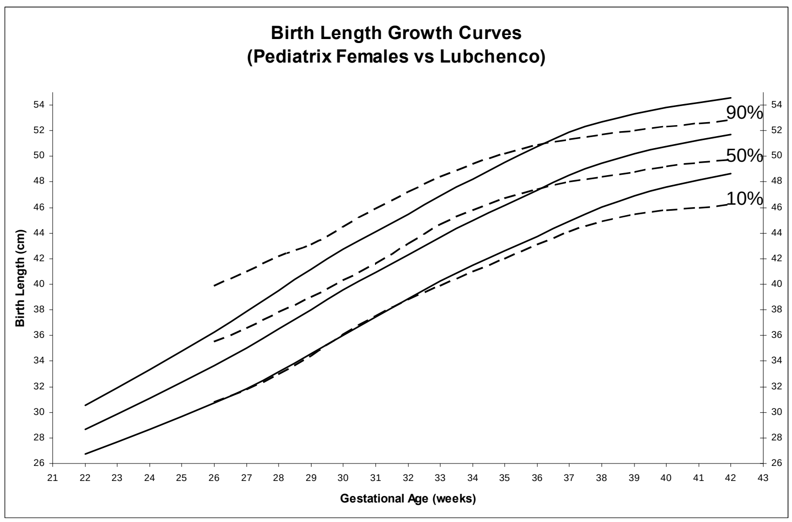


Figure 26: Female birth length: Pediatrix (solid) compared to Lubchenco combined-gender curves (dashed, start at 26 weeks)

Table 21: Female birth length: Comparison between Pediatrix curves and Lubchenco curves at selected percentiles and gestational ages

Gestational age (weeks)	10th percentile			50th percentile			90th percentile		
	Lub (cm)	Ped (cm)	% Ped greater	Lub (cm)	Ped (cm)	% Ped greater	Lub (cm)	Ped (cm)	% Ped greater
26	30.8	30.7	-0.32	35.5	33.6	-5.25	39.9	36.3	-9.05
30	36.1	36.0	-0.16	40.3	39.5	-1.87	44.5	42.7	-3.98
36	43.1	43.7	1.48	47.4	47.4	-0.09	50.9	50.8	-0.24
40	45.8	47.6	3.96	49.2	50.8	3.23	52.3	53.8	2.85

Lub refers to the Lubchenco curves; Ped refers to the new curves with Pediatrix data

Females:

Head Circum.

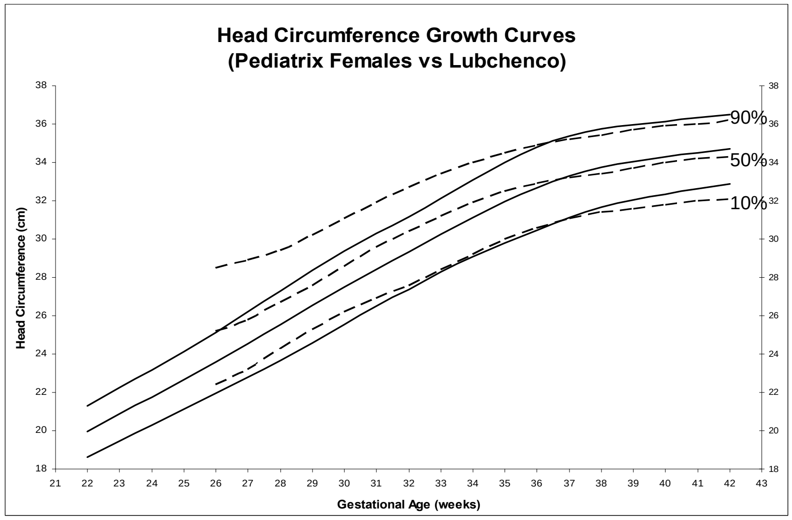


Figure 27: Female head circumference: Pediatrix (solid) compared to Lubchenco combined-gender curves (dashed, start at 26 weeks)

Table 22: Female head circumference: Comparison between Pediatrix curves and Lubchenco curves at selected percentiles and gestational ages

Gestational age (weeks)	10th percentile			50th percentile			90th percentile		
	Lub (cm)	Ped (cm)	% Ped greater	Lub (cm)	Ped (cm)	% Ped greater	Lub (cm)	Ped (cm)	% Ped greater
26	22.4	22.0	-1.97	25.2	23.6	-6.41	28.5	25.1	-11.77
30	26.2	25.6	-2.47	28.6	27.5	-3.82	31.1	29.4	-5.60
36	30.6	30.5	-0.44	32.9	32.7	-0.68	34.9	34.8	-0.30
40	31.8	32.3	1.69	34.0	34.3	0.80	35.9	36.1	0.68

Lub refers to the Lubchenco curves; Ped refers to the new curves with Pediatrix data

Males:

Weight

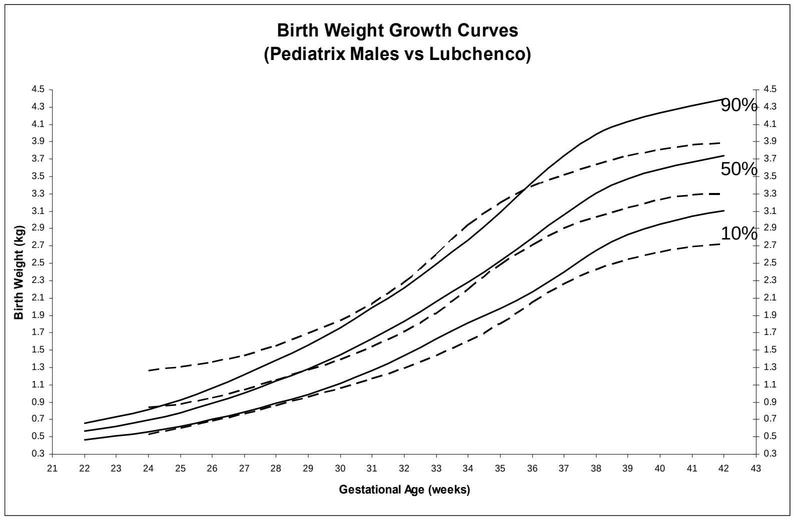


Figure 28: Male birth weight: Pediatrix (solid) compared to Lubchenco combined-gender curves (dashed, start at 24 weeks)

Table 23: Male birth weight: Comparison between Pediatrix curves and Lubchenco curves at selected percentiles and gestational ages

Gestational age (weeks)	10th percentile			50th percentile			90th percentile		
	Lub (kg)	Ped (kg)	% Ped greater	Lub (kg)	Ped (kg)	% Ped greater	Lub (kg)	Ped (kg)	% Ped greater
26	0.69	0.70	2.79	0.96	0.89	-6.83	1.36	1.07	-21.66
30	1.06	1.11	5.13	1.40	1.44	3.47	1.84	1.76	-4.31
36	2.05	2.17	5.84	2.71	2.79	3.02	3.39	3.43	1.24
40	2.63	2.95	12.17	3.23	3.58	10.79	3.82	4.23	10.92

Lub refers to the Lubchenco curves; Ped refers to the new curves with Pediatrix data

Males:

Length

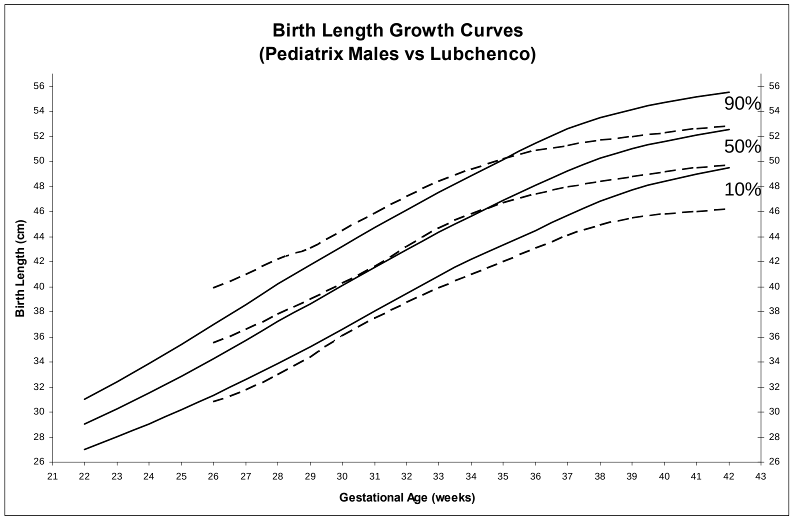


Figure 29: Male birth length: Pediatrix(solid) compared to Lubchenco combined-gender curves (dashed, start at 26 weeks)

Table 24: Male birth length: Comparison between Pediatrix curves and Lubchenco curves at selected percentiles and gestational ages

Gestational age (weeks)	10th percentile			50th percentile			90th percentile		
	Lub (cm)	Ped (cm)	% Ped greater	Lub (cm)	Ped (cm)	% Ped greater	Lub (cm)	Ped (cm)	% Ped greater
26	30.8	31.3	1.72	35.5	34.3	-3.48	39.9	37.0	-7.36
30	36.1	36.6	1.39	40.3	40.1	-0.47	44.5	43.2	-2.83
36	43.1	44.5	3.20	47.4	48.1	1.48	50.9	51.5	1.13
40	45.8	48.4	5.72	49.2	51.6	4.90	52.3	54.7	4.56

Lub refers to the Lubchenco curves; Ped refers to the new curves with Pediatrix data

Males:

Head Circum.

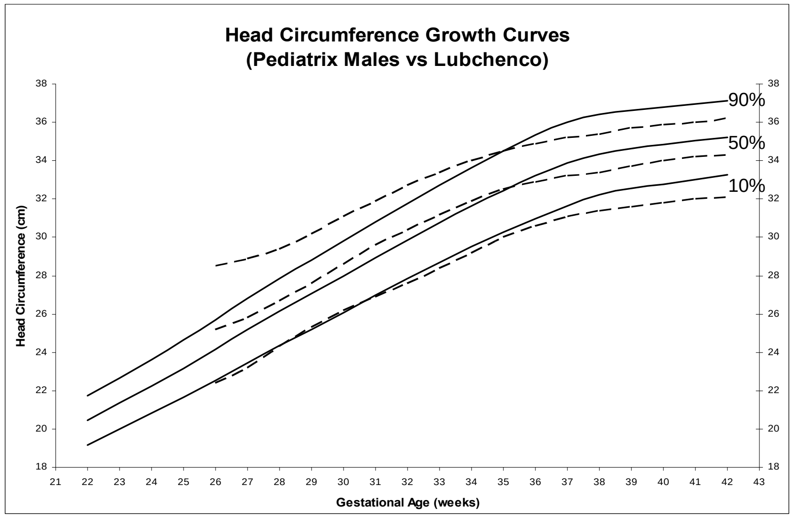


Figure 30: Male head circumference: Pediatrix(solid) compared to Lubchenco combined-gender curves (dashed, start at 26 weeks)

Table 25: Male head circumference: Comparison between Pediatrix curves and Lubchenco curves at selected percentiles and gestational ages

Gestational age (weeks)	10th percentile			50th percentile			90th percentile (cm)		
	Lub (cm)	Ped (cm)	% Ped greater	Lub (cm)	Ped (cm)	% Ped greater	Lub (cm)	Ped (cm)	% Ped greater
26	22.4	22.5	0.67	25.2	24.2	-4.11	28.5	25.7	-9.78
30	26.2	26.1	-0.50	28.6	28.0	-2.14	31.1	29.8	-4.15
36	30.6	31.0	1.23	32.9	33.2	0.95	34.9	35.3	1.26
40	31.8	32.8	3.07	34.0	34.8	2.45	35.9	36.8	2.51

Lub refers to the Lubchenco curves; Ped refers to the new curves with Pediatrix data

In order to check for the clinical significance in the differences between the new Pediatrix curves and the Lubchenco curves, the Pediatrix infants were classified as SGA (<10th percentile), AGA (10th to 90th percentiles), and LGA (>90th percentile) based on the Lubchenco curves (using Battaglia and Lubchenco's definition of the classifications) (Battaglia and Lubchenco, 1967). The percentages of male and female infants that were classified as SGA and LGA for birth weight, length and head circumference are displayed in Figures 31-33.

At most gestational ages and for all three measurements, the number of SGA males was underestimated by 40-50% or more. For female infants, the SGA classification was also often underestimated but at the younger ages was overestimated by as much as 80%. In the LGA category, males and females greater than 36 weeks of age were extremely overestimated by up to 400%. At earlier than 36 weeks of age, LGA was underestimated by up to 90% with females being especially underestimated.

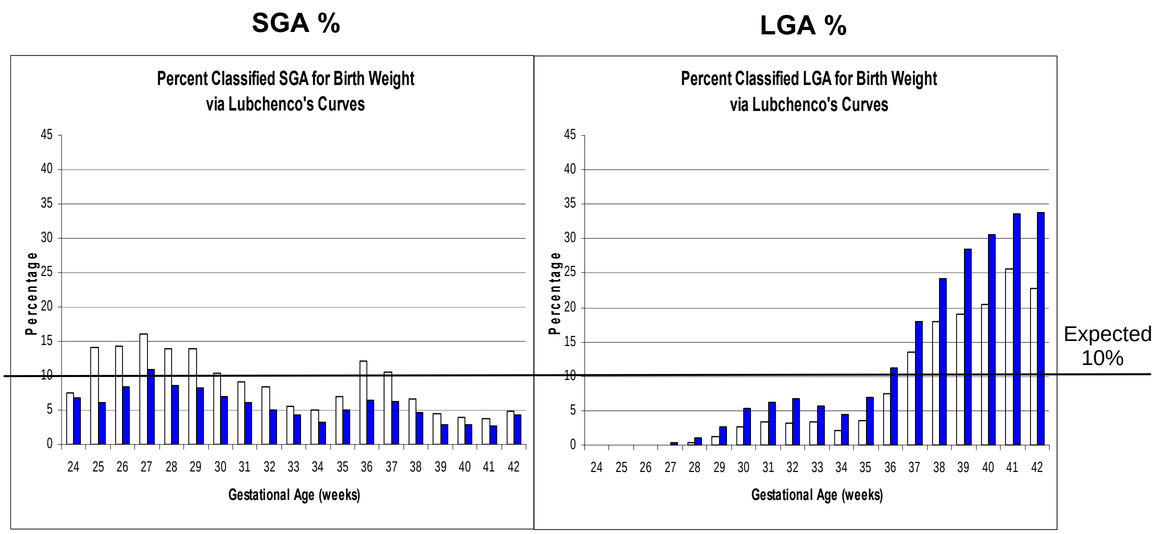
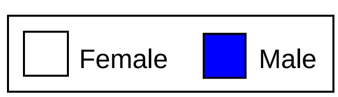


Figure 31: Birth weight: Percentage of Pediatrix females (white) and males (black) classified as SGA (left) and LGA (right)



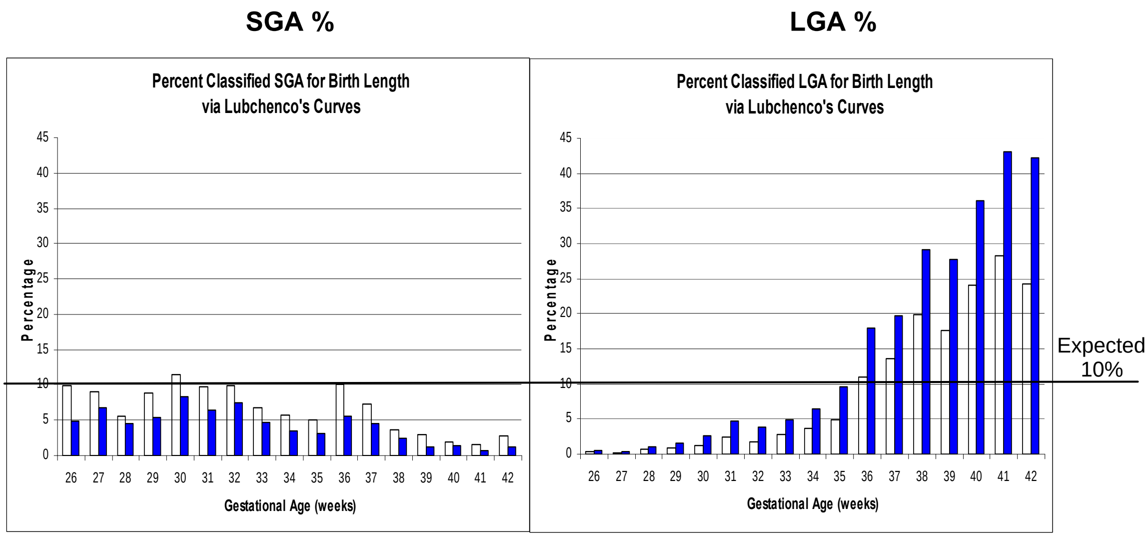


Figure 32: Birth length: Percentage of Pediatrix females (white) and males (black) classified as SGA (left) and LGA (right)

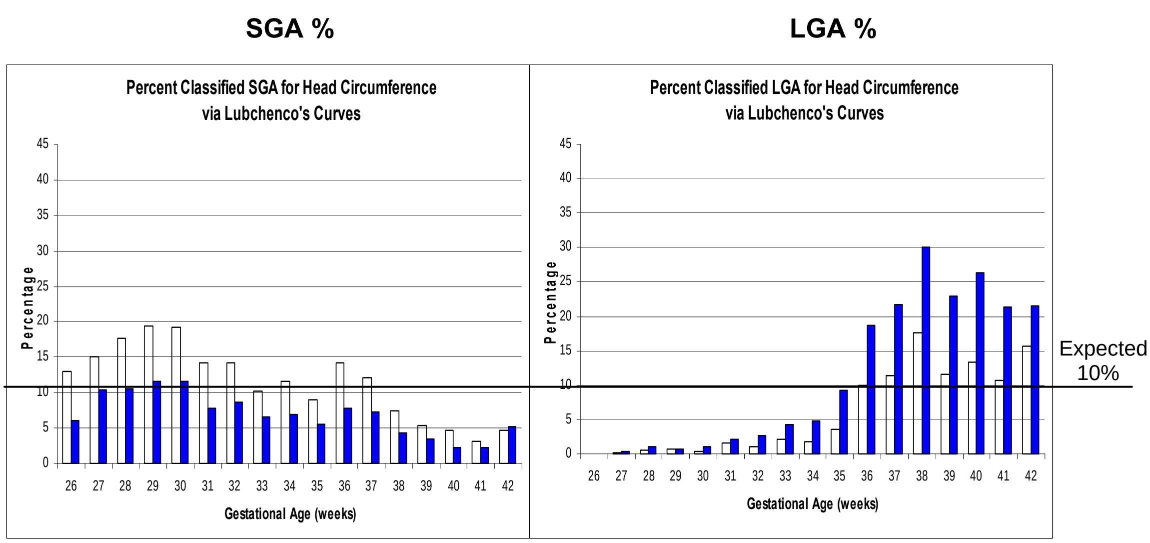
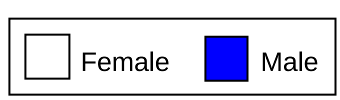


Figure 33: Head circumference: Percentage of Pediatrix females (white) and males (black) classified as SGA (left) and LGA (right)



When classifying the Pediatrix validation data set on Lubchenco's curves as SGA and LGA, as mentioned above, it was found that in many instances the Lubchenco curves underestimated or overestimated the number of SGA and LGA infants. In total (Table 26), 8 -11% of the female infants and 13-19% of the male infants in the validation sample were misclassified for either SGA, AGA or LGA for weight, length, or head circumference by the Lubchenco curves when compared to classification based on the new Pediatrix curves.

