This is to certify that the
dissertation entitled

MATERNAL PHYSICAL ACTIVITY AND BIRTH WEIGHT: A META-ANALYSIS

presented by

COOKER CANDACE D. PERKINS

has been accepted towards fulfillment
of the requirements for the

Ph.D. degree in KINESIOLOGY

James M Pivarnik
Major Professor's Signature

12/17/04
Date

MATERNAL PHYSICAL ACTIVITY AND BIRTH WEIGHT: A META-ANALYSIS

By

Cooker Candace D. Perkins

A DISSERTATION

Submitted to
Michigan State University
in partial fulfillment of the requirements
for the degree of

DOCTOR OF PHILOSOPHY

Department of Kinesiology

2004

ABSTRACT

MATERNAL PHSYICAL ACTIVITY AND BIRTH WEIGHT: A META-ANALYSIS

By

Cooker Candace D. Perkins

Introduction: Although CDC/ACSM recommendations for physical activity are endorsed by ACOG, little information is available indicating the effect of women following these guidelines on fetal growth. Results from previous studies are equivocal with respect to whether physical activity is related to birth weight (BW). The purpose of this dissertation was to conduct a meta-analysis in order to estimate the extent to which CDC/ACSM recommendations for moderate and vigorous physical activity during pregnancy influence BW. An additional purpose was to examine how this relationship may be influenced by various subject and study factors. **Methods**: A thorough literature search was performed using NLM Medline, Embase, and manual bibliographic searches. Studies were included if they contained some measure/report of physical activity during pregnancy (with a sedentary/control group) and reported BW as a continuous variable within normal BW range (2500-4500 grams). Individual effect size estimates (Cohen's *d*) were calculated for each study. The Dersimonian and Laird random effects model was used to calculate the pooled effect size (ES) of BW by maternal physical activity across all studies. A formal test for statistical heterogeneity was performed on the overall analysis. Pre-specified subgroups were used to explore potential heterogeneity in the full analysis. Meta-regression was utilized to estimate the variance in difference in BW between exercising and nonexercising women that can be explained by difference in maternal weight gain between groups. **Results**: Of the 2418 potentially relevant titles located, 192 abstracts were reviewed, 52 were selected for full-text review, and 18 studies satisfied inclusion criteria. Study participants were mostly Caucasian, nonsmokers, and nulliparous. Individual effect sizes for all studies ranged from -1.51 to 0.52, a difference in mean birth weight ranging from -509.9 to 300 grams between exercisers and nonexercisers. The pooled effect size of all 23 comparisons was significantly heterogeneous, with an overall effect size and 95% confidence interval (95% CI) of -0.12 (-0.30 – 0.06). Some of the heterogeneity in the full analysis was explained by subgroups analyses. Among these subgroup analyses, there was nonsignificant heterogeneity among moderate intensity physical activity studies (ES=0.05; $P=0.49$), studies where exercisers gained more weight than nonexercisers (ES=0.02; $P=0.06$), studies where the mean age of women was ≥27 years (ES=0.10; $P=0.06$), and in studies where physical activity data were collected via retrospective self-report (ES=0.22; $P=0.82$). Meta-regression of BW by maternal weight gain was significant (coeff=0.17; $P<0.01$). **Conclusions**: Moderate physical activity performed throughout pregnancy did not affect BW. The impact of vigorous physical activity on BW remains unclear. Differences in maternal weight gain partially explain the differences in BW among exercisers. Although the effect of maternal physical activity on BW was not uniform across all studies, future studies on BW and the incidence of later life morbidities should include careful measures of maternal physical activity.

ACKNOWLEDGEMENTS

I would like to give sincere thanks to the members of my dissertation committee, Dr. Jim Pivarnik, Dr. Mathew Reeves, Dr. Deborah Feltz, and Dr. Chris Womack, whose guidance and support made the completion of this project possible. A special note of gratitude to my advisor, Dr. Jim Pivarnik, for his unyielding dedication to my graduate studies over the past four years. I am grateful for the endless opportunities he offered and the guidance he provided on each and every task. I am also very grateful to Dr. Mathew Reeves, Dr. Mahukar Pai, and Dr. Jack Colford, for their expertise in meta-analyses and the extra time they invested in this project. To my other mentors, Karin Pfeifer and Dawn Coe, for their wisdom from experience. I am thankful for the assistance and friendship of JoAnn Janes, who was a valuable part of my graduate experience, endlessly offering her time.

I am indebted to my wonderful support group of family and friends. While advisors, mentors, and committee members provide professional direction, it is the personal support I received from friends and family that undoubtedly facilitated the completion of this project. I give the utmost thanks and appreciation to Stephanie Getz and David Rayner for their unconditional love and support. I thank Stephanie for making me laugh, laughing with me, and loving me. I thank Dave for being everything I needed.

I was fortunate to be able to experience this process with my roommate and friend, Becki Battista. I am grateful for the love and support from my sister

and her family: Carey, Nate, and Tyler Reise. In addition, I am grateful for my extended network of support in Tom and Betsy Callahan, Bill and Linda Stockton, Linda Pivarnik, The Rayner Family, The Getz Family, The Nelson Family, Beth Barlow, The Purcell Family, and the members of the HERL. I also would like to thank Erik Dayrell for making the literature search for this project possible.

To my parents, Skip and Linda Perkins. There are no words to adequately convey my gratitude, for they have given me the world.

I am blessed to have so many to thank.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER 1

INTRODUCTION

Physical inactivity is inexorably linked to the burden of chronic disease. Because of this physical activity is considered the number one *leading health indicator* by the Centers for Disease Control and Prevention (CDC) in Healthy People 2010 (CDC, 2000). This ranking is based on its high relevance to public health and chronic disease risk, as physical activity behavior is a major focus of the CDC over the next ten years. However, physical activity does not come without the risk of injury, and some populations may be advised to modify the intensity of certain activities or exercise. Careful evaluation of the risks and benefits of physical activity during pregnancy is warranted.

Adaptations to pregnancy

There are obvious body weight and biomechanical changes that accompany pregnancy. Additionally, extensive endocrine, metabolic, and physiological adaptations occur throughout gestation. Changes that accompany pregnancy are largely feto-protective and promote healthy growth of the unborn baby. However, because of these physiological changes and the intimate interaction between the mother and fetus, we cannot expect the woman's physiological responses to exercise to be identical to what occurs in the nonpregnant state (Gorski, 1985). Given this intimate interaction within the maternal-fetal unit, physiological demands of exercise on the mother are translated to the fetus. Therefore, any consideration of an exercise response in the mother should also include a potential fetal effect (Gorski, 1985).

Approximately 65% of U.S. women engage in some form of regular physical activity during pregnancy (Evanson, Savitz and Huston, 2004). Estimates suggest that 15%-25% of physically active women choose to continue a regimen of sustained, moderate- to high-intensity exercise, three or more times a week throughout pregnancy (Clapp, 1996). In addition, with the number of girls participating in high-school sports exceeding 2.6 million by the end of the 20th century (National Federation of State High School Associations [NFSHSA], 2001), it is plausible that many will choose to continue their vigorous physical activity into adulthood and as they contemplate starting families. Thus, a thorough investigation of the potential effects of maternal exercise on fetal development is warranted.

Recommendations for Physical Activity During Pregnancy

In 1985, the American College of Obstetricians and Gynecologists (ACOG) published recommendations for exercise during pregnancy and the postnatal period (ACOG , 1985). Research on exercise and human pregnancy was somewhat limited at this time and the resultant guidelines were generally conservative. Without much evidence-based data available from studies involving pregnant humans, experts believed that restrictions and/or modifications to exercise were necessary. They contended that exercise programs should be based on the individual's previous activity level with the primary concern of maternal-fetal safety (ACOG, 1985). The primary recommendations included adequate caloric intake, strenuous exercise not exceeding 15 minutes in duration, a maximum heart rate of 140 beats$\cdot$min^{-1}, and avoidance of supine exercises after the fourth month. No definition of strenuous exercise was provided in the guidelines.

Over the course of a decade, the body of knowledge on the effect of exercise on the maternal-fetal unit advanced considerably, and the ACOG recommendations were updated accordingly in 1994 (ACOG, 1994). These updated recommendations advocated regular exercise up to 30 minutes per day at least three days a week, and removed limitations regarding maximum heart rate and vigorous exercise. Overall, exercise was promoted during pregnancy and postpartum for attainment of health benefits.

The most recent 2002 ACOG committee opinion on exercise during pregnancy contends that "despite the fact that pregnancy is associated with profound anatomical and physiological changes, there are few instances that should preclude otherwise healthy, pregnant women from following the same [exercise] recommendations [as their nonpregnant counterparts]" (Artal and O'Toole, 2003). ACOG's current stand on exercise during pregnancy recognizes the CDC and American College of Sports Medicine (ACSM) combined recommendations for exercise (Pate, 1995). Specifically, accumulation of 30 minutes or more of moderate exercise a day should occur on most, if not all, days of the week [in the absence of either medical or obstetric complications] (ACOG, 2002).

The CDC and ACSM are currently updating the recommendations for physical activity in adults. Due out in late 2004 or early 2005, these updates are intended to expand, clarify and reaffirm the 1995 recommendations. The previous recommendation indicating "30 minutes of moderate intensity on most days of the week" was not clear to the general public or media (Kohl, 2004). The 1995 recommendations define *moderate* activity as "activity performed at an intensity of 3 to 6 METs – the equivalent of brisk walking at 3 to 4 mph for most healthy adults" (Pate, 1995). The updated recommendations clarify "most days of the week" as "five or more days each week" (Kohl, 2004). Additionally, the same health benefits may be achieved if the individual performs *vigorous* activity (equivalent to a jog that results in a substantially increased heart rate and shortness of breath) amounting to at least 20 minutes on three or more days each week (Kohl, 2004).

According to the most current ACOG recommendations for exercise during pregnancy (2002), the 1995 CDC/ACSM recommendations for physical activity in adults are acceptable for non-complicated pregnancy despite a lack of evidence-based research on this specific level of physical activity and birth outcomes. Given that the 1995 CDC/ACSM recommendations are both acknowledged and advocated by ACOG for pregnant women, it may be assumed that the soon-to-be updated CDC/ACSM recommendations, including vigorous activity, may be an appropriate prescription during pregnancy in the future. At this time, it is useful to evaluate the exercise and pregnancy literature to determine whether maternal physical activity performed at CDC/ACSM recommended levels results in an effect on fetal growth. This can be done by performing a systematic analysis of the effect of maternal physical activity and exercise on birth weight.

Birth Weight

Birth weight is considered the single most important determinant of risk of mortality during the newborn period as well as childhood mortality and morbidity (McCormick, 1985). Consequently, much attention has been given to identifying causal determinants of birth weight, especially low birth weight (<2,500 g) (Kramer, 1987). Although there are multiple factors that affect birth weight, it is governed primarily by gestation duration and the intrauterine growth rate (Kramer, 1987). The established direct determinants of intrauterine growth rate include infant gender, racial/ethnic origin, maternal height, prepregnancy weight, paternal height and weight, maternal birth weight, parity, prior low birth weight, gestational weight gain, caloric intake, general morbidity, cigarette smoking, and alcohol consumption (Kramer, 1987). Indirect determinants include both maternal age and socioeconomic status (Kramer, 1987).

There are obvious risks associated with clinically low (<2500 g) or high (>4500 g) birth weight. However, the risks associated with normal birth weight are not as apparent. In addition to infant morbidity and mortality, birth weight may subsequently affect health in adulthood. The fetal origins theory speculates that coronary heart disease is associated with specific patterns of disproportionate fetal growth in middle to late gestation (Godfrey and Barker, 2001). In several studies, Barker and colleagues have found an inverse relationship between birth weight and adulthood blood pressure (and other CVD risk factors) that is mostly independent of gestational age (Barker, 1995; Barker and Martyn, 1992; Law and Barker, 1994). Huxley, Neil, and Collins (2002) performed a systematic review of the fetal origins literature and found inverse associations between birth weight and subsequent blood pressures reported in 52 of the 55 included studies. Despite whether this inverse relationship holds true with appropriate control for confounding (Huxley, Neil, and Collins, 2002), there is reason to believe that the effects of lower birth weight interact with postnatal growth and subsequently affect risks for hypertension and/or obesity (Eriksson and Forsén, 2002). The magnitude of the difference in normal birth weights associated with a significant affect on later life morbidities ranges from approximately 700 g (Barker et al., 1989) to 1000 g (Adair and Cole, 2003). Additionally, birth weight within the normal range was shown to be significantly and consistently positively associated with adult body mass index (Sayer et al., 2004). Studies show that light infants experience adiposity rebound at an earlier age (1-4 y), have higher BMI at age 12 years, and a higher cumulative incidence of Type 2 diabetes as an adult (Eriksson and Forsén, 2002). If the fetal origins theory is correct, individuals who were low birth weight, or at the light end of the normal birth weight range, may be predisposed to premature morbidity and mortality.

Although our understanding of determinants of fetal growth is considerable, our understanding of the effect of maternal exercise on intrauterine growth is substantially less, and current evidence is inconclusive with respect to the benefits and disadvantages of maternal exercise on birth weight. In addition to the direct and indirect determinants of fetal growth previously listed, there are theoretical fetal growth risks associated with acute maternal exercise. More specifically, hypoxia, hyperthermia, and hypoglycemia, which may all be consequences of PA, may indirectly restrict fetal growth.

Potential Risk Factors for Fetal Growth Involving Physical Activity

Hypoxia

In order to maintain normal metabolism, growth, and development, the fetus requires a continuous supply of oxygen (Koos, Power, and Longo, 1991). The exchange of oxygen across the placenta is dependent on maternal and fetal blood flow, placental hemoglobin concentrations, and arterial O_2 tensions (Koos, Power, and Longo, 1991). Maternal-fetal oxygen exchange may also be affected by temperature, hemoglobin oxygen affinity, and placental flow rate. With the exception of placental flow rate, these factors in nonpregnant women are known to be altered during the stress of exercise.

Hypoxia, a decrease in tissue oxygenation, may result from acute physical activity when normal physiologic adaptations do not occur. In pregnancy, a decreased oxygen supply to the splanchnic beds (including the placenta) may be a threat to appropriate fetal growth. However, two studies investigating exercise effects in pregnant ewes found that despite a decrease in uterine blood flow, uteroplacental oxygen delivery remained unchanged as a result of an acute increase in maternal hemoglobin concentration during the activity (Lotgering, Gilbert, and Longo, 1983a; Lotgering, Gilbert, and Longo, 1983b). Additionally the authors found an increase in maternal and fetal arterial pH values, which may be expected to occur in response to hyperventilation associated with exercise (Lotgering, Gilbert, and Longo, 1983a; Lotgering, Gilbert, and Longo, 1983b). Based on results of these and other studies, it appears that maternal exercise is accompanied by appropriate physiologic responses in order to maintain O_2 delivery to the fetus. Maternal hemoconcentration during exercise leads to increased oxygen-carrying capacity, which augments fetal oxygenation. This response offsets the reduced uterine blood flow and increased pH of maternal and fetal blood which act to depress fetal oxygenation (Koos, Power, and Longo, 1991). Overall, it appears that the net effect is little change in oxygen transfer to the fetus.

Hyperthermia

Excluding a study by Adamson et al, 1966, the majority of research involving maternal and fetal temperature regulation has primarily been conducted on animals (Lotgering, Gilbert, and Longo, 1983b; Clapp, 1980). Under normal resting conditions, fetal temperature in humans (Adamson, 1966), baboons (Morishima, Yeh, Niemann, and James, 1977), and sheep (Power, Schroder, and Gilbert, 1984) is approximately 0.5°C higher than maternal temperature.

Heat produced by the working muscles during exercise elevates body temperature. For example, 30-60 minutes of strenuous physical activity in nonpregnant women can increase core temperature to 39°C (Clapp, Wesley, and Sleamaker, 1987; Rozycki, 1984). Both exercise (Snellen, 1969) and pregnancy (Abrams, Canton, Clapp, and Barron, 1970) are associated with increased heat production. Theoretically, there may be an additive thermal effect of exercise and pregnancy that might stress normal adaptive mechanisms and exaggerate the hyperthermic response (Clapp, Wesley, and Sleamaker 1987). Because fetal heat dissipation occurs by convection and flow-limited diffusion, fetal temperature would increase in proportion to that of the mother.

Retrospective studies by Miller, Smith, and Shepard (1978) and by Pleet, Graham, and Harvey (1980) suggest that hyperthermia can be teratogenic. The reports that suggest the teratology of hyperthermia have been viewed vigilantly by researchers and practitioners because teratogenic effects of maternal temperatures in excess of 38.9°C were found. Some investigators suggested that these temperatures can be reached by an exercising woman during intense physical activity (Romen, Masaki, and Mittelmark, 1991). Milunsky et al. (1992) performed an elaborate study on non-exercising pregnant humans to examine the effects of increased core body temperature on birth outcome. The authors found an increased relative risk for intrauterine growth restriction in women who were exposed to heat in the form of hot tub, sauna, or fever (Milunsky et al., 1992). Although this study raised concern for the teratogenic effects of increased core body temperature, these findings have not been corroborated in *exercising* pregnant women.

Clapp (1991) found that at a given submaximal exercise intensity, maternal core body temperature was decreased by ~0.3°C at the eighth week of gestation, and decreased approximately 0.1°C each month thereafter. This study (Clapp, 1991) and others (Cefalo and Hellegers, 1978; Adamson, 1966) provide evidence that thermoregulatory adaptations occur during pregnancy and the normal fetal-maternal temperature gradient diminishes (Cefalo and Hellegers, 1978) or reverses (Adamsons, 1966).

Based on animal research, concern for the effects of maternal hyperthermia in humans appears warranted. Human studies are limited, and the only one showing a negative effect of temperature on fetal development was performed on women who were at rest, but exposed to high ambient temperature load (Milunsky et al, 1992). However, studies by Clapp (1991), Cefalo and Hellegers (1978), and Adamson (1966), suggest that thermal adaptations that occur during pregnancy may be feto-protective to an exercising woman.

Energy Intake and Hypoglycemia

Caloric intake and nutrient content appear to be the most important nutritional factors affecting infant growth (Butterfield and King, 1991). Appropriate caloric intake and nutrient content are required to promote an adequate supply of energy for the fetus as well as an adequate amount of weight gain by the mother. Tafari, Naeye, and Gobezie (1980) found that women who exercised through pregnancy and consumed a low calorie diet, gained significantly less weight during pregnancy than did women who exercised through pregnancy and ate a normal calorie diet. Based largely on this study, it is now recommended that pregnant women consume an additional 300 kilocalories per day to account for the extra energy demand of pregnancy (Hytten and Chamberlain, 1980).