Table 26: Number of infants misclassified by the Lubchenco curves as SGA, AGA or LGA when compared to the new Pediatrix curves

	Females misclassified by Lubchenco's curves as:				**Males misclassified by Lubchenco's curves as:**			
	SGA	**AGA**	**LGA**	**Total (%)**	**SGA**	**AGA**	**LGA**	**Total (%)**
Birth weight[a] females(n)=54468 males(n)= 73056	227	3315	2014	5556 (10.2)	0	5294	5573	10867 (14.9)
Birth length[b] females(n)=53540 males(n)=71993	0	3740	2309	6049 (11.3)	0	5912	7579	13491 (18.7)
Head circumference[b] females(n)=53540 males(n)=71993	938	2530	847	4315 (8.1)	0	4604	5187	9791 (13.6)

[a]Weight classified starting at 24 weeks
[b]Length and head circumference classified starting at 26 weeks

Discussion of Specific Aim I

A recent large diverse data set of birth size measurements was analyzed to contribute to an understanding of whether the Lubchenco preterm infant growth curves should be replaced with newer curves. Preterm infant growth curves are significant to the health of infants as assessment of fetal and postnatal growth are key indicators of an infant's health and future adult health. Inadequate fetal growth can put an infant at higher risk for mortality, disease, complications and/or neurological delays (Kramer, Olivier, McLean, Willis et al., 1990). Slow postnatal growth can lead to neurological delays whereas fast postnatal growth has been correlated with metabolic syndrome later in life (Barker, Winter, Osmond et al., 1989; Lucas, 2005; Neu, Hauser, and Douglas-Escobar, 2007). Since growth curves are one of the primary tools used clinically and in epidemiological studies to assess fetal and postnatal growth, it is crucial that they be as accurate as possible. To this end, the Pediatrix data set (1998 – 2006) was used to create gender-specific smoothed birth weight, head circumference, and length growth curves for preterm infants. The new Pediatrix curves and the Pediatrix data were then compared to the Lubchenco curves in order to quantify the differences and to analyze the clinical significance of differences.

The new Pediatrix curves were found to be quite different from the Lubchenco curves. Through the graphical comparison (Figures 25-30), a pattern emerges which shows that the new Pediatrix curves usually cross the Lubchenco curves. All of the female and most of the male Pediatrix curves start out with smaller measurements, were similar between 30 and 36 weeks, and then were larger than the Lubchenco curves starting at approximately 36 weeks of age. This finding was in agreement with Thomas et al. (Thomas, Peabody, Turnier et al., 2000) who found similar differences. The actual differences of the 10th, 50th, and 90th percentiles had a large span ranging from the Pediatrix percentiles being 26.2% smaller to 11.8% larger.

The differences between the two set of curves are likely to have been affected by the 40 year gap between when the two data sets were collected. In the 1960's, early gestational ages infants who were smaller had lower survival rates than in recent years due to modern medical advances and therefore would not have been included in Lubchenco's data set (Lorenz, Wooliever, Jetton et al., 1998). A meta-analysis by Lorenz et al. (Lorenz, Wooliever, Jetton et al., 1998) found that survival of extremely small infants (800 g or less) rose 2.1% per year from 1976 to 1990. At ages closer to full-term, the Pediatrix curves may be larger because infants are generally being born with larger sizes since the 1960's (Oishi, Honda, Takamura et al., 2004). So, the new Pediatrix curves should be more representative of current birth sizes in the United States.

Another factor that could have contributed to differences in the curves was the difference in the demographics of the data sets. Lubchenco's data set was not a diverse data set – Caucasian infants from low-income families born in Denver while the Pediatrix data was a large heterogeneous (51% white, 24% Hispanic, 16% black, 10% other) data set from 33 states in the United States. So, the Pediatrix data was fairly representative of the United States population (Table 8) while the Lubchenco data was a smaller isolated sample.

It is also possible that statistical techniques could have created differences in the curves. Lubchenco smoothed the curves "arithmetically" which may have been a less rigorous method than the LMS method, used in this study, which combines the three cubic splines for median, variance, and skewness (Cole and Green, 1992; Lubchenco, Hansman, Dressler et al., 1963). The LMS method is a robust smoothing technique also used by the CDC to create their 2000 growth charts (Kuczmarski, Ogden, Grummer-Strawn et al., 2000). This method provided growth curves that were representative of the sample while at the same time reducing the variance appropriately. With the use of worm plots, z-scores, visual inspection and validation steps, the new Pediatrix growth curves have been proven to have an excellent goodness-of-fit to the large sample of 257,855 infants in the validation and curve samples combined.

To investigate the clinical significance of the differences in the curves, the classification of the Pediatrix validation data set on Lubchenco's curves for SGA, AGA and LGA was compared to the classification via the new Pediatrix curves. It was found that in many instances infants (8-11% of females and 13-19% of males) were misclassified as either SGA, AGA or LGA for weight, length, or head circumference based on the older curves. At most gestational ages and for all three measurements, the number of SGA males was underestimated by 40-50% or more. For female infants, the SGA classification was also often underestimated but at the younger ages was overestimated by as much as 80%. In the LGA category, males and females greater than 36 weeks of age were extremely overestimated by up to 400%. At earlier than 36 weeks of age, LGA was underestimated by up to 90% with females being especially underestimated. Use of the Lubchenco curves would result in a large number of size-for-age misclassifications, primarily underestimating SGA and overestimating LGA, and thus infants who should be categorized as high-risk may not receive the extra attention reserved for high-risk patients in the NICU.

These findings support two conclusions: the need to replace Lubchenco's curves due to misclassification and the need for gender-specific curves. At each gestational age, the Lubchenco curves either classify too many infants or too few infants as SGA or LGA. The number of children that were misclassified as SGA, LGA, or AGA was large and indicates the need for new curves. The difference in the percentage of females (8-11%) and males (13-19%) who were misclassified demonstrates the need for gender-specific curves. Many more male infants than female infants were misclassified with Lubchenco's gender-specific curves. In the next section, the differences in birth size between male and females is investigated further.

The new curves were created from a large heterogeneous data set (391,681 originally) collected recently (1998-2006) at 248 NICUs in 33 states across the United States. The racial demographics (Table 8) were similar to births in the United States in 2005. Thus, these growth curves can be generalized to use in NICUs across the country. Inclusions and exclusions were constructed with the goal of limiting the data set to the healthiest infants with accurate measurements. A rigorous smoothing technique, previously used by the CDC, was utilized (Kuczmarski, 2000). These gender-specific curves could provide a much needed update to the Lubchenco curves with recent data representative of the country.

One possible limitation of this study is the accuracy of gestational age as gestational age can be incorrect due to misstatement of the date of the last menstrual period by the mother or bleeding after conception. The gestational age in this study was determined by the neonatologist who used his or her best judgment to determine the gestational age based on the information known to him or her including ultrasounds, physical exams, and information reported by the mother. Extreme outliers were removed from the data set which should have eliminated infants with improbable sizes for their ages.

Other possible limitations were possible measurement errors and limitations in the measurement accuracy. The growth measurements were not collected in a controlled research setting. However, the measurements were made in the clinical setting where the nurses generally receive training on measurement procedures. Weight is likely to be the most accurate measurement as digital scales are used, followed by head circumference and then length (Kramer, McLean, Olivier et al., 1989). In order to reduce errors in the data set, extreme outliers were removed from the data set. Therefore, the data set used to create the new Pediatrix growth curves should be free of the majority of inaccurate measurements.

Future work could include the development of a body proportionality index, investigation of the SGA, AGA, and LGA cutoffs, as well as an analysis of secular trends in birth size. Body proportionality is important to consider along with the birth size measurements (weight, length, and head circumference) in the evaluation of growth status of infants. The method used by Lubchenco et al. (Lubchenco, Hansman, and Boyd, 1966) as a weight for length ratio is Rohrer's ponderal index (weight (g) multiplied by 100 divided by length^3 (cm^3)). Through the use of the ponderal index, Lubchenco found some aspects of growth that were not revealed in the individual weight and length curves (Lubchenco, Hansman, and Boyd, 1966). Lubchenco observed that infants increased in weight more than length in the latter part of the pregnancy (Lubchenco, Hansman, and Boyd, 1966). However, it has been shown that during fetal growth body proportions change; so one measurement, such as the ponderal index, may not be appropriate for all gestational ages and multiple indexes may be needed (Cole, Henson, Tremble et al., 1997).

Another area of future work is the analysis of the percentile-for-age classifications of SGA, AGA, and LGA. The percentile-for-age classifications were determined based on infant mortality rates from 1958 to 1961 in New York City (Battaglia and Lubchenco, 1967; Erhardt, Joshi, Nelson et al., 1964). Neonatal mortality rates were used to determine the cutoffs in order to ensure that the high risk groups be closely monitored. Since neonatal mortality rates have changed in the last 50 years, it seems logical to reassess these percentile cutoffs (Lorenz, Wooliever, Jetton et al., 1998). If the cutoffs were not appropriately categorizing infants who were at high-risk, some high-risk infants may be missed while others may be put through unnecessary tests and their parents may be unnecessarily alarmed.

Another possible area to analyze is the change in the birth sizes over time. As Pediatrix adds more data to this data set, there may be enough years of data in order to look for secular trends in birth sizes. This would be an interesting study given the current obesity epidemic. If changes in birth sizes were found, this analysis could assess a number of contributing factors, such as maternal size, socioeconomic status, racial group, and infant gender. There could be a number of determinants leading to the change in size and may signify a need for more prenatal care or changes in diet during pregnancy.

In the analysis for this specific aim, it was found that Lubchenco’s curves need to be updated due to the large numbers of infants misclassified as SGA, AGA, and LGA and that gender-specific curves were warranted. The new Pediatrix curves were created from a large heterogeneous recent data set collected in 33 states across the United States. The data set has racial demographics similar to the United States and therefore the new Pediatrix curves were representative of the country. These new contemporary rigorously created curves may provide an update to the Lubchenco curves.

<u>Results for Specific Aim II:</u>

A. Gender analysis

As found in the preliminary studies, significant differences were found in birth size measurements between males and females in this data set. Tables 27-29 below display a comparison of mean weight, lengths and head circumferences by age group between the genders. Analysis of variance was used to test for the statistical significance in the difference and all measurements for each age group were found to be statistically significant with p-values less than 0.0001. The majority of these differences would also be considered clinically significant.

Table 27: Birth weight: Means compared by gender and age group

Age group (weeks)	Female mean birth weight (kg) ± SD	Male mean birth weight (kg) ± SD	# g male larger (%)	P-value*
23-26	0.747 ±0.148	0.795 ±0.152	48 (6.39)	<.0001
27-31	1.283 ±0.316	1.363 ±0.323	80 (6.21)	<.0001
32-36	2.230 ±0.496	2.373 ±0.524	143 (6.41)	<.0001
37-41	3.275 ±0.537	3.387 ±0.547	112 (3.43)	<.0001

*Statistical significance of ANOVA (alpha=0.0125)

Table 28: Birth length: Means compared by gender and age group

Age group (weeks)	Female mean birth length (cm) ± SD	Male mean birth length (cm) ± SD	# cm male larger (%)	P-value*
23-26	32.5 ±2.4	33.0 ±2.4	0.5 (1.75)	<.0001
27-31	38.5 ±3.2	39.2 ±3.2	0.7 (1.81)	<.0001
32-36	45.0 ±3.1	45.9 ±3.2	0.9 (1.93)	<.0001
37-41	49.9 ±2.7	50.6 ±2.7	0.7 (1.41)	<.0001

*Statistical significance of ANOVA (alpha=0.0125)

Table 29: Head circumference: Means compared by gender and age group

Age group (weeks)	Female mean head circumference (cm) ± SD	Male mean head circumference (cm) ± SD	# cm male larger (%)	P-value*
23-26	22.7 ±1.5	23.3 ±1.5	0.6 (2.22)	<.0001
27-31	26.9 ±2.0	27.4 ±2.0	0.5 (2.07)	<.0001
32-36	31.2 ±1.9	31.8 ±2.0	0.6 (2.11)	<.0001
37-41	33.9 ±1.6	34.5 ±1.7	0.6 (1.71)	<.0001

*Statistical significance of ANOVA (alpha=0.0125)

The male and female new Pediatrix growth curves are overlaid in the Figures 34-36 below to show the differences in the male and female curves. Along with each set of curves, a table (31-33) is displayed which compares points on the percentile curves at gestational ages 26, 30, 36, and 40 weeks. Based on the smoothed curves, male infants were found to have larger birth weights by 2.8-9.1% depending on gestational age. Males were also found to have longer lengths by 1.2-2.0% depending on gestational age. Male head circumferences were larger than females by 1.3-2.7% depending on gestational age.

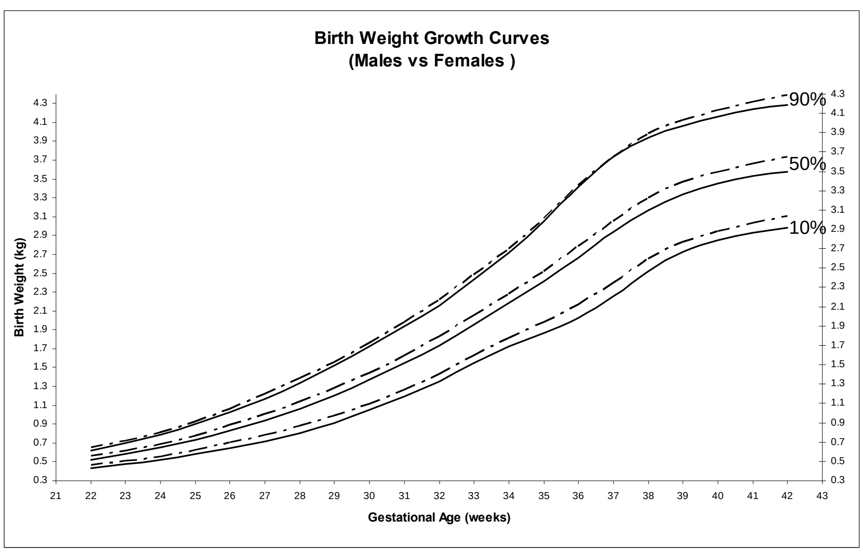


Figure 34: Birth weight: New Pediatrix curves for males (dashed) and females (solid)

Table 30: Birth weight: Comparison between male and female percentiles at selected gestational ages

Gestational age (weeks)	10th percentile			50th percentile			90th percentile		
	Female (kg)	Male (kg)	% male larger	Female (kg)	Male (kg)	% male larger	Female (kg)	Male (kg)	% male larger
26	0.645	0.704	9.13	0.827	0.890	7.53	1.004	1.065	6.11
30	1.052	1.114	5.91	1.373	1.443	5.11	1.693	1.761	3.99
36	2.028	2.170	7.00	2.664	2.792	4.78	3.339	3.432	2.77
40	2.855	2.950	3.35	3.454	3.579	3.62	4.070	4.232	3.98

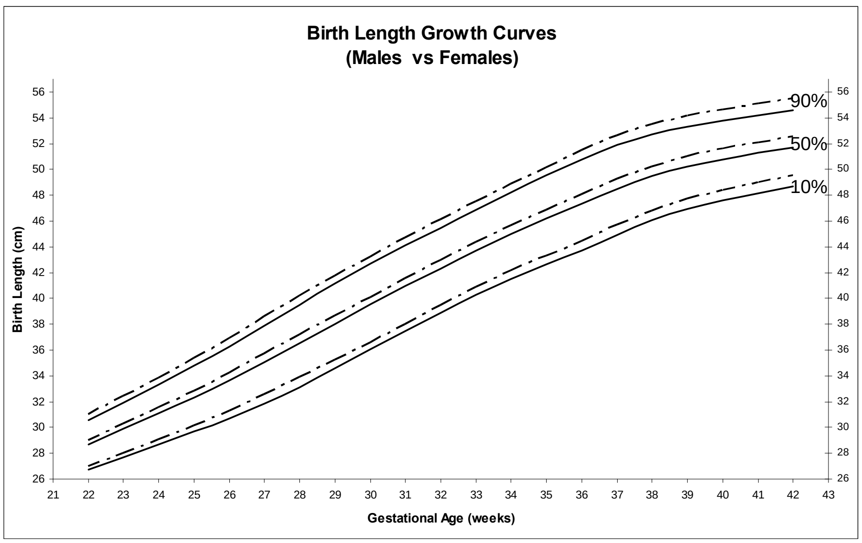


Figure 35: Birth length: New Pediatrix curves for males (dashed) and females (solid)

Table 31: Birth length: Comparison between male and female percentiles at selected gestational ages

Gestational age (weeks)	10th percentile			50th percentile			90th percentile		
	Female (cm)	Male (cm)	% male larger	Female (cm)	Male (cm)	% male larger	Female (cm)	Male (cm)	% male larger
26	30.7	31.3	2.04	33.6	34.3	1.87	36.3	37.0	1.86
30	36.0	36.6	1.55	39.5	40.1	1.43	42.7	43.2	1.20
36	43.7	44.5	1.70	47.4	48.1	1.58	50.8	51.5	1.37
40	47.6	48.4	1.70	50.8	51.6	1.62	53.8	54.7	1.66

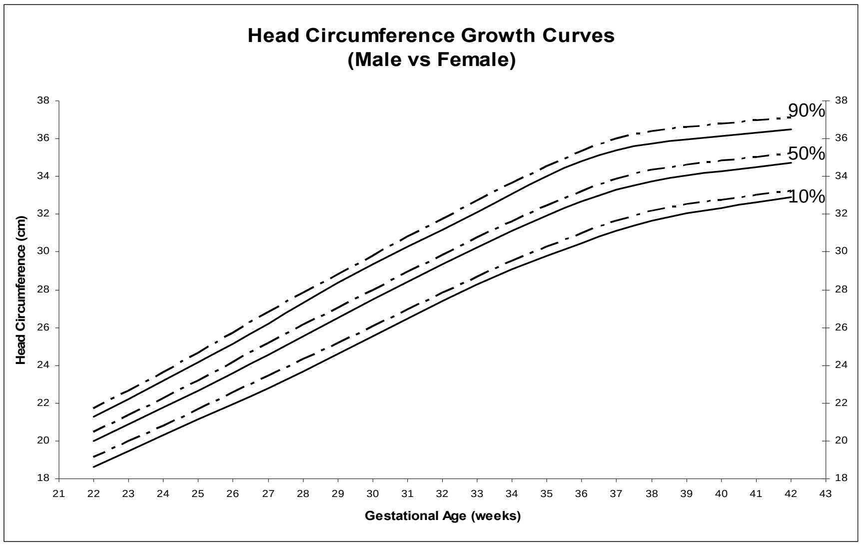


Figure 36: Head circumference: New Pediatrix curves for males (dashed) and females (solid)

Table 32: Head circumference: Comparison between male and female percentiles at selected gestational ages

Gestational age (weeks)	10th percentile			50th percentile			90th percentile		
	Female (cm)	Male (cm)	% male larger	Female (cm)	Male (cm)	% male larger	Female (cm)	Male (cm)	% male larger
26	22.0	22.5	2.69	23.6	24.2	2.46	25.1	25.7	2.26
30	25.6	26.1	2.03	27.5	28.0	1.75	29.4	29.8	1.54
36	30.5	31.0	1.68	32.7	33.2	1.64	34.8	35.3	1.57
40	32.3	32.8	1.35	34.3	34.8	1.63	36.1	36.8	1.81

B. Racial group analysis

The racial groups were analyzed for differences in birth sizes and for possible contributing factors to differences. In this analysis, the gestational ages 22 weeks and 42 weeks were omitted due to the small sample size at 22 weeks and the drop-off in size at 42 weeks. The analysis was performed on the black, Hispanic and white infants. Those infants with a race of “missing”, ”unknown data”, “other”, “Pacific Islander”, “Asian” or “American/Alaska Native” were not included in this analysis due to their small sample size. It would also not be logical to combine them as an "other" group and include them in the analysis as it would difficult to interpret results for this disparate group.