Between 1943 and 1989 the Food and Nutrition Board prepared ten editions of the report on Recommended Dietary Allowances (RDAs) for pregnant women (Institute of Medicine [IOM], 1990). Over the course of almost four decades, energy intake recommendations during pregnancy ranged from to 2,700 kcal/day in 1953 to 2,200 kcal/day in 1968 (IOM, 1990). Since 1989, the recommendation has been an energy intake of 2,500 kcal/day (IOM, 1990). The current energy intake recommendations are 300 kcal/day higher than what is recommended for nonpregnant women (National Research Council [NRC], 1989). The additional energy requirements are greatest between 10 and 30 weeks of gestation, during the period when relatively large quantities of maternal fat are deposited (IOM, 1990). These extra energy requirements do not account for the energy cost of moving a heavier maternal body mass and it has been assumed that this would be compensated for by a reduction in physical activity (IOM, 1990). Additionally, Hytten and Chamberlain (1980) suggested that the increased needs for pregnancy could be met by reductions in physical activity.

Exercise may not only diminish energy supply to the fetus, but more specifically, may limit the amount of fuel to the fetus available from carbohydrate (stored as glycogen). While a certain level of energy intake is required for the growing fetus, it is assumed that glucose is the primary fuel source (Zaidise, Mittelmark, and Bessman, 1991) and there is a direct relationship between maternal blood glucose levels and birth weight (Clapp, 2002). Soultanakis, Wiswell, and Artal (1996) examined glucose responses to exercise in athletes and showed that pregnant women utilize circulating glucose and carbohydrate stores at a greater rate than nonpregnant women. The authors concluded that in order to decrease potential fetal "starvation" and possible subsequent growth restriction, and additional supply of carbohydrate intake is needed (Soultanakis, Wiswell, and Artal, 1996).

Physical Activity and Birth Weight

Without appropriate maternal adaptations to exercise, the resultant hypoxia, hyperthermia, and hypoglycemia, may restrict fetal growth. However, there is little evidence in the literature showing a clear relationship between exercise and lower birth weight. Although the offspring of exercising mothers are rarely low birth weight, the fetal origins theory provides further impetus to explore the determinants of birth weight within the normal range. Because exercise during pregnancy may be associated with lighter birth weight babies, maternal physical activity should be considered in future studies of the fetal origins theory. It is possible that lighter birth weight as a result of maternal physical activity may not show the same association with later life morbidities as does lighter birth weight due to other causes. With the heightened interest in maintaining or beginning an exercise program during pregnancy, and the potential relationship between maternal physical activity birth weight and later childhood morbidities, it is important to gain a more clear understanding of the effect of both *moderate* and *vigorous* physical activity during pregnancy on birth weight.

There is a growing body of literature resulting from studies designed to examine whether exercise and pregnancy can safely coexist without significantly affecting birth weight in full term infants. The results from these studies are equivocal. For example, several studies have found an inverse relationship between physical activity and birth weight, such that physically active mothers have lower birth weight infants than their sedentary counterparts (Clapp 1996b; Clapp and Capeless, 1990; Clapp, Lopez, and Harvcar-Sevcik, 1999; Collings, Curet, and Mullin, 1983; Marquez-Sterling et al., 2000). The results from these studies show a difference in birth weight between exercisers and non-exercisers that ranges between -200 to -509 grams. In general, the women in these studies were performing vigorous exercise through the end of their third trimesters. In contrast, results from several other studies showed that women who exercised during pregnancy delivered babies approximately 100 to 300 grams heavier (all infants were delivered at term and were within normal birth weight range) compared to babies of women who were not physically active during pregnancy (Clapp, Kim, Burcui and Lopez, 2000; Hatch, Shu, and McLean. 1993; Lewis, Yates, and Driskell, 1988; Magann, Evans, and Newnham, 1996). Similarly, these women participated in vigorous weight-bearing aerobic exercise throughout gestation. Some authors speculate that the time in pregnancy when exercise is begun, as well as the type of exercise may be important determinants of the effect on birth weight (Clapp, Kim, Burcui and Lopez, 2000). Lastly, there is evidence that maternal physical activity is not associated with either a lower or higher birth weight infant (Hall and Kaufman, 1987; Rabkin et al., 1990; Meyer and Daling, 1985; Wong and McKenzie, 1987).

There are several theoretical explanations for the variation in study findings in this area. Two potentially influential variables are the amount (frequency and duration) and the timing (trimester) of physical activity. For example, in a prospective cohort by Clapp and Dickstein (1984), birth outcome was examined in three healthy pregnant groups: 1) sedentary prior to and during pregnancy, 2) vigorous aerobic activity prior to, and up to the 28th week of pregnancy, and 3) vigorous aerobic activity prior to and throughout gestation. Since the two exercising groups were comparable in exercise frequency and intensity, the authors were able to evaluate the effect of trimester of physical activity on birth outcome. There was no difference in birth weight between the sedentary group and those who stopped exercise by the 28th week of gestation. However, offspring of the women who continued the same intensity and frequency throughout their pregnancies were approximately 500 grams lighter than the offspring of their sedentary counterparts. The authors suggested that multiple exercise-performance variables (such as high exercise intensity) that reflect the degree of physiological stress, as well as the timing of the exercise relative to pregnancy, play a role in the interaction between physical activity and birth weight. Magann and Evans (1996) compared birth outcome in four groups of healthy pregnant women with different levels of exercise behavior. Since these women continued their exercise regimens throughout the duration of their pregnancies, trimester of physical activity was controlled for so the effect of intensity and/or frequency of physical activity on birth weight could be examined. The authors found a positive relationship between maternal physical activity and birth weight that was slightly exaggerated by exercise intensity and/or frequency of training.

In addition to variables associated with physical activity itself, measurement error or quantification of physical activity is another potential reason for differences across studies. At present, there is no universally accepted method of measuring maternal exercise and physical activity behavior throughout gestation (Hatch and Stein, 1991). Studies vary in the methods and extent to which they measure and quantify exercise mode, intensity, frequency, and duration. Furthermore, studies that examine the same exposure and outcome may have different results due to differences in study design, study population, selection criteria employed, and potential controlled or uncontrolled confounders (maternal smoking status, maternal weight gain, etc.).

The scientific community can gain from research that pools results of similar studies in an attempt to understand an overall affect of an exposure and the subsequent outcome. Most studies of the effects of physical activity on pregnancy are not randomized control trials (RCT). If and when an RCT may be conducted on this population, compliance becomes an issue if a woman develops complications during the study period. Further, RCTs are limited because it may be unethical to allocate women to respective exercise regimens during pregnancy. Thus, reviews of literature with respect to physical activity during pregnancy are typically narrative and are limited to qualitative interpretation of observational studies. Some of the variability in qualitative reviews may be due to the subjectivity of the narrative review procedure and/or author bias. The conclusions among these review articles are as mixed as the studies they include. For example, conclusions have varied from determining that "exercise is safe during pregnancy if performed in moderation" (Gorski, 1985; Morton, Paul, and Metcalfe, 1985; Sady and Carpenter, 1989), that "the addition of exercise during pregnancy is not beneficial" (Goodlin, 1984), or that "strenuous maternal activity promotes fetal growth and decreased complications" (Woodward, 1981). The diverse and inconsistent results of these review studies may be the consequence of differing study designs, subject demographics, and exercise regimens.

Meta-Analysis

Research reviews are essential tools for health care workers and public health policy makers. A systematic review allows for a more objective appraisal of the evidence than a traditional narrative review and may help to resolve the ambiguity when original research and narrative reviews disagree (Egger, Smith, and O'Rourke, 2001). More specifically, a meta-analysis provides a quantifiable interpretation of the magnitude of the effect an exposure on selected outcome(s), and can be done when more than one study has estimated this effect. Glass

(1976) was the first to propose meta-analytical techniques, which he defined as "a rigorous alternative to the casual, narrative discussions of research studies which typify our attempts to make sense of the rapidly expanding research literature". Through meta-analysis, researchers may be able to answer questions concerning the effects of differences in studies on the overall outcome.

Two meta-analyses have been conducted previously on the effect of physical activity during pregnancy on birth weight. Kramer (2001) performed a meta-analysis for the Cochrane library on six trials considered to be of adequate rigorous design. Only three of the six studies reported any fetal outcome measures and they concluded that the available data were insufficient to exclude important risks or benefits for the mother or infant (Bell, 2002). Since that time there have not been a significant number of new randomized studies to justify an updated meta-analysis of exercise in pregnancy trials. Lokey, Tran, Wells, Myers, and Tran (1991) completed a meta-analytic review of observational studies on the effects of physical exercise on pregnancy outcomes. They concluded that an exercise program using a variety of activity modes that is performed for an average of 43 $\text{min} \cdot \text{d}^{-1}$, 3 $\text{times} \cdot \text{wk}^{-1}$, at a heart rate of up to 144 $\text{beats} \cdot \text{min}^{-1}$, does not appear to be associated with adverse effects to the mother or fetus in a healthy normal pregnancy (Lokey et al., 1991). In addition to a growth in the body of literature, there has been considerable change with respect to recommendations for physical activity during pregnancy. Increasing approval of physical activity for pregnant women has resulted in the 2002 ACOG document, advocating the same CDC/ACSM physical activity recommendations for healthy, for pregnant women (ACOG, 2002). An updated meta-analysis that encompasses recently published studies, as well as the investigation of CDC/ACSM recommended physical activity levels (moderate and vigorous intensities) on birth weight, is warranted.

Meta-analyses of observational studies are not always able to produce a single summary estimate of the effect of the exposure on the outcome. Therefore, the purpose of this meta-analysis was to estimate the extent to which CDC/ACSM recommendations for moderate and vigorous physical activity participation during pregnancy, achieved by a variety of means, influences birth weight. An additional purpose was to examine how this relationship might be influenced by various subject and study design factors. Overall, this dissertation includes five research questions and hypotheses:

Research Question 1: Does a meta-analysis of the literature show a clear relationship between maternal physical activity and lower birth weight?

Hypothesis: The summary estimate of the studies included will not indicate lower birth weight babies in women who performed physical activity as compared to their sedentary, uncomplicated pregnant, counterparts. Effect size estimates from individual studies will show significant heterogeneity for birth weight.

Statistical analysis: Chi-squared, forest plot

Research Question 2: Does a meta-analysis of the literature uniformly show an inverse relationship between *intensity* of physical activity during pregnancy and birth weight?

Hypothesis: Studies grouped by physical activity intensity (*moderate* or *vigorous*) will have significantly heterogeneous effect sizes that do not indicate an inverse relationship to birth weight.

Statistical analysis: Subgroup analysis (Chi-squared), forest plot

Research Question 3: Is trimester of physical activity an effect modifier in the relationship between moderate- to vigorous physical activity and birth weight?

Hypothesis: Women who participate in moderate- to vigorous physical activity during the third trimester will not have lower birth weight babies than those who discontinue the same level of physical activity by the first and/or second trimester.

Statistical analysis: Subgroup analysis (Chi-squared), forest plot

Research Question 4: Does maternal weight gain help explain the variance in birth weight between physically active and sedentary pregnant women?

Hypothesis: Differences in maternal weight gain between physically active and sedentary pregnant women will partially and significantly explain the differences in birth weight between infants of physically active and sedentary pregnant women.

Statistical analysis: Meta-regression

CHAPTER 2

METHODS

The influence of maternal physical activity and birth weight was investigated using meta-analytic techniques. This chapter contains a description of the procedures used for the literature search, initial screening of studies, full-text review, data abstraction, calculation of the summary statistic for each study, test and examination of heterogeneity, and calculation of the pooled estimate of the effect.

Literature Search

The goal of the literature search is to obtain all available research relevant to the area of interest. For this dissertation, a meta-analysis of observational and quasi-experimental studies, the initial literature search was performed on two computer-indexed databases: *NLM Medline* (produced by the US National Library of Medicine) and *Embase*. *NLM Medline* is a vast database that is updated weekly, covering January 1, 1966, to the present, providing references for more than 4,000 journals and over 7,000,000 articles (Jadad, Moher, and Klassen, 1998). *Embase* (Excerpta Medica Databse) is considered the European equivalent of *NLM Medline*, and provides references from an additional 1,000 journals. There is approximately 40% overlap between the *NLM Medline* and *Embase* databases (Counsell, 1997).

Although it was the intent to exclude non-English language studies, an unrestricted search (with respect to language) was conducted to document the number of relevant non-English studies that were excluded from review. The two databases, in addition to extensive hand searching of reference lists from relevant articles, allowed a maximum number of eligible studies. Although this review did not include unpublished data, it is the assumption that the included studies will be a representative sample of all available studies. There was no attempt to obtain data from unpublished studies. The reviewers felt that the peer-review process associated with manuscript submission and eventual publication was an indication of study quality. Searching terms were determined for the study population, the exposure variable, and the outcome variable. Four phrases were derived from these keywords and were use to search both *NLM Medline* and *Embase* databases. The four phrases were: 1) exercise AND birth weight, 2) exercise AND fetal growth, 3) pregnancy AND "physical activity", and 4) pregnancy AND exercise. Both databases were searched without specified limitations, covering all possible publication years. In *NLM Medline*, medical subject heading (MeSH) was used on certain keywords. MeSH is *NLM's* controlled vocabulary used for indexing articles for PubMed. MeSH terminology provides a consistent way to retrieve information that may use different terminology for the same concepts. For example, when searching for the word 'pregnancy [mh]', PubMed searches the *NLM* database for articles containing the word 'pregnancy' and all words that fall below it on a hierarchical list. Therefore, in addition to the word 'pregnancy', *Medline* was also searched for the words 'labor', 'parturition', and 'obstetric', etc. *Embase* search terms were mapped to subject heading, essentially providing depth in the search strategy. Quotations were used in the search phrase to search citations that contained the word(s) exactly as entered.

Initial Literature Screening and Study Selection

The keyword search provided a list of potential eligible articles whose titles were scanned by one reviewer. All English written, primary research articles with relevant titles were selected for abstract review. Abstract screening was done on 10% of the excluded titles to ensure validity of the title screening process. Two independent reviewers screened all identified abstracts. Reviewers were not blinded to the journal title or authors. Selection criteria, inclusion and exclusion standards were established to determine the final subset of studies to be included in the full-text review. These selection criteria fit the broad review question and were utilized to employ methodological, impartial, and reliable strategies of manuscript identification.

The broad review question of this dissertation was: Do pregnant women who are physically active during pregnancy have smaller babies than their sedentary/less active pregnant counterparts? The selection criteria for the abstract screening assessed the study population, exposure variable, study design, and outcome with respect to the objectives of this meta-analysis. Therefore, abstracts were selected if they provided evidence of a healthy, pregnant, human population, the exposure of physical activity during pregnancy, and the outcome measure of birth weight (as a continuous variable) in an observational or quasi-experimental design. Observational studies included cross-sectional, cohort, and/or case-control studies. Quasi-experimental studies were those in which the women were placed into regimented exercise programs. There was no formal attempt to exclude randomized control trials (RCTs) or experimental studies, although prior baseline knowledge of this literature suggested that there would be few if any such studies available. This meta-analysis did not include studies performed on women in third-world countries, as the authors believed that the cultural and lifestyle differences (compared to Western populations) were too influential of birth weight. In the event the two independent reviewers did not agree, the abstract was automatically selected.

Inter-observer reliability estimates for abstract selection were made using the Kappa statistic (Table 1. 2x2 Table of Observer Agreement; Equation 1.1). The Kappa statistic, as opposed to percent agreement, accounts for the degree of chance occurrence agreement between the two reviewers. While there is no absolute kappa level that indicates a valid systematic review, a low kappa may indicate that the selection criteria may be ambiguous and difficult to interpret, the selection depends too much on arbitrary judgment of reviewers, the article did not clearly describe the primary study and the selection criteria cannot be applied, or there was bias on the account of one of the reviewers (Jadad, Moher, and Klassin, 1998). The guidelines for interpreting varying levels of agreement suggested by Landis and Koch (1977) for kappa levels when applied to inter-observer reliability were used on the abstract screening results. The Landis and

Koch (1977) guidelines offer the following interpretation of kappa levels: <0 = poor, 0.00-0.20 = slight, 0.21-0.40 = fair, 0.41-0.60 = moderate, 0.61-0.80 = substantial, and 0.81-1.00 almost perfect.

Table 1. Observer Agreement for Kappa statistic

		Observer 1		
		Yes	No	
Observer 2	Yes	A	B	N_1
	No	C	D	N_2
		N_3	N_4	Total

Equation 1.1 **Kappa statistic**

$$\frac{2(AD-BC)}{N_1N_4+N_2N_3}$$

Full-Text Review

Publications selected after the initial abstract screening were subjected to a more critical evaluation including data abstraction. Primary studies selected for full-text review were given an article identification name based on the first two authors, date of publication, manuscript title, and journal title. The data abstracted were determined based on the reviewer's current understanding of the literature, guided by proposed subgroup analyses and suspected sources of heterogeneity. Initially, each full-text was scanned to verify that the study still met eligibility -- that is, the manuscript contained the appropriate population, a report of physical activity during pregnancy, and birth weight measured as a continuous variable. Once these initial criteria were satisfied, a more thorough abstraction of data followed. Data abstracted from each manuscript included information regarding the study design (i.e., case/control, prospective-, retrospective-, ambispective cohort, quasi-experimental) any potential maternal and/or infant variables that may have confounded the relationship between maternal physical activity and birth weight (i.e., gestational weight gain, socio-economic status, infant gender, race), level of physical activity during pregnancy (sedentary, moderate, or vigorous), timing of physical activity during pregnancy (trimester), and birth weight. Each reviewer determined (independently) whether physical activity level was sedentary, moderate, or vigorous. Physical activity level was considered moderate if the subjects exercised an average of 5 days per week, at least 30 minutes each exercise session, at an intensity of approximately 3-5 metabolic equivalent units (METs) above resting, according to CDC/ACSM definitions. Likewise, vigorous physical activity was defined as an average of 3 days per week, at least 20 minutes per exercise session, at an intensity of at least 6 METs. As mentioned previously, these definitions are based on the soon to be released updated CDC/ACSM recommendations for physical activity in healthy populations, and were employed in this meta-analysis in order to determine the difference in exercise level and birth weight. A shortened version of the data abstraction form, illustrating the most relevant data from each study, is located in the results section (Chapter 3, Table 2).

Individual Study Effect Estimates

Meta-analysis is a two-stage process where a summary statistic is first calculated for each study followed by determination of an overall pooled estimate of the effect. The effect size for each study represents the difference between group means relative to the amount of random variation within those groups (Equation 1.5). The articles selected for this meta-analysis reported a mean birth weight (and variance) for a control group (no physical activity during pregnancy) and at least one experimental group who performed some amount of physical activity during some defined point in pregnancy. Some studies reported comparisons between a control group and several different exercise groups. For example, if a study reported a comparison of a control group (A) and two exercise groups (B and C), this allowed for two possible comparisons to be made: A versus B, and A versus C. In this case, the data from 'A' is repeated in both comparisons. Adjustment for these multiple comparisons is necessary in order to control for the lack of independence in the data. These multiple comparisons were adjusted for by reducing the number of participants in the group that was repeated (Enanoria, Ng, Saha, and Colford, 2004). In these studies, the weighting of the control group (sedentary subjects) was adjusted by distributing the population of this group equally into both comparisons. For example, in a study that contained one control group and three exercise groups, the sample size of the control group was reduced to one-third its original size.