In Table 33 below, it is shown that a greater percentage of black infants were born at earlier gestational ages. Approximately 5.5% of black infants of both genders were born at 23-26 weeks compared to approximately 2.5% of the white and Hispanic infants. This finding was also true at 27-31 weeks where approximately 16% of black infants of both genders were born in this age range compared to approximately 10% of Hispanic and whites. The trend reverses from 32 to 41 weeks where it was found that a smaller percentage of the black infants were born in the older age groups.

Table 33: Percent of infants by race born in each age group

Age group (weeks)	**Female**			**Male**		
	Black (%)	**Hispanic (%)**	**White (%)**	**Black (%)**	**Hispanic (%)**	**White (%)**
23-26	6.08	2.60	2.68	5.37	2.44	2.24
27-31	17.93	10.43	11.9	15.66	9.84	10.39
32-36	41.33	38.10	43.84	38.63	38.74	42.46
37-41	34.66	48.87	41.58	40.34	48.97	44.92

Tables 34-39 below show the mean growth measurements of the racial groups by age group. Black and Hispanic infants were compared to white infants. These tables show that black and Hispanic infants were generally smaller than white infants with black infants being the smallest of the three groups. The differences in means were most notable in the birth weight measurement in the age groups of 27-31 weeks and 32-36 weeks; 32-36 week black female infants were smaller than white females by 9.0% and 32-36 week black males were smaller by 10.4%. Hispanic female infants were also found to be smaller (2.4% at 32-36 weeks) than white females but by a much smaller amount than black infants.

Table 34: Female birth weight: Means by age and racial group

Age group (weeks)	Black		Hispanic		White	P-value*
	Mean weight (kg) ± SD	% < white	Mean weight (kg) ± SD	% < white	Mean weight (kg) ± SD	
23-26	0.729 ±0.141[a]	-3.95	0.749 ±0.147	-1.25	0.759 ±0.153[c]	0.0014
27-31	1.225 ±0.301[ab]	-6.12	1.301 ±0.307[c]	-0.31	1.305 ±0.323[c]	<.0001
32-36	2.077 ±0.456[ab]	-8.99	2.226 ±0.493[ac]	-2.44	2.282 ±0.503[bc]	<.0001
37-41	3.188 ±0.563[ab]	-4.14	3.244 ±0.524[ac]	-2.47	3.326 ±0.533[bc]	<.0001

Bonferroni (Dunn) *t*-tests used for comparison among racial groups; statistical significance defined as alpha=0.0125 where [a]different from white; [b]different from Hispanic; [c]different from black.
*Overall comparison, statistical significance of ANOVA (alpha=0.0125)

Table 35: Female birth length: Means by age and racial group

Age group (weeks)	Black		Hispanic		White	P-value*
	Mean length (cm) ± SD	% < white	Mean length (cm) ± SD	% < white	Mean length (cm) ± SD	
23-26	32.3 ±2.3	-0.91	32.4 ±2.4	-0.71	32.6 ±2.4	0.0672
27-31	38.0 ±3.2[ab]	-1.60	38.7 ±3.2[c]	0.07	38.6 ±3.3[c]	<.0001
32-36	44.3 ±3.0[ab]	-2.14	45.1 ±3.1[ac]	-0.34	45.3 ±3.1[bc]	<.0001
37-41	49.5 ±2.8[ab]	-1.13	49.8 ±2.6[ac]	-0.55	50.1 ±2.7[bc]	<.0001

Bonferroni (Dunn) *t*-tests used for comparison among racial groups; statistical significance defined as alpha=0.0125 where [a]different from white; [b]different from Hispanic; [c]different from black.
*Overall comparison, statistical significance of ANOVA (alpha=0.0125)

Table 36: Female head circumference: Means by age and racial group

Age group (weeks)	Black		Hispanic		White	P-value*
	Mean head circumference (cm) ± SD	% < white	Mean head circumference (cm) ± SD	% < white	Mean head circumference (cm) ± SD	
23-26	22.5 ±1.5[a]	-1.65	22.7 ±1.4	-0.94	22.9 ±1.4[c]	<.0001
27-31	26.5 ±1.9[ab]	-2.14	26.9 ±2.0[c]	-0.37	27.0 ±2.0[c]	<.0001
32-36	30.5 ±1.8[ab]	-3.02	31.1 ±1.9[ac]	-0.90	31.4 ±1.9[bc]	<.0001
37-41	33.5 ±1.7[ab]	-1.80	33.7 ±1.5[ac]	-1.21	34.1 ±1.6[bc]	<.0001

Bonferroni (Dunn) *t*-tests used for comparison among racial groups; statistical significance defined as alpha=0.0125 where [a]different from white; [b]different from Hispanic; [c]different from black.
*Overall comparison, statistical significance of ANOVA (alpha=0.0125)

Table 37: Male birth weight: Means by age and racial group

Age group (weeks)	Black		Hispanic		White	P-value*
	Mean weight (kg) ± SD	% < white	Mean weight (kg) ± SD	% < white	Mean weight (kg) ± SD	
23-26	0.775 ±0.145[a]	-3.77	0.802 ±0.155	-0.42	0.805 ±0.153[c]	0.0006
27-31	1.300 ±0.300[ab]	-6.26	1.359 ±0.313[ac]	-1.99	1.387 ±0.335[bc]	<.0001
32-36	2.187 ±0.489[ab]	-10.40	2.340 ±0.502[ac]	-4.12	2.441 ±0.530[bc]	<.0001
37-41	3.299 ±0.576[ab]	-4.23	3.331 ±0.539[ac]	-3.28	3.444 ±0.537[bc]	<.0001

Bonferroni (Dunn) *t*-tests used for comparison among racial groups; statistical significance defined as alpha=0.0125 where [a]different from white; [b]different from Hispanic; [c]different from black.
*Overall comparison, statistical significance of ANOVA (alpha=0.0125)

Table 38: Male birth length: Means by age and racial group

Age group (weeks)	Black		Hispanic		White	P-value*
	Mean length (cm) ± SD	% < white	Mean length (cm) ± SD	% < white	Mean length (cm) ± SD	
23-26	32.8 ±2.4	-0.86	33.0 ±2.5	-0.26	33.1 ±2.4	0.0961
27-31	38.7 ±3.1[ab]	-1.59	39.2 ±3.3[c]	-0.33	39.3 ±3.3[c]	<.0001
32-36	45.0 ±3.1[ab]	-2.60	45.9 ±3.1[ac]	-0.71	46.2 ±3.2[bc]	<.0001
37-41	50.2 ±2.8[ab]	-1.30	50.5 ±2.7[ac]	-0.75	50.8 ±2.7[bc]	<.0001

Bonferroni (Dunn) *t*-tests used for comparison among racial groups; statistical significance defined as alpha=0.0125 where [a]different from white; [b]different from Hispanic; [c]different from black.
*Overall comparison, statistical significance of ANOVA (alpha=0.0125)

Table 39: Male head circumference: Means by age and racial group

Age group (weeks)	Black		Hispanic		White	P-value*
	Mean head circumference (cm) ± SD	% < white	Mean head circumference (cm) ± SD	% < white	Mean head circumference (cm) ± SD	
23-26	23.0 ±1.4[a]	-1.81	23.3 ±1.6	-0.75	23.4 ±1.5[c]	<.0001
27-31	27.0 ±1.9[ab]	-2.41	27.4 ±1.9[ac]	-0.92	27.6 ±1.9[bc]	<.0001
32-36	31.0 ±1.9[ab]	-3.54	31.7 ±1.9[ac]	-1.36	32.1 ±2.0[bc]	<.0001
37-41	34.0 ±1.7[ab]	-2.13	34.2 ±1.6[ac]	-1.42	34.7 ±1.6[bc]	<.0001

Bonferroni (Dunn) *t*-tests used for comparison among racial groups; statistical significance defined as alpha=0.0125 where [a]different from white; [b]different from Hispanic; [c]different from black.
*Overall comparison, statistical significance of ANOVA (alpha=0.0125)

Z-scores were calculated for the infants in the validation sample using the new Pediatrix curves. Bar charts (Figures 37-42) below show the differences in mean z-scores by age group of black infants (diagonal lines), Hispanic infants (dots), and white infants (bricks). The z-scores were converted to the percentile equivalent (relative to the median) and are displayed by each bar in Figure 37-42. As discussed in the methods, a z-score is the number of standard deviations from the mean and is indicative of the infant's measurements relative to the same age peers. Therefore, the z-scores provide more information than the means alone.

The largest racial group differences in z-scores were found in the 32-36 and 37-41 week age groups indicating that fetal growth in the latter part of pregnancy was where the largest differences were found. For females in the 32-36 week age group, the mean percentile differences relative to the median between black and white infants were 14 percentiles for weight, 10 percentiles for length, and 18 percentiles for head circumference. For males in the 32-36 week age group, the mean percentile differences relative to the median between black and white infants were 19 percentiles for weight, 12 percentiles for length, and 22 percentiles for head circumference. These percentile differences illustrate a large difference in the growth status of black infants compared to white infants.

Females:

Weight

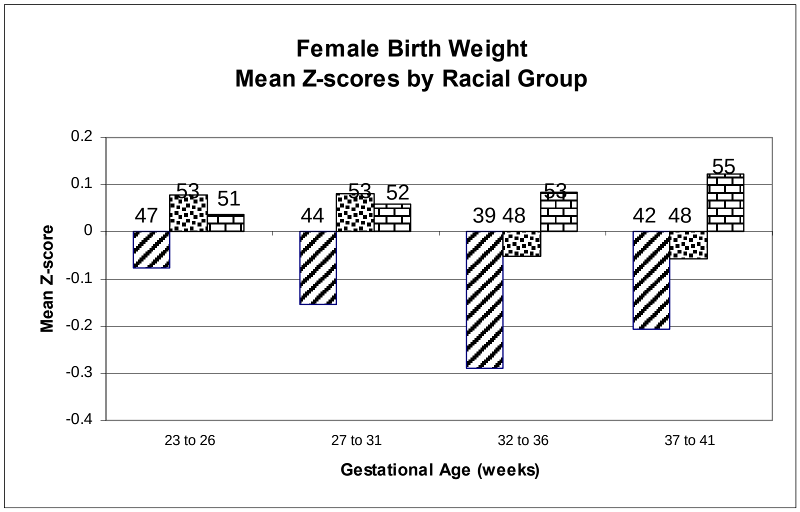


Figure 37: Female birth weight: Mean z-scores by age and racial groups
Percentile value (converted from the z-score) displayed above each bar.

Females:

Length

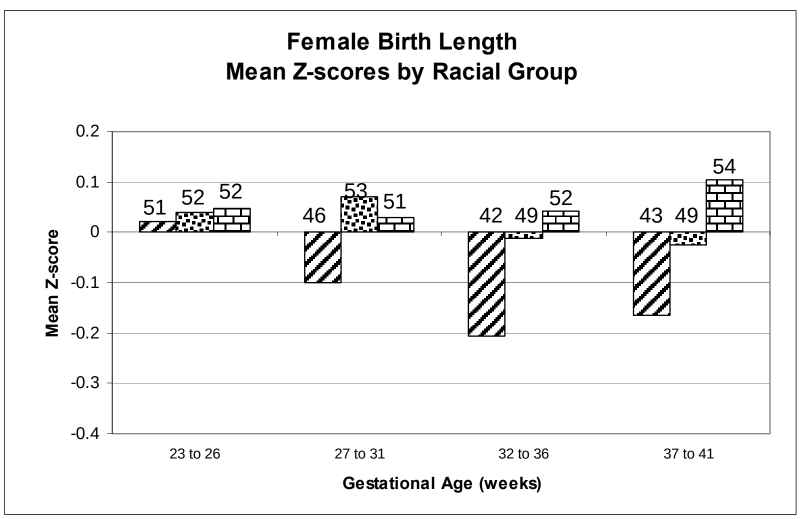


Figure 38: Female birth length: Mean z-scores by age and racial groups
Percentile value (converted from the z-score) displayed above each bar.

Females:

Head Circum.

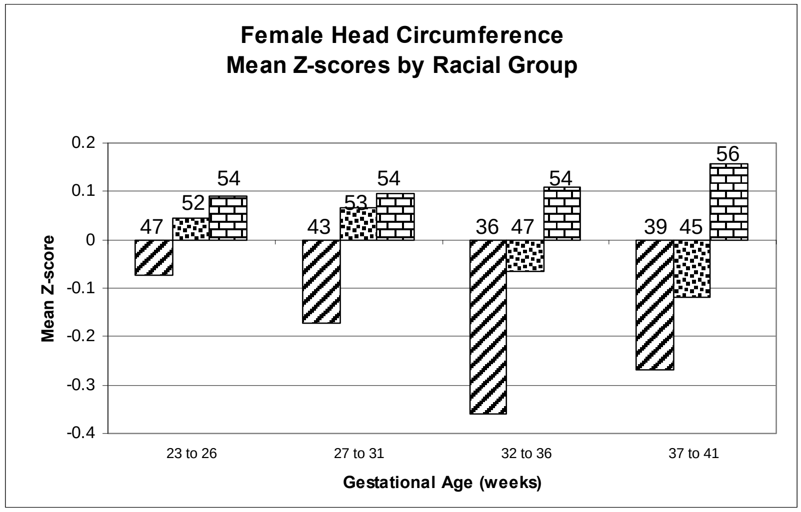


Figure 39: Female head circumference: Mean z-scores by age and racial groups Percentile value (converted from the z-score) displayed above each bar.

Males:

Weight

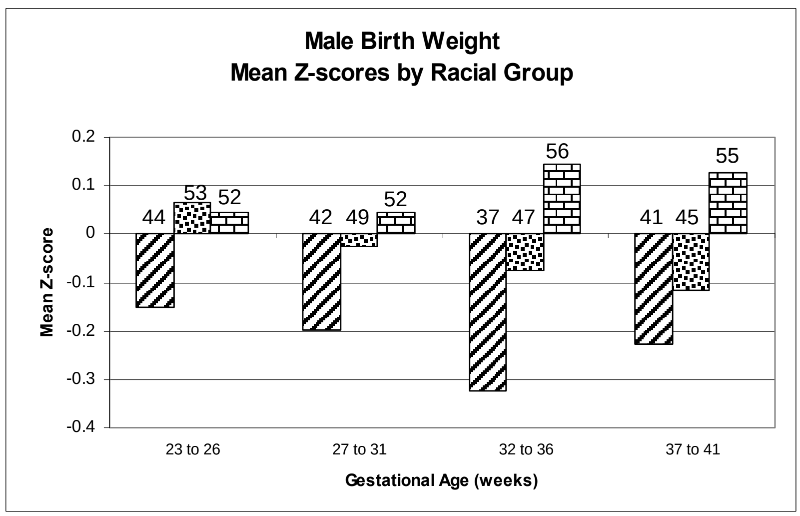


Figure 40: Male birth weight: Mean z-scores by age and racial groups Percentile value (converted from the z-score) displayed above each bar.

Males:

Length

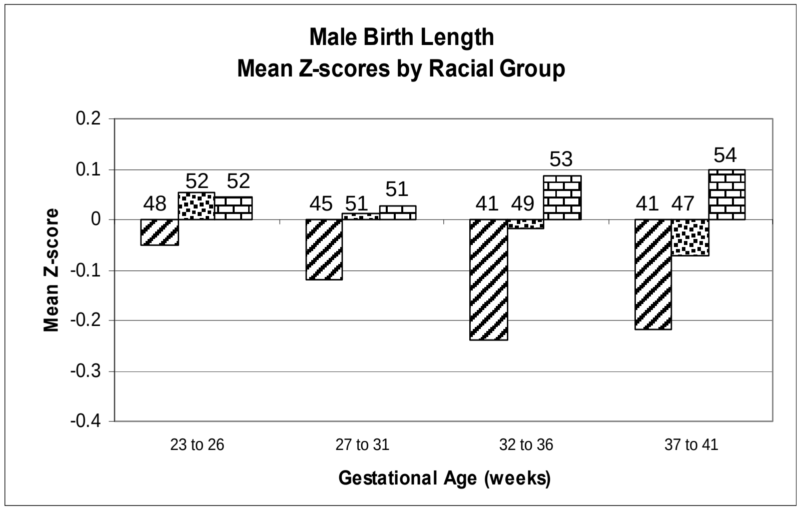


Figure 41: Male birth length: Mean z-scores by age and racial groups
Percentile value (converted from the z-score) displayed above each bar.

Males:

Head
Circum.

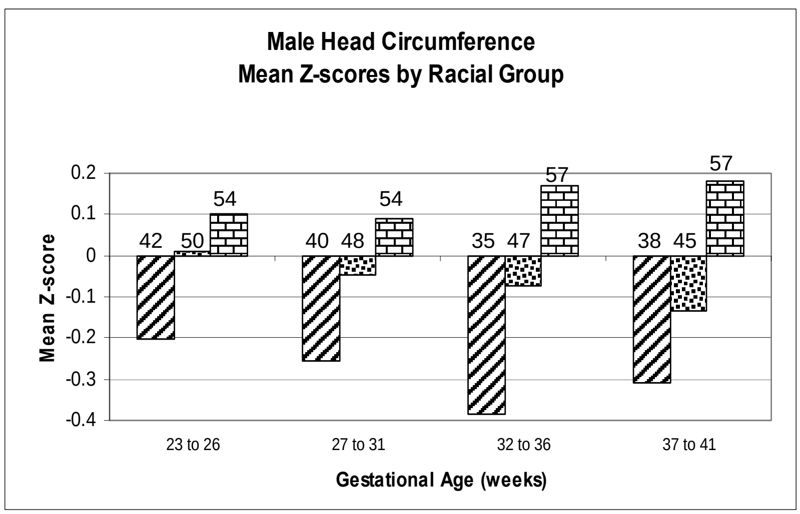


Figure 42: Male head circumference: Mean z-scores by age and racial groups
Percentile value (converted from the z-score) displayed above each bar.