For each study, an effect size (Cohen's *d*) (Equation 1.1) and standard error (SE*d*) was calculated (Equation 1.2) to express the size of the effect relative to the observed variability in the study. A pooled standard deviation (*S*) between the exercise and sedentary groups was calculated (Equation 1.3). This pooled standard deviation was used as the denominator in the effect size (Cohen's *d*) calculation.

Equation 1.1 **Cohen's *d***

$$d_i = \frac{m_{1i} - m_{2i}}{s_i}$$

Where in study *i*:
S_i represents the pooled standard deviation of the two groups
(see *Equation 1.3*)
m_1 represents the mean response in the physically active group
m_2 represents the mean responses in the sedentary group

Equation 1.2 **Standard error of Cohen's *d***

$$SEd_i = \sqrt{\frac{N_i}{n_{1i}n_{2i}} + \frac{d_i^2}{2(N_i - 2)}}$$

Where in study *i*:
n_1 represents the number of participants in the physically active group
n_2 represents the number of participants in the sedentary group
$N = n_1 + n_2$, the total number of participants in study *I*
d represents the effect size

Equation 1.3 **Pooled standard deviation**

$$S_i = \sqrt{\frac{\left((n_{1i} - 1)(SD_{1i})^2\right) + \left((n_{2i} - 1)(SD_{2i})^2\right)}{N_i - 2}}$$

Where in study *i*:
n_1 represents the number of subjects in the physically active group
n_2 represents the number of subjects in the sedentary group
SD_1 and SD_2 represent the standard deviations of the variable of interest in the experimental and control group.
$N = n_1 + n_2$, the total number of participants in study *i*

Calculation of the Pooled Estimate

The overall estimate of the effect was calculated as a weighted average of the standardized difference in means from each study according to the DerSimonian-Laird random effect model (DerSimonian and Laird, 1986). The estimate of variance (τ^2) (Equation 1.4), pooled effect size (θ_{DL}) (Equation 1.6), standard error ($SE(\theta_{DL})$) (Equation 1.7), and confidence interval for the overall effect (Equation 1.8) were calculated. A test statistic for overall effect was calculated (Equation 1.9). From this equation (Equation 1.9) a p-value was generated and we were able to determine if the overall effect was significantly different than 1.0. The DerSimonian and Laird random effects model lessens the assumption that each study is estimating exactly the same underlying effect of the exposure (physical activity). It also accounts for additional between-study variation (rather than just within study variation) and results in a more conservative estimate of the effect. Each study was weighted (Equation 1.5) to reflect the amount of information it contained (i.e., sample size).

Equation 1.4 **DerSimonian and Laird estimate of variance (τ^2)**

$$\tau^2 = \frac{Q-(k-1)}{\sum w_i - \left(\frac{\sum w_i^2}{\sum w_i}\right)}$$

Where:
Q represents the heterogeneity statistic (see *Equation 2.1*)
k represents the number of studies included in the pooled estimate
w' represents the study weight *i* (see *Equation 1.5*)

Equation 1.5 **Inverse variance**

$$w'_i = \frac{1}{SE(\theta_i)^2 + \tau^2}$$

Where:
SE(θ) represents the standard error the effect size (Cohen's *d*)

Equation 1.6 **DerSimonian and Laird pooled effect size**

$$\theta_{DL} = \frac{\sum w'_i \theta_i}{\sum w'_i}$$

Where:
θ_i represents the effect size estimate from study *i*
w_i represents the weighting of the estimate from study *i*

Equation 1.7 **Standard error of the pooled estimate**

$$SE\theta_{DL} = \frac{1}{\sqrt{\sum w'_i}}$$

Equation 1.8 **Confidence interval (95%) for the overall effect**

$$\theta-\left(1.96\times SE(\theta)\right) \text{ to } \theta+\left(1.96\times SE(\theta)\right)$$

Equation 1.9 **Test statistic for overall significance**

$$z=\frac{\theta}{SE(\theta)}$$

Test for Heterogeneity

An important component of a meta-analysis is the assessment of effect estimates (Cohen's *d*) across all studies. This is achieved using the test for heterogeneity, given by *Q* (Equation 2.1). Heterogeneity exists when effect size estimates vary between studies to a greater extent that expected based on chance alone (Sutton, 2001), and can be determined by testing whether the variance (τ^2) between studies is equal to zero. *A priori* statistical significance of the heterogeneity test was set at *P*-value < 0.05. Nonsignificant heterogeneity was not considered proof of statistical homogeneity across studies. In addition to a statistical test for heterogeneity, a graphical display (forest plot) of the results from individual studies was used for visual examination of the degree of heterogeneity between studies.

Equation 2.1 **Heterogeneity statistic**

$$Q = \sum w_i(\theta_i - \theta_{DL})^2$$

Where:
θ_{DL} represents the pooled summary estimate from the full meta-analysis

Stratified Analyses and Meta-Regression

In attempt to explain potential heterogeneity among birth weight we performed subgroup analyses according to study-level variables and participant characteristics known or suspected to affect birth weight. For the subgroup analyses, a separate meta-analysis was carried on the studies within the subgroup. These subgroups were pre-specified according to features that potentially moderate the overall effect. Figure A. illustrates the pre-specified subgroups.

Study-level characteristics included exercise intensity (moderate/vigorous), timing of physical activity (trimester), physical activity measurement methods (training sessions, prospective questionnaire, retrospective questionnaire), and study design (observational or quasi-experimental). Participant characteristics suspected or known to affect infant birth weight were pre-specified for subgroup analyses. These variables included maternal weight gain, maternal age, smoking status, and race. All subgroups were derived from the full meta-analysis. Further stratification beyond the initial subgroup analyses was dependent upon the number of studies attained in the full meta-analysis.

In order to determine whether the subgroups had differential affects on the outcome, a formal test of statistical significance was performed (Equation 2.2). In this equation, the overall unstratified analysis is Q_T, and the heterogeneity explained by differences between subgroups is Q_B. The result (Q_B) was compared with critical values of the chi-squared distribution with k-1 degrees of freedom, where k is the number of subgroups.

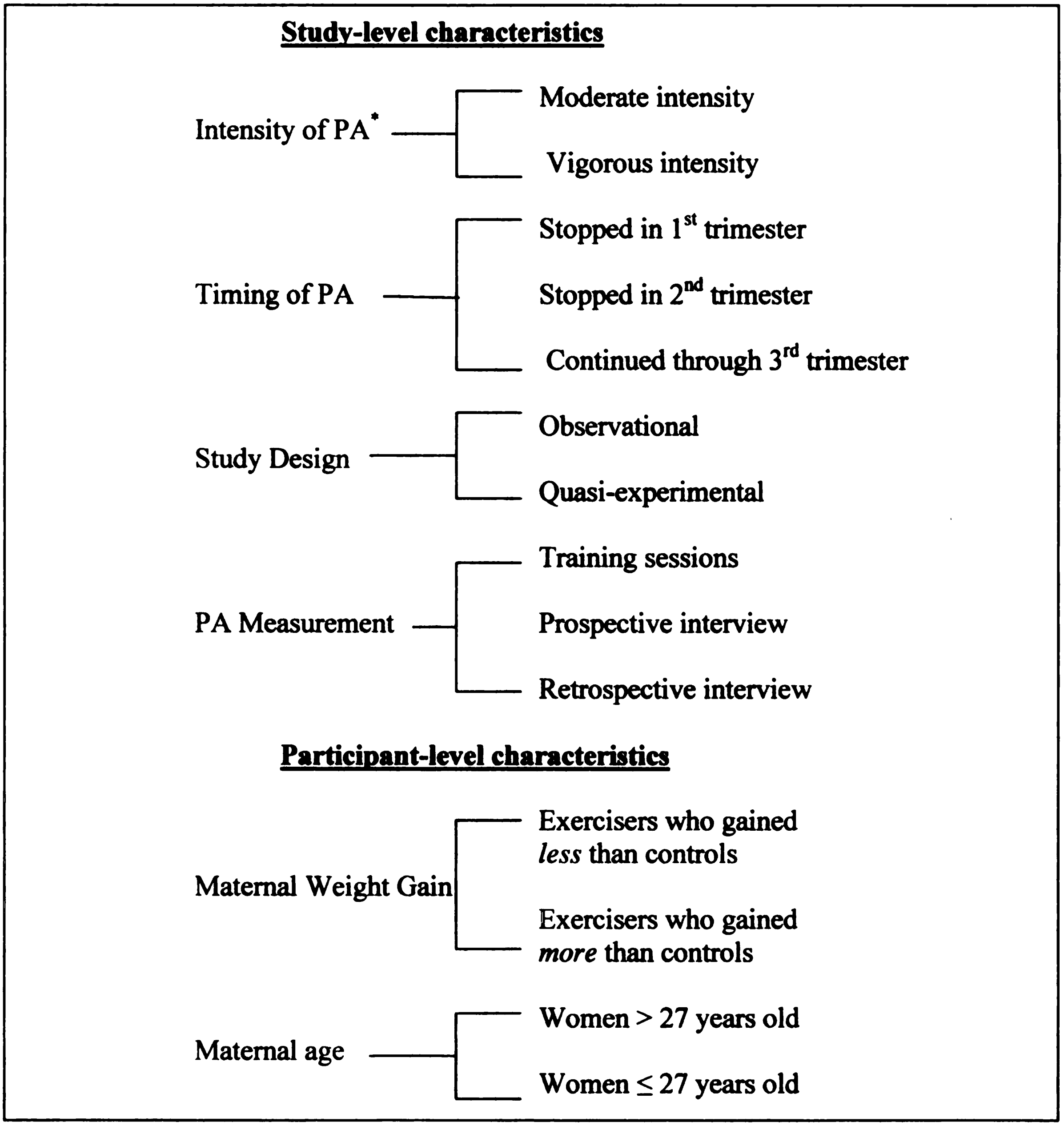
Equation 2.2 **Significance of the subgroup analysis**

$$Q_B = Q_T - \sum_k Q_k$$

Where:
Q_T represents the overall unstratified analysis
k represents the number of subgroups

It is possible that studies where maternal weight gain was different may have produced differences on the effect of maternal physical activity and birth weight (Perkins et al., 2004). Therefore, meta-regression was used to analyze the association between maternal weight gain and birth weight. From this model we were able to determine the estimated amount of increase in birth weight per unit increase in the covariate (maternal weight gain). A regression coefficient, 95% CI, and p-value was determined for maternal weight gain.