Figures 43-54 show the percentage of infants by racial group that were classified as SGA (<10th percentile) on the left or LGA (>90th percentile) on the right based on the new Pediatrix curves. With a few exceptions, black infants were found to have higher rates of SGA for all birth size measurements. Similar to the z-scores, the largest differences were found in the 32-26 and 37-41 week age groups. Black females had higher rates of SGA with ranges depending on age group (weight (8.8-14.5%), length (8.1-13.9%) or head circumference (10.1-17.2%)) compared to white females (weight (7.5-11.1%), length (8.1-11.3%) and head circumference (7.5-8.3%)) and Hispanic females (weight (7.9-10.5%), length (8.2-10.1%), and head circumference (8.3-11.7%)). For male infants, blacks had higher rates of SGA with ranges depending on age group (weight (10.8-16.6%), length (10.9-14.6%) or head circumference (13.3-18.2%)) compared to whites (weight (7.4-10.7%), length (8.0-10.6%) and head circumference (7.2-9.4%)) and Hispanics (weight (6.9-10.6%), length (8.9-11.6%), and head circumference (8.7-11.8%)). In the right column, the percentage of infants that were classified as LGA are shown. In most cases, white infants were found to have higher rates of LGA birth sizes.

Females SGA%

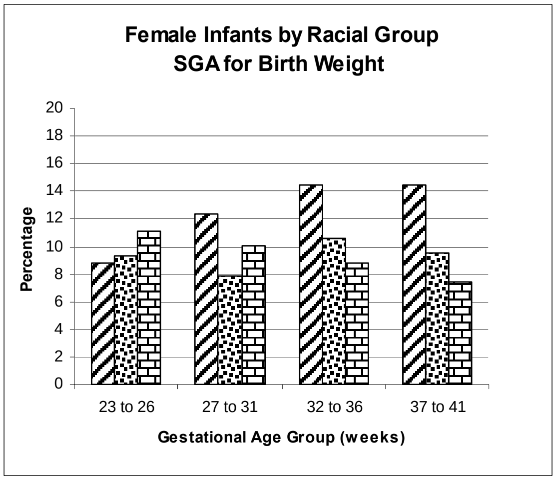


Figure 43: Female birth weight: Percentage born SGA by racial group

Females LGA %

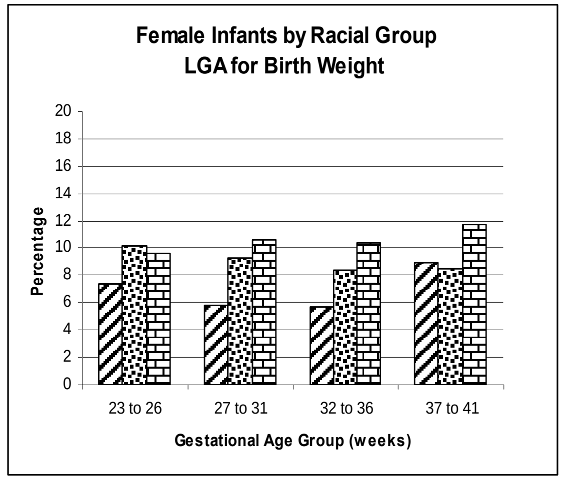


Figure 44: Female birth weight: Percentage born LGA by racial group

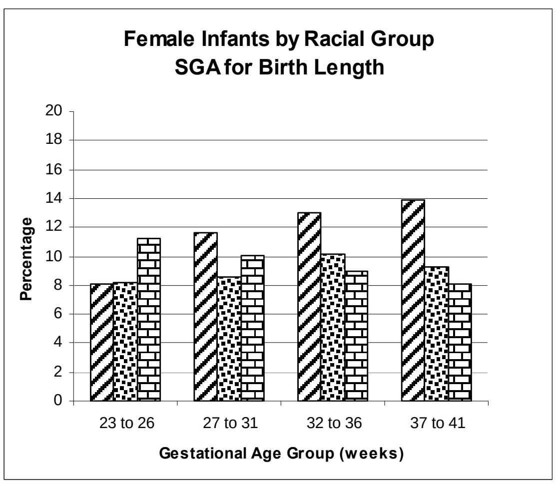


Figure 45: Female birth length: Percentage born SGA by racial group

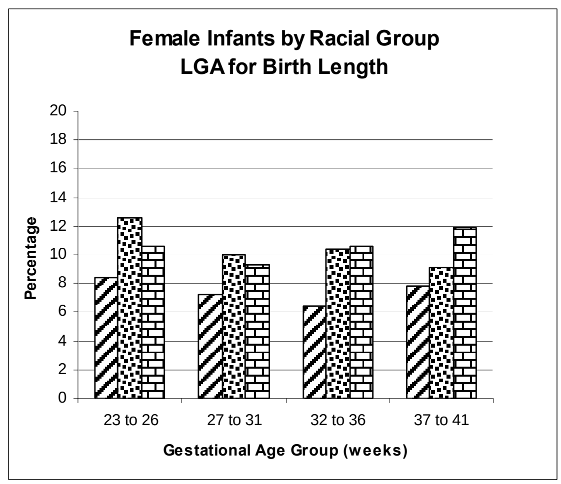


Figure 46: Female birth length: Percentage born LGA by racial group

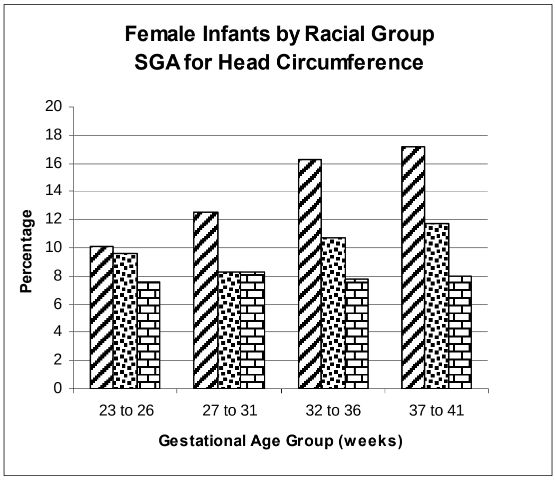


Figure 47: Female head circumference: Percentage born SGA by racial group

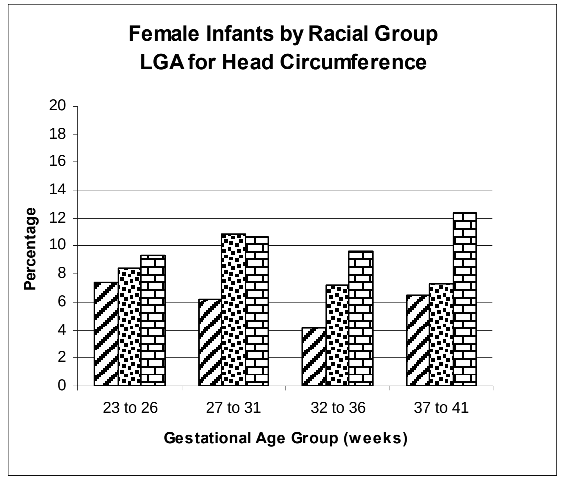


Figure 48: Female head circumference: Percentage born LGA by racial group

Males SGA%

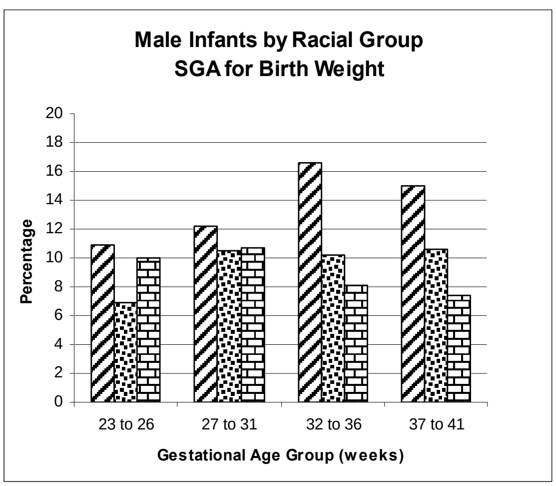


Figure 49: Male birth weight: Percentage born SGA by racial group

Males LGA %

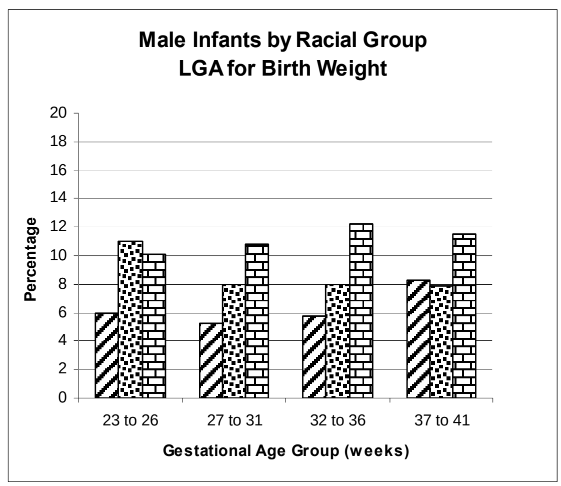


Figure 50: Male birth weight: Percentage born LGA by racial group

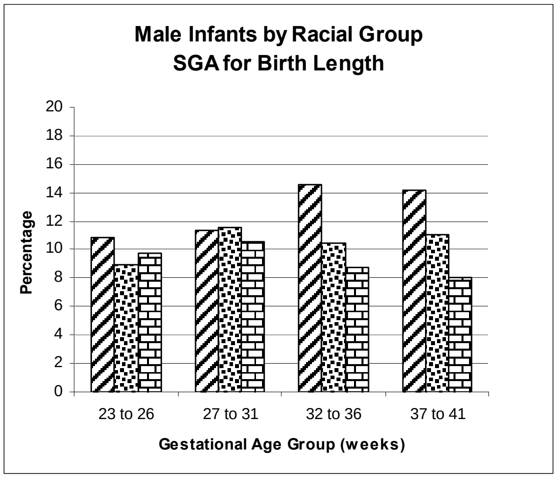


Figure 51: Male birth length: Percentage born SGA by racial group

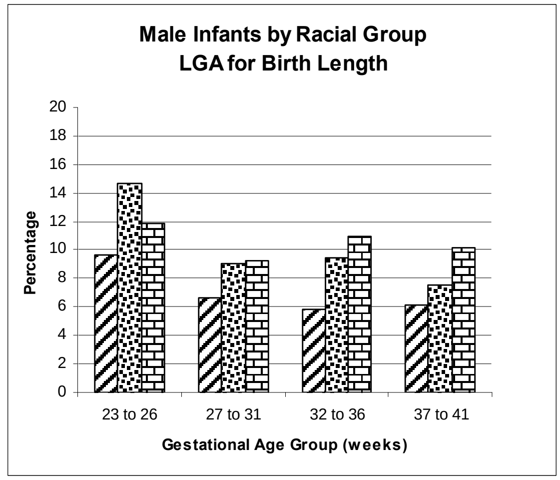


Figure 52: Male birth length: Percentage born LGA by racial group

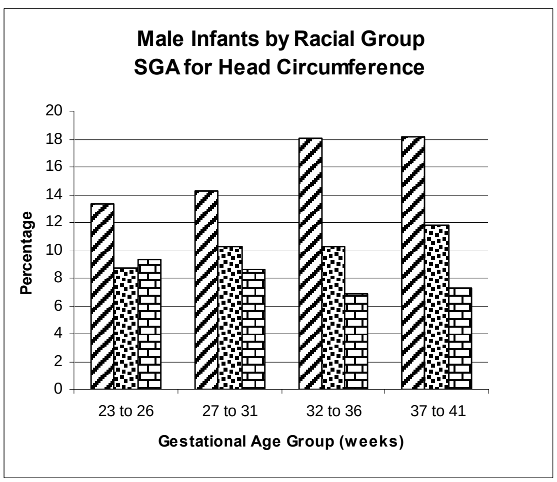


Figure 53: Male head circumference: Percentage born SGA by racial group

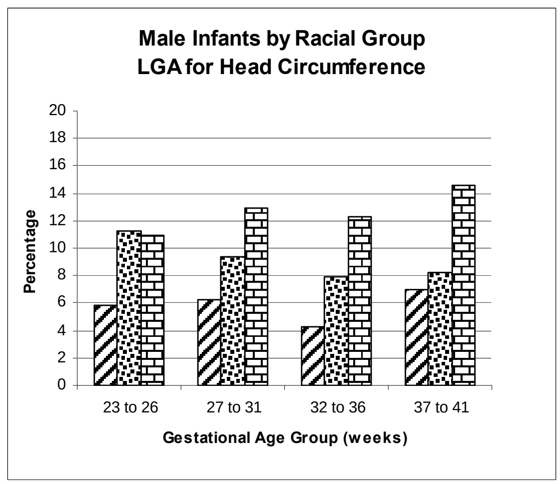


Figure 54: Male head circumference: Percentage born LGA by racial group

Logistic regression analysis with white infants as the reference group was used to estimate the odds ratio of being born SGA for all three birth size measurements (weight, length, and head circumference) by gestational age group. Tables 40-45 show the odds ratios with the confidence limits that were computed. The cells that are shaded signify a significant odds ratio indicated by confidence limits above one. Black infants had higher odds of being born SGA – especially for birth weight and head circumference. These differences were the most striking for males in the latter two age groups (32-36, 37-41 weeks) where it was found that black infants have a two times or higher risk of being born SGA than white infants.

Table 40: Female Birth Weight: Odds Ratio for being born SGA by racial group and age category

Racial group	Odds ratio*	95% Wald confidence limits		Odds ratio*	95% Wald confidence limits	
	23 to 26 weeks			27 to 31 weeks		
Black	0.772	0.531	1.121	1.259	1.045	1.516
Hispanic	0.822	0.534	1.265	0.760	0.606	0.954
	32 to 36 weeks			37 to 41 weeks		
Black	1.259	1.045	1.516	2.089	1.851	2.359
Hispanic	0.760	0.606	0.954	1.299	1.165	1.449

*White infants are reference group
Shaded cells indicate statistical significance

Table 41: Female Birth Length: Odds Ratio for being born SGA by racial group and age category

Racial group	Odds ratio*	95% Wald confidence limits		Odds ratio*	95% Wald confidence limits	
	23 to 26 weeks			27 to 31 weeks		
Black	0.694	0.474	1.016	1.175	0.972	1.421
Hispanic	0.700	0.447	1.098	0.843	0.676	1.051
	32 to 36 weeks			37 to 41 weeks		
Black	1.511	1.349	1.693	1.841	1.631	2.078
Hispanic	1.139	1.019	1.273	1.165	1.046	1.298

*White infants are reference group
Shaded cells indicate statistical significance

Table 42: Female Head Circumference: Odds Ratio for being born SGA by racial group and age category

Racial group	Odds ratio*	95% Wald confidence limits		Odds ratio*	95% Wald confidence limits	
	23 to 26 weeks			27 to 31 weeks		
Black	1.370	0.928	2.023	1.572	1.297	1.906
Hispanic	1.302	0.829	2.047	0.996	0.792	1.252
	32 to 36 weeks			37 to 41 weeks		
Black	2.292	2.054	2.556	2.376	2.119	2.665
Hispanic	1.421	1.271	1.589	1.526	1.378	1.689

*White infants are reference group
Shaded cells indicate statistical significance

Table 43: Male Birth Weight: Odds Ratio for being born SGA by racial group and age category

Racial group	Odds ratio*	95% Wald confidence limits		Odds ratio*	95% Wald confidence limits	
	23 to 26 weeks			27 to 31 weeks		
Black	1.103	0.780	1.560	1.160	0.970	1.386
Hispanic	0.668	0.433	1.030	0.974	0.811	1.170
	32 to 36 weeks			37 to 41 weeks		
Black	2.257	2.043	2.494	2.223	2.008	2.461
Hispanic	1.282	1.164	1.412	1.499	1.371	1.638

*White infants are reference group
Shaded cells indicate statistical significance

Table 44: Male Birth Length: Odds Ratio for being born SGA by racial group and age category

Racial group	Odds ratio*	95% Wald confidence limits		Odds ratio*	95% Wald confidence limits	
	23 to 26 weeks			27 to 31 weeks		
Black	1.133	0.799	1.605	1.077	0.897	1.293
Hispanic	0.912	0.611	1.360	1.103	0.923	1.318
	32 to 36 weeks			37 to 41 weeks		
Black	1.782	1.608	1.975	1.895	1.711	2.099
Hispanic	1.226	1.116	1.348	1.420	1.302	1.549

*White infants are reference group
Shaded cells indicate statistical significance

Table 45: Male Head Circumference: Odds Ratio for being born SGA by racial group and age category

Racial group	Odds ratio*	95% Wald confidence limits		Odds ratio*	95% Wald confidence limits	
	23 to 26 weeks			27 to 31 weeks		
Black	1.488	1.065	2.079	1.769	1.481	2.111
Hispanic	0.923	0.616	1.385	1.220	1.009	1.476
	32 to 36 weeks			37 to 41 weeks		
Black	2.999	2.712	3.316	2.843	2.580	3.133
Hispanic	1.558	1.411	1.720	1.721	1.578	1.878

*White infants are reference group
Shaded cells indicate statistical significance

Logistic regression was performed to test if the difference in black infants being at higher risk for being born SGA could be predicted by any of the maternal characteristics in the data set (Table 3). The odds ratios are available in Appendix 6, Tables 67-72. Controlling for race, the regression analysis showed that preeclampsia (pregnancy-induced hypertension and proteinuria) or eclampsia (preeclampsia with convulsions) was a predictor of SGA for weight, length, and head circumference for both genders and all age groups. Also controlling for race, report of smoking (reported by mother) was found to be a predictor of SGA for weight, length, and head circumference for the latter two age groups (32-36, 37-41 weeks) and both genders. Antenatal steroids (steroids given during a pregnancy at risk for preterm delivery to help the fetus's lungs to mature) were found to be a predictor in a few cases of race and age group. However, one would expect antenatal steroids to be found as a predictor for SGA because the steroids are given when the mother is at risk for going into premature labor. Diabetes and insulin were found to be predictor of a lower risk of SGA in some cases. This is also expected because women who experience diabetes with or without insulin generally have larger infants. Preeclampsia/eclampsia and smoking were the factors that may be actual predictors and will be discussed further.

Since smoking and preeclampsia/eclampsia were found to be consistent predictors of an increased risk of being born SGA for birth size, the prevalence (Table 46-47) of smoking and preeclampsia/eclampsia within these populations was examined. The prevalence of smoking as reported by the mothers was consistent within all age groups and both genders with whites having the highest reported levels, then black women, and finally Hispanics having the lowest levels.

The prevalence of pre-eclampsia/eclampsia was found to be the lowest among Hispanic mothers for both genders and in all age groups. In the latter two age groups (32-36, 37-41 weeks), black women were found to have the highest prevalence of pre-eclampsia/eclampsia for both male and female infants. Since pre-eclampsia/eclampsia was found to be a significant predictor of SGA birth weights for all infants, this factor may be a contributor to a higher rate of black infants with SGA births weights in both genders in the 32-36 week age group.