Figure 1. Diagram of Proposed Subgroup Analyses



*PA = physical activity

CHAPTER 3

RESULTS

The search strategy yielded 2418 potentially relevant citations (Figure 2). Of these 2418 studies, 192 abstracts were reviewed and 52 were selected for full-text review. The Kappa agreement between the two independent reviewers for abstract screening was 0.94. A final 18 articles satisfied the inclusion criteria for this meta-analysis. However, because of multiple physical activity groups in three of the studies (Clapp 1984, Magann 1996, Magann 2002), 23 comparisons were made between a sedentary control group and a physically active group. In this dissertation, *n* refers to the total number of studies and *k* refers to the number of comparisons. For example, a subgroup analysis may contain 10 studies (n=10), but because of multiple exercise groups in 2 studies there were 13 possible comparisons to be made between exercising and nonexercising groups (k=13). Appendix A includes a complete list of studies included in this meta-analysis, organized by physical activity level. Appendix B includes the 34 studies excluded after full-text review.

Each reviewer independently classified the described physical activity as *moderate* (3-5 METs; $\geq$30 min; $\geq$5 $\mathrm{d \cdot wk^{-1}}$) or *vigorous* ($\geq$6 METs; $\geq$20 min; $\geq$3 $\mathrm{d \cdot wk^{-1}}$) according to CDC/ACSM recommendations. However, some groups did not meet CDC/ACSM recommendations for moderate physical activity. These studies met moderate recommendations with respect to intensity and duration, but demonstrated low frequency of physical activity performance (< 5 $\mathrm{d \cdot wk^{-1}}$).

The 23 studies were grouped into the following levels of physical activity; 2 studies met CDC/ACSM moderate physical activity recommendations, 8 did not meet CDC/ACSM moderate physical activity recommendations, and 13 studies met CDC/ACSM recommendations for vigorous physical activity. For statistical analysis, the studies that met CDC/ACSM moderate criteria (Magann, Evans, and Newnham, 1996; Magann, Evans, Weitz, and Newnham, 2002) were combined with the 8 studies that did not meet CDC/ACSM moderate criteria. The main characteristics of each study, including physical activity level, number of participants, mean age, birth weight, gestational weight gain, study design, and physical activity measurement method are presented in Table 2.

Research Question 1: Does a meta-analysis of the literature show a clear relationship between maternal physical activity and lower birth weight?
Hypothesis: The summary estimate of the studies included will not indicate lower birth weight babies in women who performed physical activity as compared to their sedentary, uncomplicated pregnant, counterparts. Effect size estimates from individual studies will show significant heterogeneity for birth weight.

Individual effect sizes (Cohen's *d*) for all studies ranged from -1.51(Clapp and Dickstein, 1984) to 0.52 (Lewis, Yates, and Driskell, 1988) and the difference in mean birth weight varied from -509.9 to 300.0 grams. A pooled analysis of all 23 comparisons, adjusting for repeated use of some control groups, showed significant heterogeneity (X^2=109.11; df=22; P<0.01) with an overall effect size and 95% confidence interval (95% CI) of –0.12 (-0.30 to 0.06) using the

DerSimonian-Laird (D-L) random effects model. In the forest plot (Figure 3) of this analysis, and all subsequent forest plots (Figures 5 through 10), the *location* of the black square with respect to the x-axis indicates the effect sizes from each study, while the *size* of the black square indicates the sample size. The vertical line represents an effect size of zero, or no effect. Squares located to the left of the vertical line represent studies where the physically active women had lighter babies than their sedentary counterparts. The squares located to the right of the vertical line represent studies where the physically active women had heavier babies than their sedentary counterparts. As a result of this analysis we accept the hypothesis to research question 1; although the overall effect size was negative, the studies indicated significant heterogeneity.

Inspection of the forest plot shown in Figure 3 revealed that two of the 18 studies may be outliers. A *post hoc* sensitivity analysis was performed by repeating the meta-analysis on subsets of the original data set in order to determine the influence of each study on the overall result. In a sensitivity analysis, the influence of a given study is estimated by deleting it from the analysis and noting the degree to which the size and significance of the effect changes (Deeks, Altman, and Bradburn, 2001). This analysis allowed examination of the stability of effect estimates across the differing study designs, approaches to exposure ascertainment, and selection of study participants. It is evident from Table 4 and Figure 4 (graphical display of the sensitivity analysis) that the effect size estimates determined with the exclusion of any individual study are virtually identical to the combined effect. Because the sensitivity analysis was unable to isolate the two potential outliers simultaneously, a separate meta-analysis was employed on the remaining 21 studies when the extremely vigorously active group in Clapp and Dickstein (1984) and Clapp et al (1990) were removed. With these studies removed, the pooled effect size was close to zero (ES=0.03; X^2=32.79) and approached nonsignificant heterogeneity (*P=0.04*). The results from the sensitivity analyses prevent misleading conclusions from the exploration of heterogeneity via subgroup analysis.

Research Question 2: Does a meta-analysis of the literature uniformly show an inverse relationship between intensity of physical activity during pregnancy and birth weight?

Hypothesis: Studies grouped by physical activity intensity (*moderate* or *vigorous*) will have significantly heterogeneous effect sizes that do not indicate an inverse relationship to birth weight.

When studies were stratified by physical activity level (moderate and vigorous), moderate exercise studies were nonsignificantly heterogeneous (ES=0.05; X^2=8.41; *P*=0.49), while significant heterogeneity remained among the vigorous studies (ES=$^{-}$0.27; X^2=98.72; *P*<0.01) (Table 3). The forest plots for moderate and vigorous physical activity are shown in Figure 5. Figure 5 illustrates the forest plot of all studies, subgrouped by moderate or vigorous physical activity intensity. In this figure, the moderate intensity studies are shown by white squares and the vigorous intensity studies are shown by black squares. The pooled effect sizes for each subgroup, moderate and vigorous, are shown by the white and black diamonds, respectively. Although there was an inverse relationship present among the vigorously active group, these studies were significantly heterogeneous. We accept the hypothesis to research question 2 with respect to vigorous physical activity. However, there was no effect of moderate physical activity on birth weight and studies in this group had homogeneous effect sizes. We do not accept the hypothesis to research question 2 with respect to moderate physical activity.

The heterogeneity among the 2 subgroups (moderate X^2=8.41 and vigorous X^2=8.41) explained a majority of the heterogeneity in the unstratified analysis (X^2=109.11). Based on a significance value of *P*=0.05 and one degree of freedom, the heterogeneity remaining after the studies were stratified into groups based on physical activity intensity was not significant (Q_B=1.98).

Research Question 3: Is trimester of physical activity an effect modifier in the relationship between moderate- to vigorous physical activity and birth weight?
Hypothesis: Women who participate in moderate- to vigorous physical activity during the third trimester will not have lower birth weight babies than those who discontinue the same level of physical activity by the first and/or second trimester.

The final group of studies selected for this meta-analysis did not differ with respect to trimester of physical activity. Twenty-one of the 23 studies included in this meta-analysis contained physically active participants who continued physical activity through the third trimester. Thus, we were unable to draw any conclusions regarding trimester of physical activity and birth weight.

Research Question 4: Does maternal weight gain help explain the variance in birth weight between physically active and sedentary pregnant women?
Hypothesis: Difference in maternal weight gain between physically active and sedentary pregnant women will partially and significantly explain the differences in birth weight between infants of physically active and sedentary pregnant women.

Each study was classified into one of two maternal weight gain groups, exercisers who gained less (-MWG; k=10) or exercisers who gained more than their sedentary controls (+MWG; k=9). When analyzed, the -MWG studies were significantly heterogeneous (ES=¯0.30; $P<0.01$), while the +MWG studies were not significantly heterogeneous (ES=0.02; $P=0.06$) (Table 3). The forest plot of these analyses is shown Figure 6. In Figure 6, the –MWG studies are shown by white squares, while the +MWG studies are shown by black squares. The pooled effect sizes for each subgroup, -MWG and +MWG, are shown by white and black diamonds, respectively.

Meta-regression of maternal weight gain on birth weight was significant (P<0.01) among the studies where weight gain data were available (k=19). The resultant coefficient (95% CI) was 0.17 (0.09-0.25), indicating that the difference in weight gain can account for approximately 17% of the variance in the difference in birth weight between the offspring of physically active and sedentary women. The hypothesis to research question 4 was accepted; differences in maternal weight gain explain a significant amount of differences in birth weight between physically active and sedentary pregnant women.

Additional Stratified Analyses

Studies were further stratified by several group- or study-level characteristics including maternal age, year of publication, method of physical activity measurement, and study design. Studies (k=20) that provided information regarding the subjects' mean age were dichotomized into two groups, women ≤27 years (AGE$_{\leq 27}$), and women >27 years of age (AGE$_{>27}$). This dichotomization of age at 27 years was selected after examination of the mean ages of women in these studies. More specifically, no one study had women that were both less than or equal to 27 years of age *and* women who were greater than 27 years of age. Studies that contained women less than or equal to 27 years of age (k=7) were nonsignificantly heterogeneous and showed no effect of physical activity on birth weight (ES=0.10; X^2=12.00; *P*=0.06) (Table 3; Figure 7). Studies that contained women greater than 27 years of age (k=13) showed a negative effect size and significant heterogeneity (ES=$^{-}$0.21; X^2=48.73; *P*<0.01) (Table 3; Figure 7).

The final 23 studies that were included in this meta-analysis measured physical activity one of three ways. Therefore, all studies were classified into one of three physical activity measurement categories: training session/researcher monitored (PAmeas-1), prospective interview/questionnaire throughout gestation

(PAmeas-2), retrospective interview/questionnaire (PAmeas-3). Dersimonian-Laird random effects analysis of these subgroups showed significant heterogeneity in PAmeas-1 (ES=$^{-}$0.28; X^2=29.82; *P*<0.01) and PAmeas-2 (ES=$^{-}$0.30; X^2=39.08; *P*<0.01). There was nonsignificant heterogeneity in the studies that obtained physical activity data via retrospective questionnaire (ES=0.22; X^2=2.94; *P*=0.82). The forest plot of PAmeas-1, PAmeas-2, and PAmeas-3 is shown in Figures 8. In this figure, the PAmeas-1 is shown by black squares, PAmeas-2 is shown by white squares, and PAmeas-3 is shown by the striped squares. The pooled effect sizes for each of these subgroups are represented by diamonds, using the same color scheme. The size of the diamond incorporates the 95% confidence interval of the pooled effect size. The cumulative heterogeneity (X^2=71.84) of this subgroup did not explain a significant amount of the heterogeneity in the full meta-analysis (X^2=109.11), based on *P*=0.05 and 2 degrees of freedom.

All studies were stratified by publication year (those published prior to 1995 and those published after 1995). This cutoff was chosen as it was one year after the second, less restrictive 1994 ACOG Guidelines for physical activity and pregnancy were published. The studies published prior to 1995 (ES=$^{-}$0.25; X^2=64.55; *P*<0.01) and those published after 1995 (ES=$^{-}$0.01; X^2=28.69; *P*<0.01) were significantly heterogeneous (Table 3). The forest plot for these subgroups is shown in Figure 9. In this figure, studies published prior to and after 1995 are represented by the white and black squares, respectively. The heterogeneity accounted for by publication year (X^2=98.24) did not account for a significant amount of the heterogeneity in the full meta-analysis (X^2=109.11).

Studies were stratified by study design (observational (Observ) and quasi-experimental (Q-Exp)). Observational studies (k=9) were significantly heterogeneous (ES=$^{-}$0.02; X^2=59.59; P<0.01) (Table 3). Quasi-experimental studies (k=9) were significantly heterogeneous (ES=$^{-}$0.32; X^2=31.34; P<0.01) (Table 3). The forest plot of quasi-experimental studies and observational studies is shown in Figure 10, where they are represented by white and black squares, respectively. The heterogeneity accounted for by study design (X^2=90.93) did not account for a significant amount of the heterogeneity in the full meta-analysis (X^2=109.11).

Figure 2. Flow diagram of reviewed literature

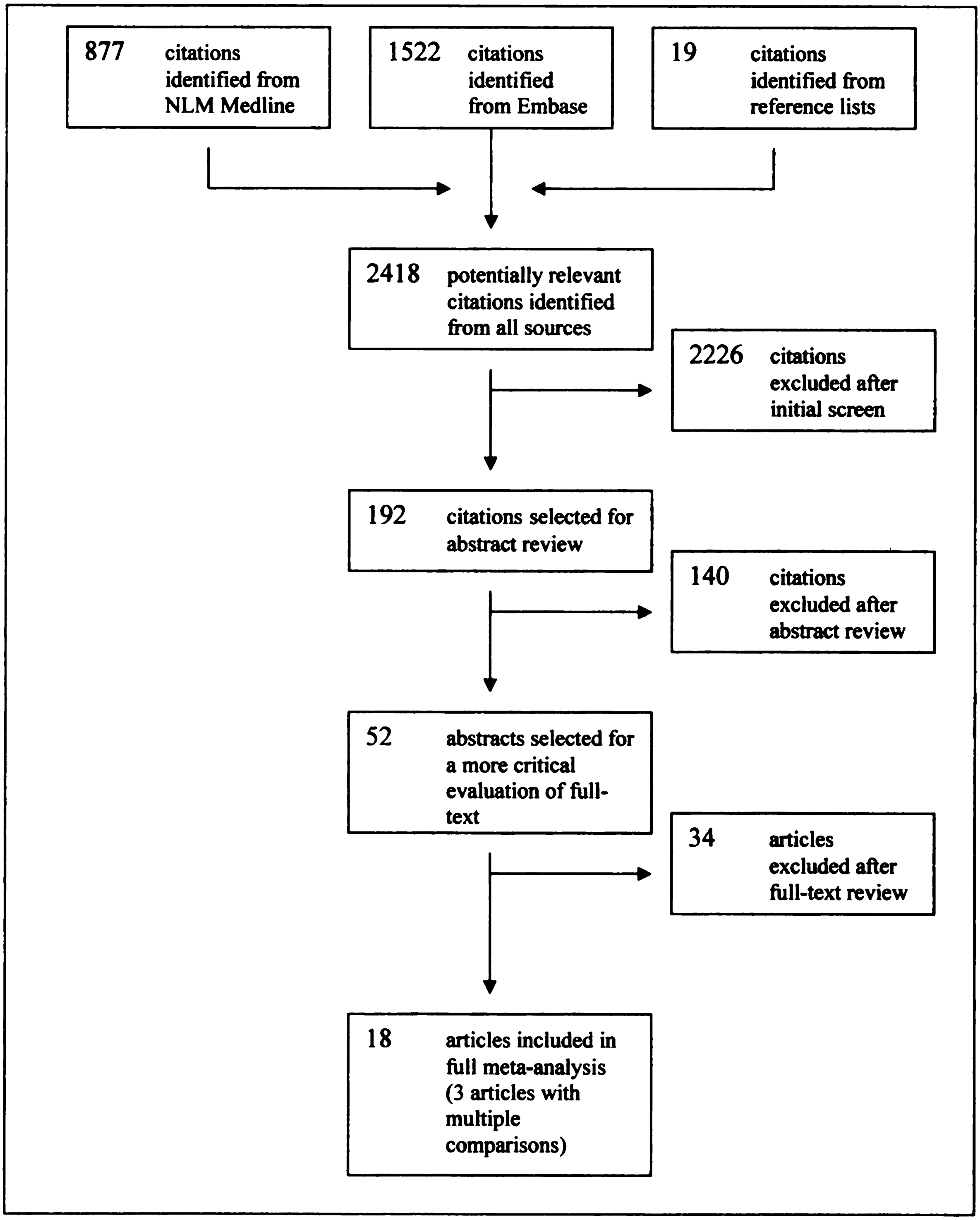


Table 2. Characteristics of included studies (n=18 studies)

Reference	PA level	Total number of participants	Age (y) (mean ± SD)	Birth Weight (g) (mean ± SD)	Maternal weight gain (kg) (mean ± SD)	PA measurement
Beckman, 1990^{A}	Sed	50	30.2 ± 3.4	3649.0 ± 340.0	13.3. ± 2.7	retrospective questionnaire
	MnR	50	28.9 ± 1.8	3689.0 ± 298.0	11.8 ± 1.6	
Bell, 1995^{B}	Sed	41	31.6 ± 4.7	3364.0 ± 412.0	na	prospective questionnaire
	MnR	58	31.8 ± 2.7	3353.0 ± 589.0	na	
Botkin, 1991^{A}	Sed	25	27.2 ± 5.5	3523.0 ± 351.0	14.5 ± 3.5	retrospective questionnaire
	MnR	19	28.1 ± 5.1	3663.8 ± 318.4	14.4 ± 4.1	
Clapp, 1984^{A}	Sed	152 (76)*	27.2 ± 0.4	3518.0 ± 528.4	14.6 ± 0.4	prospective questionnaire
	VR	47	28.2 ± 0.6	3577.0 ± 624.0	16.8 ± 0.8	
	VR	29	28.2 ± 0.6	3009.0 ± 529.2	12.2 ± 0.6	
Clapp, 1990^{B}	Sed	44	na	3776.0 ± 401.0	17.2 ± na	PA session monitored
	VR	87	na	3369.0 ± 318.0	13.6 ± na	
Clapp, 1996^{B}	Sed	19	31.0 ± 1.0	3640.0 ± 297.0	16.2 ± 1.4	prospective questionnaire
	VR	19	31.0 ± 1.0	3400.0 ± 339.0	12.8 ± 1.1	
Clapp, 1998^{B}	Sed	15	31.0 ± 1.0	3580.0 ± 700.0	17.5 ± 1.7	prospective questionnaire
	VR	52	31.0 ± 1.0	3380.0 ± 600.0	13.6 ± 1.4	
Clapp, 1999^{B}	Sed	31	32.0 ± 1.0	3640.0 ± 500.0	17.9 ± 1.6	prospective questionnaire
	VR	34	32.0 ± 1.0	3440.0 ± 600.0	12.8 ± 1.3	
Clapp, 2000^{B}	Sed	24	31.0 ± 1.0	3490.0 ± 700.0	16.3 ± 0.7	PA session monitored
	VR	22	31.0 ± 1.0	3750.0 ± 800.0	15.7 ± 1.0	
Collings, 1983^{B}	Sed	12	28.0 ± 3.7	3596.3 ± 479.8	14.0 ± 3.7	PA session monitored
	VR	8	26.9 ± 2.8	3353.0 ± 415.0	15.8 ± 3.6	
Horns, 1996^{A}	Sed	53	27.2 ± 3.8	3467.0 ± 434.0	17.4 ± 5.8	retrospective questionnaire
	MnR	48	28.4 ± 4.1	3496.0 ± 486.0	16.3 ± 5.3	