Table 46: Females: Prevalence of selected maternal characteristics by race and age group

Age group (weeks)	**Racial group**	**% with preeclampsia or eclampsia**	**P-value***	**% that reported smoking**	**P-value***
23-26 (n=1627)	Black Hispanic White	4.32 3.21 5.22	0.3228	2.52 0.87 9.34	<.0001
27-31 (n=6248)	Black Hispanic White	10.24 8.36 10.83	0.0394	3.35 1.02 7.36	<.0001
32-36 (n=20711)	Black Hispanic White	7.06 5.10 6.36	0.0004	3.36 0.98 5.97	<.0001
37-41 (n=20909)	Black Hispanic White	1.48 1.02 1.27	0.1329	2.68 0.62 4.81	<.0001
All groups (n=49495)	Black Hispanic White	5.53 3.40 4.74	<.0001	3.07 0.80 5.74	<.0001

*Statistical significance of chi-square (alpha=0.0125)

Table 47: Males: Prevalence of selected maternal characteristics by race and age group

Age group (weeks)	**Racial group**	**% with preeclampsia or eclampsia**	**P-value***	**% that reported smoking**	**P-value***
23-26 (n=1849)	Black Hispanic White	4.21 2.98 4.15	0.5294	2.63 1.38 7.35	<.0001
27-31 (n=7336)	Black Hispanic White	6.62 6.15 8.07	0.0184	3.49 0.97 7.63	<.0001
32-36 (n=27029)	Black Hispanic White	6.15 4.86 5.49	0.0135	3.88 0.84 5.54	<.0001
37-41 (n=29963)	Black Hispanic White	1.36 1.26 1.20	0.6948	2.34 0.59 4.36	<.0001
All groups (n=66177)	Black Hispanic White	4.18 3.18 3.80	<.0001	3.13 0.75 5.27	<.0001

*Statistical significance of chi-square (alpha=0.0125)

As a measure of health at birth, the APGAR scores (a combined test of five indicators of health at birth) were analyzed to investigate how the sickness of the infant correlated with the size-for-age classifications. A chi-square test (alpha=0.0125) was computed to analyze the prevalence of APGAR score at 1 minute with a score below 3 (APGAR1<3 – a common cutoff for identifying the sickest infants at birth) by SGA, AGA, and LGA categories for weight, length and head circumference for both genders and all four age groups. A higher percentage of infants (Tables 48-49) classified as SGA for weight had APGAR1<3. However, the chi-square test indicated that the differences in percentages were not significant in females in the 23-26 and 32-36 week age groups. The differences were significant for all other groups. The strength of the relationship between APGAR1<3 and SGA weight was calculated with logistic regression analysis. The significant odds ratios for infants with an APGAR1<3 of being born SGA for weight ranged from 1.3 – 1.6 depending on gender and age group (Appendix 7, Tables 73-74). This demonstrates that for most of the age groups, the infants that were SGA for weight were also sicker at birth.

Table 48: Females: Prevalence of APGAR at 1 minute < 3 by weight-for-age classification

Age group (weeks)	Classification for Weight			P-value*
	SGA%	AGA%	LGA%	
23-26	26.74	21.38	22.58	0.2722
26-31	12.97	8.88	6.13	<.0001
32-36	5.03	4.62	3.85	0.1660
37-41	9.77	7.84	8.04	0.0072

*Statistical significance of chi-square (alpha=0.0125)

Table 49: Males: Prevalence of APGAR at 1 minute < 3 by weight-for-age classification

Age group (weeks)	Classification for Weight			P-value*
	SGA%	AGA%	LGA%	
23-26	30.96	24.44	19.15	0.0255
26-31	13.36	9.49	9.45	0.0016
32-36	5.75	4.69	4.39	0.0218
37-41	9.59	6.99	7.06	<.0001

*Statistical significance of chi-square (alpha=0.0125)

As above, a chi-square test (alpha=0.0125) was also computed to analyze the prevalence of APGAR1<3 by SGA, AGA, and LGA categories for length and head circumference for both genders and all four age groups (data not shown). Statistically significant differences were only found for female head circumference classification in the age group of 27 to 31. There were no significant differences in female length classifications or male length and head circumference classifications. So, it appears that sickness at birth may not be associated with shorter lengths or smaller head circumferences.

The APGAR score at 5 minutes was analyzed for a score below 5 (APGAR5<5) which is also used to identify sick infants. The APGAR score at 5 minutes is thought of as a more important indicator of sickness than the APGAR score at 1 minute because it expresses a sustained level of sickness. As above, chi-square analysis (Appendix 8, Tables 75-76) was performed to analyze the prevalence of APGAR5<5 for the males and females by SGA, AGA, and LGA for all measurements (weight, length and head circumference) for all four age groups. The chi-square for APGAR5<5 was significant (alpha=0.0125) for only one comparison: male birth weight classification for age group 37-41 weeks (Appendix 8, Table 76).

The APGAR1<3 (Tables 50-51) was evaluated for the prevalence among the racial groups to determine if health at birth varied among the racial groups. Chi-square was computed to analyze the differences in prevalence of APGAR1<3 among the racial groups. In most age groups, black infants had the highest prevalence of an APGAR score at 1 minute of less than 3, Hispanics had the next highest rates, and whites had the lowest rates. This indicates a trend of black infants being more sick at birth. However, significant differences were not found in 23-26 week females and 23-26 and 27-31 week males.

The APGAR5<5 (Tables 52-53) was also evaluated for the prevalence among racial groups. It was found that black infants almost always had the greatest prevalence of an APGAR5<5 but was only a significant difference in 37-41 week females and 32-36 and 37-41 week males.

Table 50: Females: Prevalence of APGAR at 1 minute < 3 by racial group

Age group (weeks)	Black (%)	Hispanic (%)	White (%)	P-value*
23-26	24.28	20.6	20.41	0.3103
27-31	10.91	9.31	5.16	<.0001
32-36	5.79	4.65	3.58	<.0001
37-41	11.35	8.17	6.6	<.0001

*Statistical significance of chi-square (alpha=0.0125)

Table 51: Males: Prevalence of APGAR at 1 minute < 3 by racial group

Age group (weeks)	Black (%)	Hispanic (%)	White (%)	P-value*
23-26	26.14	26.45	19.95	0.0454
27-31	11.68	10.03	7.85	0.0026
32-36	6.34	4.56	4.27	<.0001
37-41	9.98	7.29	5.93	<.0001

*Statistical significance of chi-square (alpha=0.0125)

Table 52: Females: Prevalence of APGAR at 5 minutes < 5 by racial group

Age group (weeks)	Black (%)	Hispanic (%)	White (%)	P-value*
23-26	10.97	11.95	8.10	0.0833
27-31	3.90	3.20	3.50	0.5748
32-36	2.96	2.23	2.17	0.0162
37-41	5.14	3.77	3.91	0.0031

*Statistical significance of chi-square (alpha=0.0125)

Table 53: Males: Prevalence of APGAR at 5 minutes < 5 by racial group

Age group (weeks)	Black (%)	Hispanic (%)	White (%)	P-value*
23-26	12.28	8.94	12.22	0.1676
27-31	4.94	4.10	4.01	0.2729
32-36	3.00	2.44	2.10	0.0024
37-41	4.58	3.10	3.42	<.0001

*Statistical significance of chi-square (alpha=0.0125)

Discussion of Specific Aim II

The Pediatrix data set and new Pediatrix curves created in the previous section were analyzed to investigate the relationship of gender and race of the infants with birth size (weight, length and head circumference) and gestational age. The analysis of gender differences was primarily aimed at gathering information on whether gender-specific growth curves were warranted. The racial group analysis was aimed at examining differences in birth size by racial group with the intent to test if differences were significant in this population, to understand the clinical significance of the differences, and to determine if differences could be explained to some degree with the information in the data set. Overall, it was found that male infants were larger than females and that black infants were smaller than Hispanic infants who were smaller than white infants. It was also found that in many age ranges black infants were at a greater risk for being born SGA for all three growth measurements compared to white infants.

In the gender analysis, it was found that there were statistically significant differences between all of the mean birth size measurements (weight, length, and head circumference) of male and female infants as well as clinically significant differences in most of the comparisons. The largest difference was in birth weight where males were larger by 3-9% depending on percentile and gestational age. In birth length and head circumference, smaller but still significant differences were found in the 1-3% range. These findings were consistent with Thomas et al. (Thomas, Peabody, Turnier et al., 2000), Hindmarsh et al. (Hindmarsh, Geary, Rodeck et al., 2002) and Kramer et al. (Kramer, Olivier, McLean, Dougherty et al., 1990). Thomas et al. (Thomas, Peabody, Turnier et al., 2000) found that male infants (22-42 weeks) were larger by 0.095 kg, were

0.6 cm longer, and had larger head circumferences by 0.6 cm on average. This was also consistent with Hindmarsh et al. (Hindmarsh, Geary, Rodeck et al., 2002) who found significant differences between full-term male and female infants of 0.13 kg in weight, 0.8 cm in length, and 0.6 cm in head circumference. Kramer et al. (Kramer, Olivier, McLean, Dougherty et al., 1990) found that male infants weighed 4% more than females.

Although the differences in birth size were 3-9% for weight and just 1-3% for length and head circumference, these differences were statistically significant in all comparisons and clinically significant in the majority of comparisons; therefore gender-specific growth curves were warranted. As shown in Tables 27-29, depending on age group males are heavier by 48-143 g, longer by 0.5-0.9 cm, and have larger heads by 0.5-0.6 cm. The differences in length and head circumference may seem small but are similar to the weekly goals for growth in premature infants (~1.0 cm in length, 0.5-1.0 cm in head circumference) (Ellard, Olsen, and Sun, 2004) and are therefore overall clinically significant. The statistically significant differences in all birth size measurements in this and other studies (Hindmarsh, Geary, Rodeck et al., 2002; Kramer, Olivier, McLean, Dougherty et al., 1990; Thomas, Peabody, Turnier et al., 2000) also support the need for gender-specific growth curves.

As shown in Figures 25-30, when male and female infants were compared to combined-gender curves, such as the Lubchenco curves, greater percentages of female infants were classified as SGA compared to male infants. Greater percentages of male infants were classified as LGA compared to female infants. Across all gestational ages, 4.8% of males were classified on the Lubchenco curves as SGA for weight compared to 7.5% of females. In total, 15.4% of males were classified as LGA for weight while 9.9%

of females were classified as LGA. So, a large number of high-risk infants, especially males, would be misclassified as AGA instead of SGA for weight via the combined-gender Lubchenco curves and therefore may not receive the special attention in the NICU reserved for SGA infants. Thus, gender-specific curves are needed to prevent the increased misclassification of male infants.

In the analysis of racial group differences, it was found that more of the black infants were born at younger ages compared to white and Hispanic infants. Twenty-four percent of black females and 21% of black males were born at 23-31 weeks compare to only 15% of white females and 13% of white males. So, it seems that black women in the data set had earlier preterm deliveries. This was consistent with United States national statistics (Table 2) which show increased rates of preterm delivery for black women (B. E. Hamilton, Minino, Martin et al., 2007). This difference may indicate a need for improved prenatal health care for black women and increased research into the causes of the early preterm deliveries by black women.

Significant differences were found in birth size among the racial groups. Black infants were substantially smaller than white infants in all measurements across all gestational ages. Hispanic infants were found to be slightly smaller, not always significantly, than white infants but were larger than black infants. The finding regarding black infant size was consistent with United States national statistics (Table 2) which showed higher low birth weight rates for black women (B. E. Hamilton, Minino, Martin et al., 2007). The mean birth sizes (Tables 34-39) showed significant differences by racial group in almost all age groups for weight, length and head circumference. The largest difference was in male birth weight – black males were smaller than white males by 10.4% (0.254 kg) in the 32-36 week age group.

As discussed in the methods, the z-scores by racial group (Figures 37-42), calculated relative to the new Pediatrix curves, show the distribution of the infant measurements as a function of the standard deviation. The z-scores show large differences which were confirmed by the differences found in the mean weight, length, and head circumference between the racial groups. When the z-scores were expressed as percentiles, in the 32-36 weeks age groups, black infants range from 10 to 22 percentiles (relative to the 50th percentile) depending on gender and measurement below white infants. This demonstrates that the black infants in this sample would generally be plotted on growth curves at percentiles below white infants of the same gestational age. This indicates a difference in fetal growth of black infants compared to white infants which may impact the health of black infants.

The differences in birth size by racial group were consistent with the finding of differences in birth weight by Thomas et al. (Thomas, Peabody, Turnier et al., 2000) and Denham et al. (Denham, Schell, Gallo et al., 2001). Thomas et al. (Thomas, Peabody, Turnier et al., 2000) found that white and Hispanic infants (22-42 weeks) had larger birth sizes measurements than black infants – on average, larger birth weights by 90 g, longer lengths by 0.4 cm and larger head circumferences by 0.4 cm. These findings were also consistent with another United States study, Denham et al. (Denham, Schell, Gallo et al., 2001) who found significant differences between black and white full-term male infants in New York State with mean differences of 303 g in weight, 1.0 cm in length and 0.7 cm in head circumference.

Despite the differences in birth size among the racial groups, race-specific growth curves were not recommended. Race-specific growth curves would imply that the birth sizes of the various groups are appropriate for each group and should be accepted as the norm. However, since the reasons for the differences in birth size are not fully known and may be related to socio-economic status, these birth sizes should not currently be accepted as the norm for each group.

The percentage of infants that were classified as SGA or LGA by racial group based on the 10^{th} and 90^{th} percentiles from the new Pediatrix curves adds to the evidence that black infants were smaller and were more likely to be born SGA. With only a couple of exceptions, black infants were consistently found to have higher rates of SGA for all birth size measurements especially in the 32-26 and 37-41 week age groups. For example, at 32-36 weeks, 16.6% of black male infants were SGA for weight while only 8.1% white male infants were SGA for weight. Conversely, in most age ranges, white infants were found to have higher rates of being born LGA for birth size measurements. At 32-36 weeks, 5.8% of black male infants were LGA for weight while 12.2% of white male infants were LGA for weight.

The odds ratios for the risk of being born SGA also confirm these findings. It was found that black infants had higher odds of being born SGA – especially for birth weight and head circumference. These differences were the most striking for males in the latter two age groups (32-36, 37-41 weeks) where black infants had two times or higher risk of being born SGA than white infants.

The higher percentage of black infants classified as SGA and the higher odds ratios of black infants for being born SGA for all birth size measurements (weight, length, head circumference) provide information on the clinical significance of the difference in racial group birth sizes. Since SGA is considered the high-risk size-for-age category, this finding that black infants were more likely to be classified as SGA indicates that more black infants were at higher-risk for health complications at birth. This may suggest a need for more prenatal health care and nutrition programs for black women as well as more research into the causes of the increased SGA rates.

The information that was available in the data set on the mother and infant was used to try to explain why black infants had higher rates of being born SGA for weight, length, and head circumference. Controlling for race, logistic regression found that preeclampsia/eclampsia was a consistent predictor of being born SGA for weight, length and head circumference across all age groups. When comparing the odds ratios of the simple regression analysis (Tables 40-45) to the multivariate regression analysis (Appendix 6, Tables 67-72), the odds ratios for the racial groups remained consistent signifying that the effects of race and preeclampsia/eclampsia were independent. In the middle two age groups (32-36, 37-41 weeks), black women had a higher prevalence of preeclampsia/eclampsia than white or Hispanic women. This higher rate of preeclampsia/eclampsia was consistent with a study by Coonrod et al. (Coonrod, Hickok, Zhu et al., 1995) who found preeclampsia was more common in younger black women. The slightly increased prevalence of preeclampsia/eclampsia in black women may have contributed to their higher rates of SGA births.

The same regression analysis as above also found smoking to be a consistent predictor of being born SGA for weight, length and head circumference in the latter two age groups (32-36, 37-41 weeks). However, when the prevalence of smoking among the racial groups was analyzed it was found that white women consistently had the highest rates of smoking. So, perhaps fewer black women smoked during pregnancy but it was also possible that some black women did not report that they smoked during their pregnancy. Also, the amount of smoking was not included in the data set – so, heavy smokers were not differentiated from light smokers. However, there were likely other maternal risk-factors (such as prenatal care, nutrition, pre-pregnancy size, and pregnancy weight gain) involved in determining the infant's birth size which make these results difficult to interpret.

As a measure of the level of health of the infants, the APGAR scores were analyzed for their relationship with SGA birth sizes and with the racial groups. Infants with SGA birth weights were 1.3 – 1.6 more likely to be sicker at birth (APGAR at 1 minute of less than 3). However, this relationship was not found with APGAR at 5 minutes of less than 5 which may indicate that small size does not correlate well with sickness at birth. The APGAR score tests respiratory rate, reflexes, heart rate, color and muscle tone, and these factors may not be consistent with small birth weight.

In most age groups, black infants had the highest prevalence of an APGAR score at 1 minute of less than 3, Hispanics had the next highest rates, and whites had the lowest rates. This correlates with the birth size differences found by racial groups. The APGAR score at 5 minutes showed smaller differences but the same trend. So, the black infants in the data set were smaller and sicker at birth than white and Hispanic infants.

In general, this analysis was limited by the variables in the data set and the quality of the data. The data set included limited information on the mother, child and pregnancy which affected the completeness of the logistic regression analysis. Information which can affect birth size, such as socio-economic status, maternal body mass index, pregnancy weight gain, birth order, and amount of smoking, were not included in the data set. Another factor was that the race/ethnicities in the data set was determined based on the race of the mother, so the influence of multi-racial infants was not considered. However, this was consistent with many other studies of racial differences in birth size (Alexander, Kogan, and Himes, 1999; Overpeck, Hediger, Zhang et al., 1999). The racial demographics of the data set were similar to the United States (Table 8); so the conclusions can be generalized for the United States population and for use in NICUs across the country. The “other” group (10%) of infants which included Asian infants and unknown races was left out of the racial group analysis. In the future, research on birth size of the “other” racial groups in the United States would be worth investigating.

Another potential limitation in the data set was in the accuracy of the data reported by the mother. This information (such as smoking) was not guaranteed to be accurate since the mother may not have accurately answered questions, such as those regarding personal habits. Smoking may have been underreported since mothers may not have wanted to admit to smoking during pregnancy. If smoking was underreported evenly across races and gestational ages of the infants, then this would not have an effect on the results; however, this could not be tested in this study.

In the future, one could study the effect of socio-economic status on birth size. The zip code of the facility where the child was born could be linked to the census tract and the poverty levels could be analyzed. This would be very interesting to explore what percentage of size differences in the racial groups could be explained by socio-economic status after adjusting for race. If socio-economic status explains much of the differences in birth size, this would provide reinforcement for public health efforts which provide assistance to low-income women. Based on the results of this study, more prenatal health care may be needed in black women as indicated by lower infant APGAR scores and higher rates of maternal preeclampsia/eclampsia. So, future analysis could correlate APGAR scores and preeclampsia/eclampsia with socio-economic status.

Another area of future research could be neonatal mortality and morbidity within this sample population as stratified by race and gender. One could also investigate a correlation between neonatal mortality and morbidity and birth size as well as socio-economic status. Mortality rates would be another indication of the risk of the subpopulations. Depending on the results, this could also help to emphasize the need for improved public health for certain populations.

In this study, it was found that males were significantly larger than females and black infants were significantly smaller than Hispanic and white infants. These size differences may have important consequences. The size differences between male and female infants indicated that with combined-gender curves, males were underestimated as SGA in particular, and thus gender-specific growth curves were needed. The size differences between the racial groups demonstrated that black infants were at higher risk for being born SGA. It was also found that black infants were sicker at birth and are born earlier. While race-specific curves were not recommended, these findings indicated the need for more research into why black infants were smaller, younger, and sicker than white and Hispanic infants. Based on these results, improved prenatal care and increased nutrition programs for black women may be warranted.

CHAPTER 6. SUMMARY AND CONCLUSIONS

This study used a large heterogeneous data set of birth size measurements of 22 to 42 week infants collected in recent years (1998-2006) to examine the need for an update of the Lubchenco growth curves and to investigate differences in birth size between the genders and racial groups. The results of this study are significant to the health of infants as fetal and postnatal growth are key indicators of an infant's health and future adult health. Inadequate fetal growth can put an infant at higher risk for mortality, disease, complications and/or neurological delays (Kramer, Olivier, McLean, Willis et al., 1990). Slow postnatal growth can lead to neurological delays whereas fast postnatal growth has been correlated with metabolic syndrome later in life (Barker, Winter, Osmond et al., 1989; Lucas, 2005; Neu, Hauser, and Douglas-Escobar, 2007). Since growth curves are one of the primary tools used clinically and in epidemiological studies to assess fetal and postnatal growth, it is crucial that these are as accurate as possible.

Preterm infant growth curves that are currently in use are primarily the Lubchenco, Babson and Benda, and the Fenton curves. These curves are limited by older homogenous data sets, smaller sample sizes, varying ranges of gestational ages, and/or combined gender curves. The Fenton curves are also limited by the disparate data sources used for the different measurements. The new curves created in this study were compared to the Lubchenco curves because the actual percentile values were available for comparison and the curves are a good model that may be in need of an update.

In the first specific aim, new smoothed growth curves were created for males and females for the measurements of weight, length and head circumference using a large heterogeneous data set (from the Pediatrix Medical Group) with racial demographics similar to the United States collected in 33 states from 1998-2006. The curves were fit to the data using the LMS method by Cole and Green. The goodness of fit was assessed through the use of worm plots, z-scores, and visual inspection. The curves were then validated for the population by calculating the z-scores of the validation set relative to the new smoothed Pediatrix curves and ensuring that the z-scores had a mean close to zero and a standard deviation close to one. These curves were then compared to the Lubchenco curves visually, at selected points, and via the percentages of infants found to be SGA and LGA.

In comparing to the Lubchenco curves, it was found that the Lubchenco curves generally crossed the new Pediatrix curves. The new Pediatrix percentiles had smaller measurements until about 30 weeks, were somewhat similar between about 30 and 36 weeks, and then were larger after 36 weeks. This crossing of the curves led to varying underestimates and overestimates of infants classified as SGA or LGA depending on gestational age and gender. The underestimations and overestimations led to a large number of misclassifications of infants as SGA, AGA, or LGA by the Lubchenco curves. Approximately, 10% of females and 15% of males in the validation data set were misclassified by the Lubchenco curves compared to the new Pediatrix curves. This was a substantial number of misclassifications that warrant the replacement of the Lubchenco curves with new contemporary curves, such as the new Pediatrix curves created in this study. The new Pediatrix curves were created using a large data heterogeneous data set using rigorous statistical methods and could provide a needed update to the Lubchenco curves.

The second specific aim investigated the differences in size between males and females and among the racial groups. The analysis of the gender differences was primarily aimed at gathering information on whether gender-specific growth curves were warranted. It was found that the male and female curves were significantly different. Male infants had heavier weights by approximately 6% and the lengths and head circumferences were larger by approximately 2%. The increased misclassification by the Lubchenco curves of the male infants (15%) as compared to females (10%) was most likely due to the fact that the Lubchenco are combined-gender curves. So, individual curves were needed for males and female infants.

The racial group analysis was aimed at investigating differences in birth size with the intent to determine if the differences were significant in this population, to try to understand the impact of the differences, and to determine if differences could be explained by information in the data set. Significant differences were found in the mean gestational age among the racial groups. Black infants were born earlier than Hispanic and white infants. Significant differences in size were also found – black infants were always smaller than Hispanic and white infants. White infants were the largest. These differences were also found in the z-score distributions – when the z-scores were converted to percentiles relative to the median, black infants placed up to 19 percentiles below white infants depending on the gestational age and measurement. Black infants were also found to have higher odds ratios of being born SGA than both Hispanic and white infants – three times higher odds of male infants being SGA for head circumference.

To understand this phenomenon of size differences in greater detail, the maternal characteristics as well as infant APGAR scores were examined. For maternal characteristics, preeclampsia/eclampsia and smoking were found to be predictors of SGA for weight, length and head circumference. So, the prevalence of these factors among the various racial groups were examined and a possible link between higher rates of SGA for weight with preeclampsia/eclampsia in black women was found. For infant APGAR scores, the prevalence of APGAR at 1 minute less than 3 was found to be greater in black infants and in infants classified as SGA for weight. This indicates that black infants were likely to be sicker at birth and that their smaller weight was also correlated to sickness at birth.

Accurate preterm infant growth curves are crucial to the assessment of growth status and therefore are vital to the health of infants. This study found strong evidence for replacement of the Lubchenco growth curves with contemporary gender-specific curves. Based on this more recent heterogeneous data set demographically representative of the United States, many infants are being miscategorized as SGA, AGA or LGA by the Lubchenco curves. This study also found significant differences between male and female infants and among the racial groups. The differences between male and female infants reinforce the need for gender-specific curves. Whereas the differences among the racial groups in birth size, gestational age, and APGAR scores indicate that black infants are smaller, younger and sicker than white and Hispanic infants. This may signify a need for further research into the differences among racial groups as well as increased prenatal care for black women.

REFERENCES

Alexander, G. R., Himes, J. H., Kaufman, R. B., Mor, J., and Kogan, M. (1996). A United States national reference for fetal growth. *Obstet Gynecol, 87*(2), 163-168.

Alexander, G. R., Kogan, M. D., and Himes, J. H. (1999). 1994-1996 U.S. singleton birth weight percentiles for gestational age by race, Hispanic origin, and gender. *Matern Child Health J, 3*(4), 225-231.

American Academy of Pediatrics, Committee on Nutrition. Nutritional needs of low-birth-weight infants. (1977). *Pediatrics, 60*(4), 519-530.

Anderson, D. M. (2004). Nutrition for Low-Birth-Weight Infants. In L. K. Mahan & S. Escott-Stump (Eds.), *Krause's Food, Nutrition, and Diet Therapy* (11th ed., pp. 234-258). Philadelphia: Saunder's.

Arbuckle, T. E., Wilkins, R., and Sherman, G. J. (1993). Birth weight percentiles by gestational age in Canada. *Obstet Gynecol, 81*(1), 39-48.

Astolfi, P., and Zonta, L. A. (1999). Risks of preterm delivery and association with maternal age, birth order, and fetal gender. *Hum Reprod, 14*(11), 2891-2894.

Babson, S. G., and Benda, G. I. (1976). Growth graphs for the clinical assessment of infants of varying gestational age. *J Pediatr, 89*(5), 814-820.

Barker, D. J., Winter, P. D., Osmond, C., Margetts, B., and Simmonds, S. J. (1989). Weight in infancy and death from ischaemic heart disease. *Lancet, 2*(8663), 577-580.

Battaglia, F. C., and Lubchenco, L. O. (1967). A practical classification of newborn infants by weight and gestational age. *J Pediatr, 71*(2), 159-163.

Beeby, P. J., Bhutap, T., and Taylor, L. K. (1996). New South Wales population-based birthweight percentile charts. *J Paediatr Child Health, 32*(6), 512-518.

Bonellie, S., Chalmers, J., Gray, R., Greer, I., Jarvis, S., and Williams, C. (2008). Centile charts for birthweight for gestational age for Scottish singleton births. *BMC Pregnancy Childbirth, 8*, 5.

Brenner, W. E., Edelman, D. A., and Hendricks, C. H. (1976). A standard of fetal growth for the United States of America. *Am J Obstet Gynecol, 126*(5), 555-564.

Cameron, N. (1999). The use and abuse of growth charts. In F. E. Johnston, B. Zemel & P. B. Eveleth (Eds.), *Human growth in context*.

Cochran, W. D., and Lee, K. G. (2004). Assessment of the Newborn. In J. P. Cloherty, E. C. Eichenwald & A. R. Stark (Eds.), *Manual of Neonatal Care* (pp. 35 - 56). Philadelphia: Lippincott Williams, and Wilkins.

Cole, T. J., Freeman, J. V., and Preece, M. A. (1998). British 1990 growth reference centiles for weight, height, body mass index and head circumference fitted by maximum penalized likelihood. *Stat Med, 17*(4), 407-429.

Cole, T. J., and Green, P. J. (1992). Smoothing reference centile curves: the LMS method and penalized likelihood. *Stat Med, 11*, 1305-1319.

Cole, T. J., Henson, G. L., Tremble, J. M., and Colley, N. V. (1997). Birthweight for length: ponderal index, body mass index or Benn index? *Ann Hum Biol, 24*(4), 289-298.

Colle, E., Schiff, D., Andrew, G., Bauer, C. B., and Fitzhardinge, P. (1976). Insulin responses during catch-up growth of infants who were small for gestational age. *Pediatrics, 57*(3), 363-371.

Cooke, R. W., Lucas, A., Yudkin, P. L., and Pryse-Davies, J. (1977). Head circumference as an index of brain weight in the fetus and newborn. *Early Hum Dev, 1*(2), 145-149.

Coonrod, D. V., Hickok, D. E., Zhu, K., Easterling, T. R., and Daling, J. R. (1995). Risk factors for preeclampsia in twin pregnancies: a population-based cohort study. *Obstet Gynecol, 85*(5 Pt 1), 645-650.

Dancis, J., O'Connell, J. R., and Holt, L. E. (1948). A grid for recording the weight of premature infants. *J Pediatr, 33*, 570-572.

Denham, M., Schell, L. M., Gallo, M., and Stark, A. (2001). Neonatal size of low socio-economic status Black and White term births in Albany County, NYS. *Ann Hum Biol, 28*(2), 172-183.

Ehrenkranz, R. A. (2007). Estimated fetal weights versus birth weights: should the reference intrauterine growth curves based on birth weights be retired? *Arch Dis Child Fetal Neonatal Ed, 92*(3), F161-162.

Ehrenkranz, R. A., Younes, N., Lemons, J. A., Fanaroff, A. A., Donovan, E. F., Wright, L. L., et al. (1999). Longitudinal growth of hospitalized very low birth weight infants. *Pediatrics, 104*(2 Pt 1), 280-289.

Ellard, D., Olsen, I., and Sun, Y. (2004). Nutrition. In J. P. Cloherty, E. C. Eichenwald & A. R. Stark (Eds.), *Manual of Neonatal Care* (pp. 115 - 137). Philadelphia: Lippincott Williams, and Wilkins.

Erhardt, C. L., Joshi, G. B., Nelson, F. G., Kroll, B. H., and Weiner, L. (1964). Influence of Weight and Gestation on Perinatal and Neonatal Mortality by Ethnic Group. *Am J Public Health Nations Health, 54*, 1841-1855.

Eriksson, J. G., and Forsen, T. J. (2002). Childhood growth and coronary heart disease in later life. *Ann Med, 34*(3), 157-161.

Fenton, T. R. (2003). A new growth chart for preterm babies: Babson and Benda's chart updated with recent data and a new format. *BMC Pediatr, 3*(1), 13.

Fenton, T. R., McMillan, D. D., and Sauve, R. S. (1990). Nutrition and growth analysis of very low birth weight infants. *Pediatrics, 86*(3), 378-383.

Flegal, K. M. (1999). Curve smoothing and transformations in the development of growth curves. *Am J Clin Nutr, 70*(1), 163S-165S.

Gibson, R. (2005). *Principles of Nutritional Assessment* (2nd ed.). New York: Oxford University Press.

Goldenberg, R. L., Cutter, G. R., Hoffman, H. J., Foster, J. M., Nelson, K. G., and Hauth, J. C. (1989). Intrauterine growth retardation: standards for diagnosis. *Am J Obstet Gynecol, 161*(2), 271-277.

Hamilton, B. E., Martin, J. A., and Ventura, S. J. (2006). Births: Preliminary data for 2005. , *Health E-Stats.* (Released November 21, 2006. ed.).

Hamilton, B. E., Minino, A. M., Martin, J. A., Kochanek, K. D., Strobino, D. M., and Guyer, B. (2007). Annual summary of vital statistics: 2005. *Pediatrics, 119*(2), 345-360.

Hindmarsh, P. C., Geary, M. P., Rodeck, C. H., Kingdom, J. C., and Cole, T. J. (2002). Intrauterine growth and its relationship to size and shape at birth. *Pediatr Res, 52*(2), 263-268.

Hospital discharge of the high-risk neonate--proposed guidelines. American Academy of Pediatrics. Committee on Fetus and Newborn. (1998). *Pediatrics, 102*(2 Pt 1), 411-417.

Hovasi Cox, J., and Doorlag, D. (2000). Nutritional Concerns at Transfer or Discharge. In S. Groh-Wargo, M. Thompson, J. Hovasi Cox & J. V. Hartline (Eds.), *Nutritional Care for High-Risk Newborns* (3rd ed., pp. 549 - 565). Chicago: Precept Press, Inc.

Joseph, K. S., Kramer, M. S., Allen, A. C., Mery, L. S., Platt, R. W., and Wen, S. W. (2001). Implausible birth weight for gestational age. *Am J Epidemiol, 153*(2), 110-113.

Khoury, M. J., Erickson, J. D., Cordero, J. F., and McCarthy, B. J. (1988). Congenital malformations and intrauterine growth retardation: a population study. *Pediatrics, 82*(1), 83-90.

Kramer, M. S., McLean, F. H., Olivier, M., Willis, D. M., and Usher, R. H. (1989). Body proportionality and head and length 'sparing' in growth-retarded neonates: a critical reappraisal. *Pediatrics, 84*(4), 717-723.

Kramer, M. S., Olivier, M., McLean, F. H., Dougherty, G. E., Willis, D. M., and Usher, R. H. (1990). Determinants of fetal growth and body proportionality. *Pediatrics, 86*(1), 18-26.

Kramer, M. S., Olivier, M., McLean, F. H., Willis, D. M., and Usher, R. H. (1990). Impact of intrauterine growth retardation and body proportionality on fetal and neonatal outcome. *Pediatrics, 86*(5), 707-713.

Kramer, M. S., Platt, R. W., Wen, S. W., Joseph, K. S., Allen, A., Abrahamowicz, M., et al. (2001). A new and improved population-based Canadian reference for birth weight for gestational age. *Pediatrics, 108*(2), E35.

Kuczmarski, R. J., Ogden, C. L., Grummer-Strawn, L. M., Flegal, K. M., Guo, S. S., Wei, R., et al. (2000). CDC growth charts: United States. *Adv Data*(314), 1-27.

Lorenz, J. M., Wooliever, D. E., Jetton, J. R., and Paneth, N. (1998). A quantitative review of mortality and developmental disability in extremely premature newborns. *Arch Pediatr Adolesc Med, 152*(5), 425-435.

Lubchenco, L., Hansman, C., and Boyd, E. (1966). Intrauterine growth in length and head circumference as estimated from live births at gestational ages from 26 to 42 weeks. *Pediatrics, 37*(3), 403-408.

Lubchenco, L., Hansman, C., Dressler, M., and Boyd, E. (1963). Intrauterine growth as estimated from liveborn birth-weight data at 24 to 42 weeks of gestation. *Pediatrics, 32*, 793-800.

Lucas, A. (2005). Long-term programming effects of early nutrition -- implications for the preterm infant. *J Perinatol, 25 Suppl 2*, S2-6.

Marsal, K., Persson, P. H., Larsen, T., Lilja, H., Selbing, A., and Sultan, B. (1996). Intrauterine growth curves based on ultrasonically estimated foetal weights. *Acta Paediatr, 85*(7), 843-848.

Mathews, T. J., and MacDorman, M. F. (2006). Infant mortality statistics from the 2003 period linked birth/infant death data set. *Natl Vital Stat Rep, 54*(16), 1-29.

National Vital Statistics System. (2008). Retrieved June 3, 2008, 2008, from http://209.217.72.34/VitalStats/ReportFolders/reportFolders.aspx

Neu, J., Hauser, N., and Douglas-Escobar, M. (2007). Postnatal nutrition and adult health programming. *Semin Fetal Neonatal Med, 12*(1), 78-86.

Nevin-Folino, N. L. (2000). Neurological Impairment. In S. Groh-Wargo, M. Thompson, J. Hovasi Cox & J. V. Hartline (Eds.), *Nutritional Care for High-Risk Newborns* (3rd ed., pp. 521 - 533). Chicago: Precept Press, Inc.

Niklasson, A., Ericson, A., Fryer, J. G., Karlberg, J., Lawrence, C., and Karlberg, P. (1991). An update of the Swedish reference standards for weight, length and head circumference at birth for given gestational age (1977-1981). *Acta Paediatr Scand, 80*(8-9), 756-762.

Oishi, K., Honda, S., Takamura, N., Kusano, Y., Abe, Y., Moji, K., et al. (2004). Secular trends of sizes at birth in Japanese healthy infants born between 1962 and 1988. *J Physiol Anthropol Appl Human Sci, 23*(5), 155-161.

Oken, E., Kleinman, K. P., Rich-Edwards, J., and Gillman, M. W. (2003). A nearly continuous measure of birth weight for gestational age using a United States national reference. *BMC Pediatr, 3*, 6.

Overpeck, M. D., Hediger, M. L., Zhang, J., Trumble, A. C., and Klebanoff, M. A. (1999). Birth weight for gestational age of Mexican American infants born in the United States. *Obstet Gynecol, 93*(6), 943-947.

Reagan, P. B., and Salsberry, P. J. (2005). Race and ethnic differences in determinants of preterm birth in the USA: broadening the social context. *Soc Sci Med, 60*(10), 2217-2228.

Royston, P., and Wright, E. M. (2000). Goodness-of-fit statistics for age-specific reference intervals. *Stat Med, 19*(21), 2943-2962.

Schlesinger, E. R., and Allaway, N. C. (1955). The combined effect of birth weight and length of gestation on neonatal mortality among single premature births. *Pediatrics, 15*(6), 698-704.

Schmitt, S. K., Sneed, L., and Phibbs, C. S. (2006). Costs of newborn care in California: a population-based study. *Pediatrics, 117*(1), 154-160.

Thomas, P., Peabody, J., Turnier, V., and Clark, R. H. (2000). A new look at intrauterine growth and the impact of race, altitude, and gender. *Pediatrics, 106*(2), E21.

Thureen, P. J., and Hay, W. W. (2001). Early aggressive nutrition in preterm infants. *Semin Neonatol, 6*, 403-415.

Tucker, J., and McGuire, W. (2004). Epidemiology of preterm birth. *BMJ, 329*(7467), 675-678.

Tukey, J. W. (1977). *Exploratory Data Analysis*. Don Mills, Ontario: Addison-Wesley.

Usher, R., and McLean, F. (1969). Intrauterine growth of live-born Caucasian infants at sea level: standards obtained from measurements in 7 dimensions of infants born between 25 and 44 weeks of gestation. *J Pediatr, 74*(6), 901-910.

van Buuren, S., and Fredriks, M. (2001). Worm plot: a simple diagnostic device for modelling growth reference curves. *Stat Med, 20*, 1259-1277.

Villar, J., and Belizan, J. M. (1982). The relative contribution of prematurity and fetal growth retardation to low birth weight in developing and developed societies. *Am J Obstet Gynecol, 143*(7), 793-798.

Waterlow, J. C. (1972). Classification and definition of protein-calorie malnutrition. *Br Med J, 3*(5826), 566-569.

WHO Expert Committee. The newborn infant. (1995). In *Physical Status: The Use and Interpretation of Anthropometry* (pp. 161-262). Geneva: World Health Organization.

The World Health Organization (2007). *WHO Statistical Information System (WHOSIS): Low-birthweight newborns (percentage)*. Retrieved December 13, 2007, from http://www.who.int/whosis/indicators/2007LBW/en/.

Williams, R. L., Creasy, R. K., Cunningham, G. C., Hawes, W. E., Norris, F. D., and Tashiro, M. (1982). Fetal growth and perinatal viability in California. *Obstet Gynecol, 59*(5), 624-632.

Zeitlin, J., Saurel-Cubizolles, M. J., De Mouzon, J., Rivera, L., Ancel, P. Y., Blondel, B., et al. (2002). Fetal sex and preterm birth: are males at greater risk? *Hum Reprod, 17*(10), 2762-2768.

APPENDICES

APPENDIX 1: Institutional Review Board Approval

DREXEL UNIVERSITY
COLLEGE OF MEDICINE

Office of Research Compliance

APPROVAL NOTICE (EXEMPT)

TO: Irene E. Olsen-Shingara , Ph.D, RD
College of Arts & Science / Bioscience & Biotechnology
Mailstop:

FROM: Sreekant Murthy, Ph.D.

Sreekant Murthy, Ph.D.
Vice Provost for Research Compliance
Drexel University College of Medicine
245 N. 15th Street, MS 444, Philadelphia, PA 19102
Tel: 215-762-5078 Fax: 215-762-3722

SUBJECT: EXEMPT APPROVAL
TITLE: New Preterm Infant Growth Curves
SPONSOR: Drexel Grid Application
PROJECT No: 71443, PROTOCOL No: 16708 , ACTION No: 46981 Type: Amendment Period: 2 Seq: 1 , DETAIL No: 238675
CURRENT APPROVAL PERIOD: 09/27/2007, EXPIRES: 09/26/2008

RE: 09/27/07 - Approved Univ. Amend #1 - Request to increase the number of records to be anlayzed from 200,000 to 391, 861.

Date: 9/27/2007

On behalf of the Committee, I am pleased to inform you that the subject protocol has been reviewed and approved as **EXEMPT research** (45 CFR 46, 101(b)(1)) for the period indicated above. We operate under many Government requirements. As a result, this approval is granted with the following understandings:

1. If this is a sponsored project, then the study may not be activated until the Clinical Research Group has received BOTH a fully executed sponsored agreement AND appropriate letter(s) of indemnification by the sponsor. If this is not a sponsored study (designated "internal"), the costs of the project must be identified and a cost center designated. Please call 215-762-3453 if you have any questions regarding these procedures.
2. You must advise the IRB of the activation date. Use the attached form for this purpose.
3. Protected Health Information (PHI) cannot be collected without a Waiver of Authorization per HIPAA regulations.
4. Any change to the protocol must be submitted in writing and approved by the IRB in advance.
5. Any adverse reaction must be reported to the IRB as soon as it occurs.
6. Should the IRB decide to monitor your project directly, please cooperate fully. Failure to do so may result in withdrawal of this approval and notification to the sponsor and/or Federal agencies. Specific

245 North 15th Street, Mail Stop 444, Suite 2105 NCB • Philadelphia, PA 19102-1192 • Phone 215-762-3453 • Fax 215-762-3722
www.research.drexel.edu • www.drexelmed.edu

In the tradition of Woman's Medical College of Pennsylvania and Hahnemann Medical College®

Philadelphia Health & Education Corporation d/b/a Drexel University College of Medicine is a separate not-for-profit subsidiary of Drexel University.

information regarding monitoring appears in the book: "Guidelines for Biomedical and Behavioral Research Involving Human Subjects", obtainable through this office or vi the website http://research.drexel.edu.

7. Whether or not this protocol is activated, the IRB will conduct a Continuing Review at least annually. Should you fail to respond to this Federally-required progress report, the project may become ineligible for re-approval and the IRB may choose not to consider other projects for approval.
8. A final progress report must be submitted to the IRB in format similar to that of a periodic report.

The IRB welcomes your research project into the list of approved protocols. Your compliance with the above conditions will help to protect the continuation of all research activity at the University. With your project and others like it, we look forward to additions to knowledge of human health and benefits to science, our patients, and society.

cc: Dept Chair, Tenet, and Drexel

APPENDIX 2: Worm plot progression example

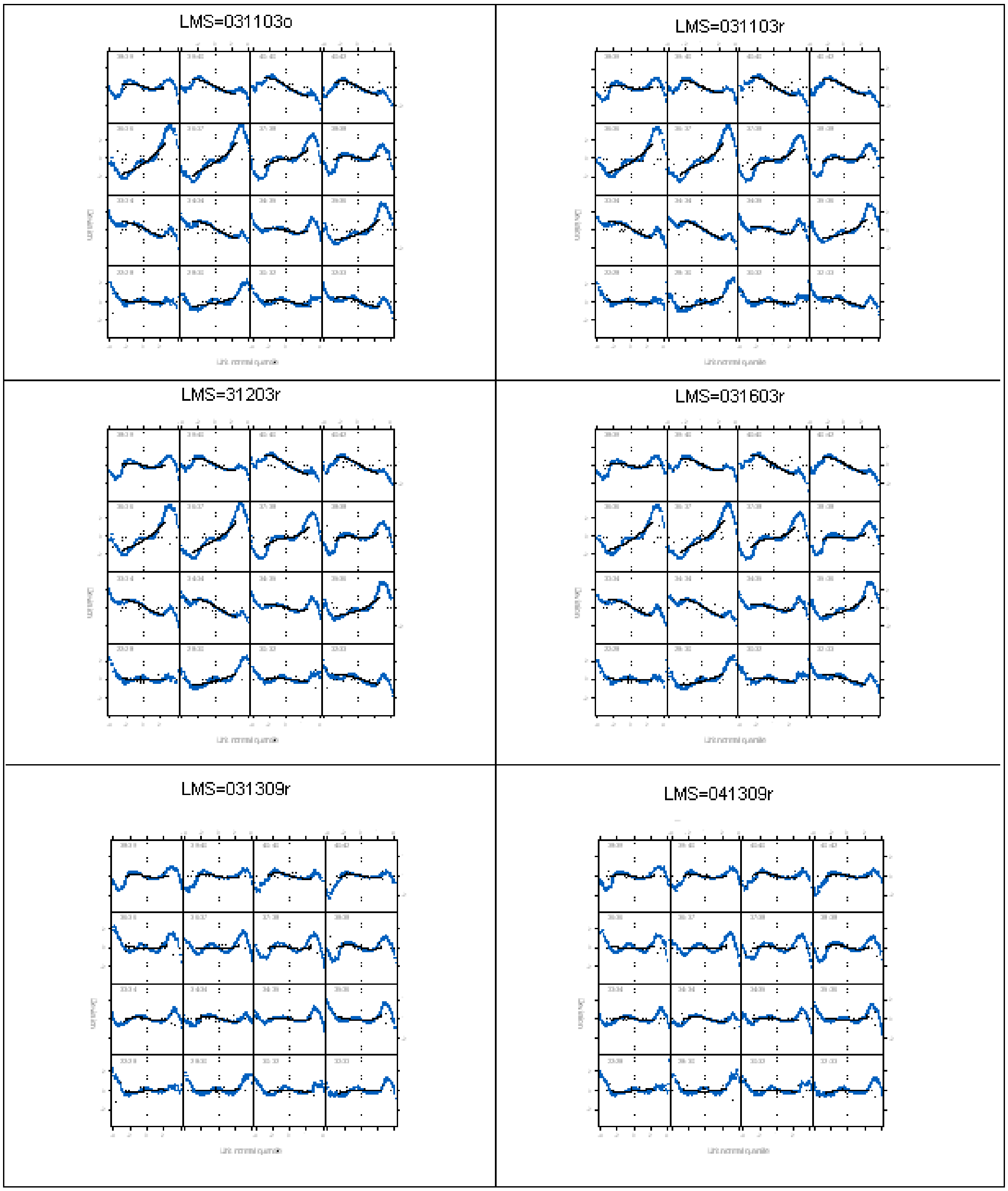


Figure 55: Female birth weight worm plots

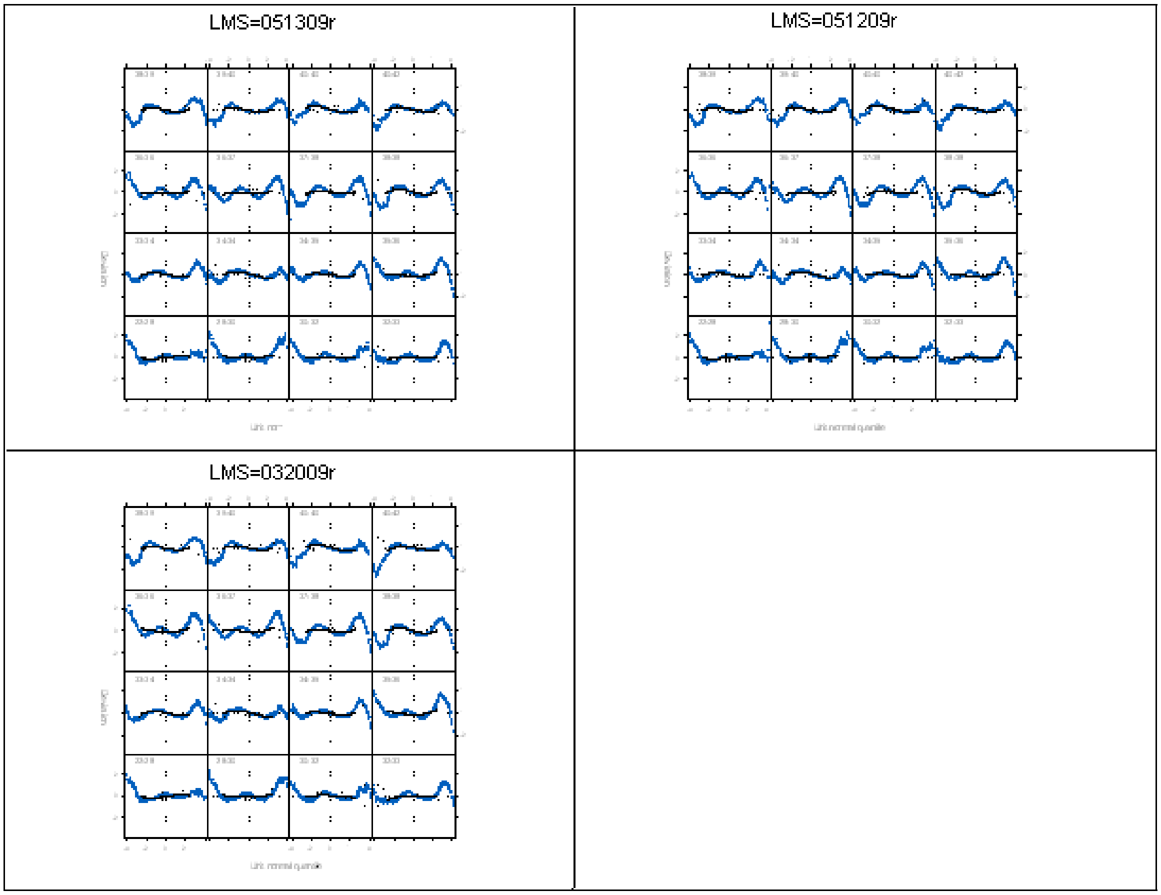


Figure 56: Female birth weight worm plots, continued

APPENDIX 3: LMS models

Table 54: LMS models analyzed

Gender	Measurement	LMS Model
Female	Birth weight	031103o, 031103r, 031203r, 031309r, 031603r, 032009r, 041309r, 051209r, 051309r
	Birth length	030703r, 030707r, 030907r, 031107r, 031607r, 070707r, 070907r, 071207r, 072007r
	Head circumference	030503r, 030703r, 030807r, 040808r, 042008r
Male	Birth weight	030908r, 031008r, 031107r, 031108r, 031203o, 031203r, 031208r, 031303r, 031608r, 031703r, 031808r, 032008r
	Birth length	030603r, 030606r, 030607r, 030705r, 030807r, 032007r, 040607r, 050807r, 050809r
	Head circumference	030705r, 030906r, 030907r, 030908r, 032008r, 032011r

APPENDIX 4: Preterm infant growth curve percentiles

Table 55: Female birth weight percentiles (kg)

Gestational age (weeks)	3rd	10th	25th	50th	75th	90th	97th
22	0.388	0.432	0.475	0.522	0.568	0.609	0.649
23	0.426	0.477	0.528	0.584	0.639	0.687	0.734
24	0.464	0.524	0.585	0.651	0.715	0.772	0.828
25	0.511	0.584	0.657	0.737	0.816	0.885	0.953
26	0.558	0.645	0.732	0.827	0.921	1.004	1.085
27	0.615	0.719	0.822	0.936	1.047	1.147	1.244
28	0.686	0.807	0.928	1.061	1.193	1.310	1.425
29	0.778	0.915	1.052	1.204	1.354	1.489	1.621
30	0.902	1.052	1.204	1.373	1.542	1.693	1.842
31	1.033	1.196	1.361	1.546	1.731	1.897	2.062
32	1.177	1.352	1.530	1.731	1.933	2.116	2.297
33	1.356	1.545	1.738	1.956	2.178	2.379	2.580
34	1.523	1.730	1.944	2.187	2.434	2.661	2.888
35	1.626	1.869	2.123	2.413	2.711	2.985	3.261
36	1.745	2.028	2.324	2.664	3.015	3.339	3.667
37	1.958	2.260	2.575	2.937	3.308	3.651	3.997
38	2.235	2.526	2.829	3.173	3.525	3.847	4.172
39	2.445	2.724	30.012	3.338	3.670	3.973	4.276
40	2.581	2.855	3.136	3.454	3.776	4.070	4.363
41	2.660	2.933	3.214	3.530	3.851	4.142	4.433
42	2.710	2.983	3.264	3.580	3.900	4.191	4.481

Table 56: Female birth length percentiles (cm)

Gestational age (weeks)	3rd	10th	25th	50th	75th	90th	97th
22	25.8	26.7	27.6	28.7	29.7	30.5	31.4
23	26.7	27.7	28.7	29.9	31.0	31.9	32.9
24	27.5	28.7	29.8	31.1	32.3	33.3	34.3
25	28.3	29.7	31.0	32.3	33.6	34.8	35.9
26	29.2	30.7	32.1	33.6	35.1	36.3	37.4
27	30.2	31.9	33.4	35.0	36.6	37.9	39.1
28	31.4	33.1	34.8	36.5	38.1	39.5	40.8
29	32.8	34.6	36.3	38.0	39.7	41.2	42.5
30	34.3	36.0	37.7	39.5	41.3	42.7	44.1
31	35.7	37.5	39.2	41.0	42.7	44.1	45.5
32	37.1	38.9	40.6	42.3	44.0	45.5	46.9
33	38.6	40.3	41.9	43.7	45.4	46.9	48.3
34	39.8	41.5	43.2	45.0	46.7	48.2	49.7
35	40.9	42.6	44.3	46.2	48.0	49.5	51.0
36	42.0	43.7	45.5	47.4	49.2	50.8	52.3
37	43.2	44.9	46.6	48.5	50.3	51.9	53.4
38	44.4	46.1	47.7	49.5	51.2	52.7	54.2
39	45.3	46.9	48.5	50.2	51.9	53.3	54.7
40	46.1	47.6	49.1	50.8	52.4	53.8	55.1
41	46.7	48.2	49.7	51.3	52.8	54.2	55.5
42	47.2	48.7	50.1	51.7	53.2	54.6	55.8

Table 57: Female head circumference percentiles (cm)

Gestational age (weeks)	3^{rd}	10^{th}	25^{th}	50^{th}	75^{th}	90^{th}	97th
22	18.0	18.6	19.3	20.0	20.7	21.3	21.9
23	18.8	19.5	20.1	20.9	21.6	22.2	22.9
24	19.6	20.3	21.0	21.8	22.5	23.2	23.8
25	20.4	21.1	21.9	22.7	23.4	24.1	24.8
26	21.2	22.0	22.7	23.6	24.4	25.1	25.9
27	21.9	22.8	23.6	24.5	25.4	26.2	27.0
28	22.7	23.7	24.6	25.5	26.5	27.3	28.1
29	23.6	24.6	25.5	26.5	27.5	28.4	29.2
30	24.6	25.6	26.5	27.5	28.5	29.4	30.2
31	25.5	26.5	27.4	28.4	29.4	30.3	31.1
32	26.5	27.4	28.3	29.3	30.3	31.2	32.0
33	27.3	28.3	29.2	30.2	31.2	32.1	33.0
34	28.1	29.1	30.1	31.1	32.2	33.1	34.0
35	28.8	29.8	30.8	31.9	33.0	34.0	34.9
36	29.4	30.5	31.5	32.7	33.8	34.8	35.8
37	30.1	31.1	32.2	33.3	34.4	35.4	36.3
38	30.7	31.7	32.7	33.7	34.8	35.7	36.7
39	31.1	32.0	33.0	34.0	35.1	36.0	36.9
40	31.4	32.3	33.3	34.3	35.3	36.1	37.0
41	31.7	32.6	33.5	34.5	35.5	36.3	37.1
42	32.0	32.9	33.8	34.7	35.7	36.5	37.3

Table 58: Male birth weight percentiles (kg)

Gestational age (weeks)	3rd	10th	25th	50th	75th	90th	97th
22	0.417	0.464	0.511	0.562	0.611	0.654	0.696
23	0.454	0.509	0.563	0.621	0.677	0.727	0.775
24	0.497	0.561	0.623	0.690	0.756	0.813	0.869
25	0.550	0.626	0.700	0.780	0.857	0.926	0.992
26	0.613	0.704	0.794	0.890	0.983	1.065	1.145
27	0.680	0.789	0.895	1.009	1.120	1.218	1.312
28	0.758	0.884	1.007	1.141	1.271	1.385	1.496
29	0.845	0.988	1.128	1.280	1.429	1.560	1.688
30	0.955	1.114	1.272	1.443	1.612	1.761	1.906
31	1.093	1.267	1.441	1.631	1.818	1.984	2.147
32	1.246	1.433	1.622	1.829	2.034	2.218	2.398
33	1.422	1.625	1.830	2.057	2.284	2.488	2.688
34	1.589	1.810	2.035	2.285	2.536	2.763	2.987
35	1.728	1.980	2.238	2.527	2.819	3.084	3.348
36	1.886	2.170	2.462	2.792	3.127	3.432	3.737
37	2.103	2.401	2.708	3.056	3.411	3.736	4.060
38	2.356	2.652	2.959	3.306	3.661	3.986	4.312
39	2.545	2.833	3.131	3.469	3.813	4.129	4.446
40	2.666	2.950	3.245	3.579	3.919	4.232	4.545
41	2.755	3.039	3.333	3.666	4.007	4.319	4.633
42	2.825	3.109	3.403	3.737	4.077	4.389	4.703

Table 59: Male birth length percentiles (cm)

Gestational age (weeks)	3rd	10th	25th	50th	75th	90th	97th
22	26.0	27.0	28.0	29.0	30.1	31.0	31.9
23	26.9	28.0	29.1	30.3	31.4	32.4	33.4
24	27.9	29.1	30.3	31.5	32.8	33.9	34.9
25	28.8	30.2	31.5	32.9	34.2	35.4	36.5
26	29.9	31.3	32.8	34.3	35.7	37.0	38.2
27	31.0	32.6	34.1	35.7	37.3	38.6	39.8
28	32.2	33.9	35.5	37.2	38.8	40.2	41.5
29	33.5	35.2	36.9	38.7	40.3	41.7	43.1
30	34.8	36.6	38.3	40.1	41.8	43.2	44.6
31	36.2	38.0	39.8	41.6	43.3	44.7	46.1
32	37.7	39.5	41.2	43.0	44.7	46.1	47.5
33	39.1	40.9	42.6	44.4	46.1	47.5	48.9
34	40.4	42.2	43.9	45.7	47.4	48.9	50.3
35	41.5	43.3	45.0	46.9	48.6	50.2	51.6
36	42.7	44.5	46.2	48.1	49.9	51.5	53.0
37	44.0	45.7	47.4	49.3	51.1	52.6	54.1
38	45.2	46.8	48.5	50.2	52.0	53.5	55.0
39	46.1	47.7	49.3	51.0	52.7	54.2	55.6
40	46.9	48.4	49.9	51.6	53.2	54.7	56.1
41	47.5	49.0	50.5	52.1	53.7	55.1	56.5
42	48.1	49.5	51.0	52.6	54.2	55.6	56.9

Table 60: Male head circumference percentiles (cm)

Gestational age (weeks)	3^{rd}	10^{th}	25^{th}	50^{th}	75^{th}	90^{th}	97th
22	18.5	19.2	19.8	20.5	21.1	21.7	22.3
23	19.3	20.0	20.6	21.3	22.0	22.7	23.3
24	20.1	20.8	21.5	22.2	23.0	23.6	24.3
25	20.9	21.7	22.4	23.2	23.9	24.6	25.3
26	21.8	22.5	23.3	24.2	25.0	25.7	26.4
27	22.6	23.5	24.3	25.2	26.0	26.8	27.6
28	23.5	24.3	25.2	26.1	27.1	27.9	28.6
29	24.3	25.2	26.1	27.1	28.0	28.8	29.6
30	25.1	26.1	27.0	28.0	29.0	29.8	30.6
31	26.0	27.0	27.9	28.9	29.9	30.8	31.6
32	26.9	27.8	28.8	29.9	30.9	31.8	32.6
33	27.7	28.7	29.7	30.8	31.8	32.7	33.6
34	28.5	29.5	30.5	31.6	32.7	33.6	34.6
35	29.2	30.3	31.3	32.4	33.6	34.5	35.5
36	29.9	31.0	32.1	33.2	34.3	35.3	36.3
37	30.6	31.7	32.7	33.9	35.0	36.0	36.9
38	31.2	32.2	33.2	34.4	35.5	36.4	37.3
39	31.5	32.5	33.5	34.6	35.7	36.6	37.6
40	31.8	32.8	33.8	34.8	35.9	36.8	37.7
41	32.0	33.0	34.0	35.0	36.1	37.0	37.8
42	32.3	33.3	34.2	35.2	36.2	37.1	38.0

APPENDIX 5: LMS values

Table 61: Female birth weight (031309r)
L, M, and S curve values

Gestational age (weeks)	L curve value	M curve value	S curve value
22	1.20883	0.522158	0.132266
23	1.194818	0.584096	0.14013
24	1.179832	0.650553	0.148521
25	1.1606	0.73716	0.159025
26	1.140124	0.827439	0.168996
27	1.115621	0.935625	0.178258
28	1.086843	1.061267	0.184654
29	1.053726	1.203761	0.185821
30	1.015403	1.373208	0.182064
31	0.975137	1.545694	0.177066
32	0.930277	1.730702	0.172321
33	0.877743	1.956463	0.166532
34	0.826687	2.186888	0.166241
35	0.785509	2.413013	0.180568
36	0.764455	2.664464	0.192323
37	0.771449	2.936585	0.185106
38	0.796105	3.172795	0.1626
39	0.821395	3.337971	0.146082
40	0.841025	3.453644	0.137347
41	0.85465	3.530244	0.133659
42	0.863869	3.580331	0.131683

Table 62: Female birth length (070707r)
L, M, and S curve values

Gestational age (weeks)	L curve value	M curve value	S curve value
22	1.430482	28.66194	0.052
23	1.613453	29.86142	0.0551
24	1.798933	31.07351	0.0583
25	2.005058	32.32348	0.0616
26	2.233793	33.63787	0.0645
27	2.395378	35.04654	0.0666
28	2.395816	36.52238	0.0677
29	2.231128	38.04121	0.0675
30	2.124795	39.54493	0.0657
31	2.21724	40.96124	0.0632
32	2.171717	42.33434	0.060643
33	1.938254	43.7008	0.0588
34	1.784309	44.97777	0.0582
35	1.745924	46.16707	0.0583
36	1.749782	47.35613	0.0579
37	1.809164	48.5012	0.0558
38	1.839674	49.47562	0.0524
39	1.882797	50.21408	0.0495
40	1.921904	50.78839	0.0474
41	1.958261	51.27105	0.045743
42	1.996645	51.69974	0.0443

Table 63: Female head circumference (040808r)
L, M, and S curve values

Gestational age (weeks)	L curve value	M curve value	S curve value
22	1.266161	19.97377	0.0523
23	1.337864	20.86304	0.0518
24	1.411841	21.75909	0.0514
25	1.49991	22.66695	0.0516
26	1.598815	23.58433	0.0527
27	1.684759	24.54104	0.0544
28	1.740237	25.52902	0.0556
29	1.770731	26.52545	0.0553
30	1.77809	27.50703	0.053901
31	1.773913	28.43328	0.0519
32	1.754456	29.33278	0.0502
33	1.719627	30.24052	0.049566
34	1.677405	31.12611	0.0503
35	1.647172	31.94374	0.0513
36	1.62873	32.67503	0.0516
37	1.613228	33.28869	0.0499
38	1.601345	33.74182	0.0472
39	1.59663	34.04006	0.045
40	1.595098	34.27305	0.0433
41	1.594508	34.49748	0.041771
42	1.594251	34.71341	0.0403

Table 64: Male birth weight (031208r)
L, M, and S curve values

Gestational age (weeks)	L curve value	M curve value	S curve value
22	1.327546	0.562053	0.131449
23	1.317082	0.620887	0.136491
24	1.304797	0.690427	0.142397
25	1.289169	0.77979	0.149732
26	1.269933	0.889783	0.157988
27	1.248475	1.008829	0.165406
28	1.223924	1.140789	0.171068
29	1.196221	1.28031	0.17407
30	1.162495	1.443446	0.174382
31	1.121586	1.63053	0.1713
32	1.074563	1.828643	0.167196
33	1.021127	2.057311	0.163514
34	0.967863	2.284874	0.162752
35	0.917	2.527055	0.170555
36	0.873458	2.791799	0.176554
37	0.83925	3.056463	0.170541
38	0.81368	3.306359	0.157555
39	0.796418	3.468678	0.145879
40	0.784917	3.578544	0.13982
41	0.775949	3.666472	0.136355
42	0.76872	3.736578	0.133754

Table 65: Male birth length (030807r)
L, M, and S curve values

Gestational age (weeks)	L curve value	M curve value	S curve value
22	1.468228	29.03106	0.0541
23	1.60782	30.28212	0.0566
24	1.747826	31.54272	0.0592
25	1.889089	32.86174	0.061772
26	2.029245	34.26546	0.0639
27	2.158662	35.73151	0.0653
28	2.264038	37.22118	0.0658
29	2.335083	38.66919	0.0653
30	2.366663	40.10986	0.0642
31	2.357345	41.55544	0.0623
32	2.313944	42.97448	0.0602
33	2.244372	44.36844	0.0581
34	2.162911	45.65795	0.0568
35	2.07372	46.87566	0.0568
36	1.973078	48.10245	0.0566
37	1.86604	49.27039	0.0546
38	1.768593	50.24886	0.0517
39	1.689483	51.00745	0.0491
40	1.625893	51.61172	0.0473
41	1.57212	52.1219	0.04591
42	1.523402	52.58268	0.0447

Table 66: Male head circumference (030908r)
L, M, and S curve values

Gestational age (weeks)	L curve value	M curve value	S curve value
22	1.392138	20.46179	0.0488
23	1.450425	21.34904	0.049
24	1.509582	22.24597	0.0494
25	1.569887	23.17379	0.05
26	1.629859	24.16381	0.051
27	1.685084	25.17773	0.051898
28	1.726563	26.14511	0.0523
29	1.749613	27.06205	0.0522
30	1.755509	27.98724	0.0521
31	1.747174	28.93489	0.0516
32	1.734149	29.85488	0.051188
33	1.726922	30.75831	0.051
34	1.728923	31.63045	0.050867
35	1.736439	32.44773	0.0511
36	1.743926	33.21239	0.0512
37	1.747784	33.86927	0.0499
38	1.749735	34.35767	0.0478
39	1.750805	34.63842	0.0462
40	1.751585	34.83157	0.0451
41	1.752418	35.03332	0.0439
42	1.753254	35.23404	0.0428

APPENDIX 6: Logistic regression analysis of predictors of SGA at birth

Table 67: Female birth weight: Odds ratios of predictors of SGA

Age group (weeks)	Effect*	Odds ratio	95% Wald confidence limits	
23 to 26	Diabetes	3.354	1.775	6.340
	Preeclampsia	9.396	5.812	15.190
27 to 31	Hispanic	0.782	0.625	0.980
	Maternal age	1.016	1.004	1.028
	Black	1.277	1.063	1.534
	Preeclampsia	2.793	2.277	3.428
32 to 36	Insulin	0.463	0.351	0.610
	Maternal age	1.008	1.001	1.014
	Antenatal steroids	1.207	1.102	1.321
	Other	1.319	1.124	1.549
	Hispanic	1.350	1.206	1.512
	Black	1.844	1.648	2.064
	Smoking	2.289	1.923	2.724
	Preeclampsia	2.573	2.241	2.953
37 to 41	Insulin	0.495	0.291	0.843
	Diabetes	0.532	0.382	0.742
	Hispanic	1.395	1.249	1.559
	Other	1.698	1.465	1.967
	Antenatal steroids	1.973	1.332	2.923
	Black	2.218	1.961	2.508
	Preeclampsia	2.391	1.753	3.261
	Smoking	3.199	2.649	3.862

*Effects are ordered by magnitude of the odds ratio

Preeclampsia and smoking are bold as they are the consistent predictors

Table 68: Female birth length: Odds ratios of predictors of SGA

Age group (weeks)	Effect*	Odds ratio	95% Wald confidence limits	
23 to 26	**Preeclampsia**	3.883	2.290	6.584
27 to 31	Maternal age	1.014	1.002	1.026
	Black	1.216	1.017	1.454
	Insulin	1.758	1.208	2.558
	Preeclampsia	2.453	1.989	3.027
32 to 36	Diabetes	0.658	0.545	0.794
	Antenatal steroids	1.139	1.039	1.247
	Hispanic	1.216	1.087	1.361
	Other	1.368	1.170	1.600
	Black	1.548	1.381	1.736
	Preeclampsia	1.630	1.395	1.903
	Smoking	2.136	1.794	2.545
37 to 41	Diabetes	0.570	0.456	0.712
	Hispanic	1.222	1.095	1.363
	Other	1.317	1.131	1.533
	Preeclampsia	1.478	1.036	2.110
	Antenatal steroids	1.636	1.084	2.471
	Black	1.909	1.690	2.157
	Smoking	2.537	2.087	3.085

*Effects are ordered by magnitude of the odds ratio
Preeclampsia and smoking are bold as they are the consistent predictors

Table 69: Female head circumference: Odds ratios of predictors of SGA

Age group (weeks)	Effect*	Odds ratio	95% Wald confidence limits	
23 to 26	**Preeclampsia**	2.052	1.099	3.832
	Diabetes	3.333	1.782	6.237
27 to 31	Antenatal steroids	1.217	1.004	1.475
	Black	1.542	1.292	1.840
	Preeclampsia	1.555	1.224	1.977
32 to 36	Insulin	0.549	0.420	0.716
	Antenatal steroids	1.389	1.269	1.521
	Hispanic	1.484	1.330	1.657
	Preeclampsia	1.534	1.309	1.798
	Smoking	1.851	1.535	2.232
	Black	2.305	2.071	2.565
37 to 41	Diabetes	0.552	0.445	0.685
	Maternal age	0.987	0.980	0.993
	Hispanic	1.549	1.396	1.720
	Antenatal steroids	1.630	1.096	2.422
	Preeclampsia	1.648	1.189	2.285
	Other	1.726	1.499	1.987
	Smoking	2.265	1.860	2.759
	Black	2.402	2.138	2.698

*Effects are ordered by magnitude of the odds ratio
Preeclampsia and smoking are bold as they are the consistent predictors

Table 70: Male birth weight: Odds ratios of predictors of SGA

Age group (weeks)	Effect*	Odds ratio	95% Wald confidence limits	
23 to 26	Hispanic	0.648	0.428	0.981
	Maternal age	1.024	1.002	1.048
	Preeclampsia	8.764	5.481	14.013
27 to 31	Maternal age	1.014	1.003	1.025
	Antenatal steroids	1.240	1.045	1.472
	Black	1.243	1.048	1.475
	Preeclampsia	3.737	3.067	4.555
32 to 36	Diabetes	0.470	0.392	0.565
	Antenatal steroids	1.259	1.161	1.364
	Hispanic	1.364	1.237	1.505
	Other race	1.509	1.316	1.731
	Smoking	1.944	1.650	2.291
	Black	2.300	2.079	2.545
	Preeclampsia	2.896	2.548	3.290
37 to 41	Diabetes	0.518	0.426	0.629
	Maternal age	0.986	0.980	0.992
	Hispanic	1.546	1.411	1.694
	Other	1.674	1.478	1.895
	Preeclampsia	1.893	1.440	2.489
	Black	2.274	2.050	2.522
	Antenatal steroids	2.306	1.698	3.132
	Smoking	2.767	2.341	3.270

*Effects are ordered by magnitude of the odds ratio
Preeclampsia and smoking are bold as they are the consistent predictors

Table 71: Male birth length: Odds ratios of predictors of SGA

Age group (weeks)	Effect*	Odds ratio	95% Wald confidence limits	
23 to 26	**Preeclampsia**	4.869	2.976	7.967
27 to 31	Maternal age	1.014	1.003	1.025
	Smoking	1.373	1.021	1.846
	Preeclampsia	2.578	2.089	3.181
32 to 36	Diabetes	0.592	0.501	0.700
	Antenatal steroids	1.111	1.025	1.205
	Hispanic	1.287	1.169	1.416
	Other	1.384	1.209	1.584
	Black	1.812	1.634	2.009
	Smoking	1.842	1.565	2.168
	Preeclampsia	1.876	1.631	2.158
37 to 41	Diabetes	0.649	0.545	0.774
	Maternal age	0.987	0.981	0.993
	Hispanic	1.445	1.323	1.580
	Preeclampsia	1.447	1.080	1.938
	Other	1.501	1.328	1.697
	Black	1.915	1.727	2.124
	Antenatal steroids	1.945	1.419	2.668
	Smoking	2.136	1.793	2.543

*Effects are ordered by magnitude of the odds ratio
Preeclampsia and smoking are bold as they are the consistent predictors

Table 72: Male head circumference: Odds ratios of predictors of SGA

Age group (weeks)	Effect*	Odds ratio	95% Wald confidence limits	
23 to 26	Black	1.496	1.106	2.025
	Preeclampsia	3.685	2.218	6.122
27 to 31	Hispanic	1.231	1.022	1.483
	Diabetes	1.399	1.041	1.882
	Black	1.774	1.493	2.109
	Preeclampsia	1.816	1.435	2.298
32 to 36	Diabetes	0.562	0.469	0.673
	Maternal age	0.988	0.982	0.994
	Antenatal steroids	1.333	1.228	1.447
	Preeclampsia	1.384	1.182	1.621
	Smoking	1.527	1.274	1.831
	Hispanic	1.590	1.438	1.758
	Other	1.703	1.480	1.959
	Black	2.971	2.683	3.290
37 to 41	Diabetes	0.587	0.490	0.704
	Maternal age	0.976	0.971	0.982
	Preeclampsia	1.430	1.067	1.915
	Other	1.610	1.420	1.826
	Hispanic	1.700	1.555	1.859
	Antenatal steroids	1.792	1.293	2.483
	Smoking	1.977	1.650	2.370
	Black	2.805	2.541	3.095

*Effects are ordered by magnitude of the odds ratio
Preeclampsia and smoking are bold as they are the consistent predictors

APPENDIX 7: Logistic regression analysis of APGAR at 1 minute < 3 as a predictor of SGA birthweight

Table 73: Female birth weight: Odds ratios of APGAR at 1 minute < 3 as a predictor of SGA birth weight

Age group (weeks)	Odds ratio	95% Wald confidence limits	
23 to 26	1.333	0.932	1.908
27 to 31	1.583	1.247	2.011
32 to 36	1.114	0.914	1.357
37 to 41	1.270	1.093	1.475

Table 74: Male birth weight: Odds ratios of APGAR at 1 minute < 3 as a predictor of SGA birth weight

Age group (weeks)	Odds ratio	95% Wald confidence limits	
23 to 26	1.429	1.037	1.969
27 to 31	1.471	1.190	1.818
32 to 36	1.251	1.061	1.475
37 to 41	1.411	1.244	1.600

APPENDIX 8: Prevalence of APGAR at 5 minutes < 5 by weight-for-age classification

Table 75: Females: Prevalence of APGAR at 5 minutes < 5 by weight-for-age classification

Age group (weeks)	Classification for Weight			P-value*
	SGA%	AGA%	LGA%	
23-26	11.05	9.80	14.84	0.1420
26-31	4.03	3.54	3.71	0.7985
32-36	2.41	2.41	2.02	0.5571
37-41	4.68	3.91	4.13	0.2089

*Statistical significance of chi-square (alpha=0.0125)

Table 76: Males: Prevalence of APGAR at 5 minutes < 5 by weight-for-age classification

Age group (weeks)	Classification for Weight			P-value*
	SGA%	AGA%	LGA%	
23-26	14.21	11.48	7.45	0.1069
26-31	4.76	4.13	5.50	0.1897
32-36	2.32	2.41	1.78	0.0965
37-41	4.73	3.47	2.79	0.0001

*Statistical significance of chi-square (alpha=0.0125)