continued on next page

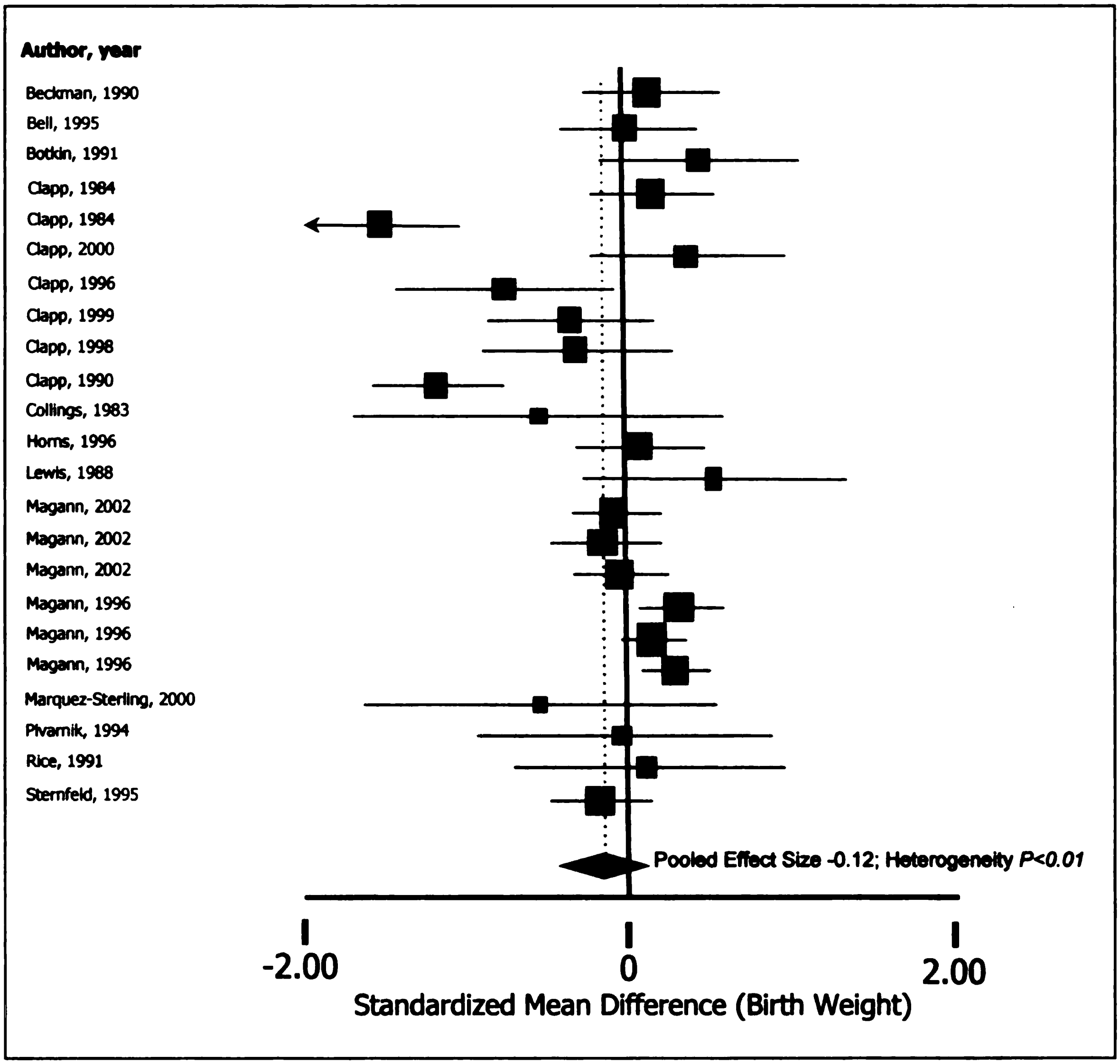
continued from previous page (Table 2.)

Reference	PA level	Total number of participants	Age (y) (mean ± SD)	Birth Weight (g) (mean ± SD)	Maternal weight gain (kg) (mean ± SD)	PA measurement
Lewis, 1988[B]	Sed	10	na	3400.0 ± 700.0	na	prospective questionnaire
	MnR	18	na	3700.0 ± 500.0	na	
Magann, 1996[A]	Sed	359 (119)*	23.5 ± 6.0	3214.0 ± 546.0	9.0 ± 3.7	retrospective questionnaire
	MR	1059	29.0 ± 5.6	3300.0 ± 608.0	8.3 ± 3.7	
	VR	608	27.0 ± 5.8	3360.0 ± 510.0	9.0 ± 3.6	
	VR	172	27.2 ± 5.8	3397.0 ± 611.0	9.3 ± 3.7	
Magann, 2002[A]	Sed	217 (72)*	23.7 ± 4.9	3423.0 ± 617.0	14.7 ± 6.2	PA session monitored
	MR	222	24.1 ± 5.0	3362.0 ± 717.0	15.6 ± 7.5	
	VR	73	25.4 ± 4.7	3323.0 ± 662.0	15.6 ± 6.5	
	VR	238	23.7 ± 4.9	3390.0 ± 638.0	16.1 ± 7.9	
Marquez-Sterling, 2000[B]	Sed	6	27.8 ± 3.1	3722.0 ± 504.6	15.7 ± 4.0	PA session monitored
	MnR	9	31.1 ± 3.1	3515.4 ± 274.9	16.2 ± 3.4	
Pivarnik, 1994[A]	Sed	5	29.0 ± 5.0	3564.0 ± 435.0	na	prospective questionnaire
	VR	9	29.0 ± 4.0	3543.0 ± 709.0	na	
Rice, 1991[A]	Sed	11	26.2 ± 5.1	3450.0 ± 449.0	13.3 ± 5.2	retrospective questionnaire
	MnR	12	23.3 ± 3.6	3496.0 ± 318.0	13.9 ± 3.6	
Sternfeld, 1995[A]	Sed	242	na	3574.0 ± 504.8	na	prospective questionnaire
	MnR	53	na	3488.0 ± 465.7	na	

*adjusted group size for multiple comparisons; [A]observational study; [B]quasi-experimental study; na = data not available

Note: participant groups across all studies were homogeneous with respect to race and smoking status.

Figure 3. The Effect of Maternal Physical Activity on Birth Weight



Note: The vertical line represents no difference in the birth weight of babies born to physically active and sedentary women. The diamond (ES=-0.27; *P*<0.01) indicates the summary estimate of comparisons (N=23) included in the full meta-analysis.

Table 3. Effect of exercise on birth weight by subgroups of comparisons (Random Effects)

Comparison	**no. of studies (k)**	**SMD (95%CI)**	**Test for heterogeneity**	**Difference in BW (g)**	**Q_B**
PA v. Sed	18 (23)	-0.12 (-0.30 – 0.06)	X^2=109.11; df=22; p<0.01	~ -80 g	
Subgroups					
VR	10 (13)	-0.27 (-0.59 – 0.04)	X^2=98.72; df=12; p<0.01	~ -150 g	
MR	10 (10)	0.05 (-0.07 – 0.16)	X^2=8.41; df=9; p=0.49	~ 25 g	
					PA-I 1.98
MWG+	7 (8)	0.02 (-0.15 – 0.18)	X^2=10.46; df=7; p=0.16	~ 20 g	
MWG-	10 (10)	-0.30 (-0.70 – 0.10)	X^2=80.34; df=9; p<0.01	~ -200 g	
AGE$_{\le 27}$	3 (7)	0.10 (-0.04 – 0.24)	X^2=12.00; df=6; p=0.06	~ 40 g	
AGE$_{>27}$	12 (13)	-0.21 (-0.51 – 0.10)	X^2=48.73; df=12; p<0.01	~ -140 g	
YEAR$_{<1995}$	7 (8)	-0.25 (-0.82 – 0.32)	X^2=64.55; df=7; p<0.01	~ -150 g	
YEAR$_{>1995}$	10 (13)	-0.01 (-0.15 – 0.13)	X^2=28.69; df=13; p<0.01	~ -10 g	
					YEAR 15.87¥
PAmeas-1	5 (7)	-0.28 (-0.64 – 0.08)	X^2=29.82; df=6; p<0.01	~ -150 g	
PAmeas-2	8 (9)	-0.30 (-0.66 – -0.07)	X^2=39.08; df=8; p<0.01	~ -200 g	
PAmeas-3	5 (7)	0.22 (-0.11 – 0.32)	X^2=2.94; df=6; p=0.82	~ 120 g	
					PAmeas 37.27¥
Observ	9 (14)	-0.02 (-0.21 – 0.17)	X^2=59.59; df=14; p<0.01	~ -15 g	
Q-Exp	9 (9)	-0.32 (-0.77 – 0.07)	X^2=31.34; df=8; p<0.01	~ -220 g	
					ST-DES 18.18¥

PA-I=physical activity intensity subgrouping; PA=physically active; Sed=sedentary; MR=moderate intensity physical activity; VR=vigorous intensity physical activity; Q_B=heterogeneity difference between subgroups; PAmeas=type of physical activity measurement; Observ=observational design; Q-Exp=quasi-experimental design; ST-DES=study design.

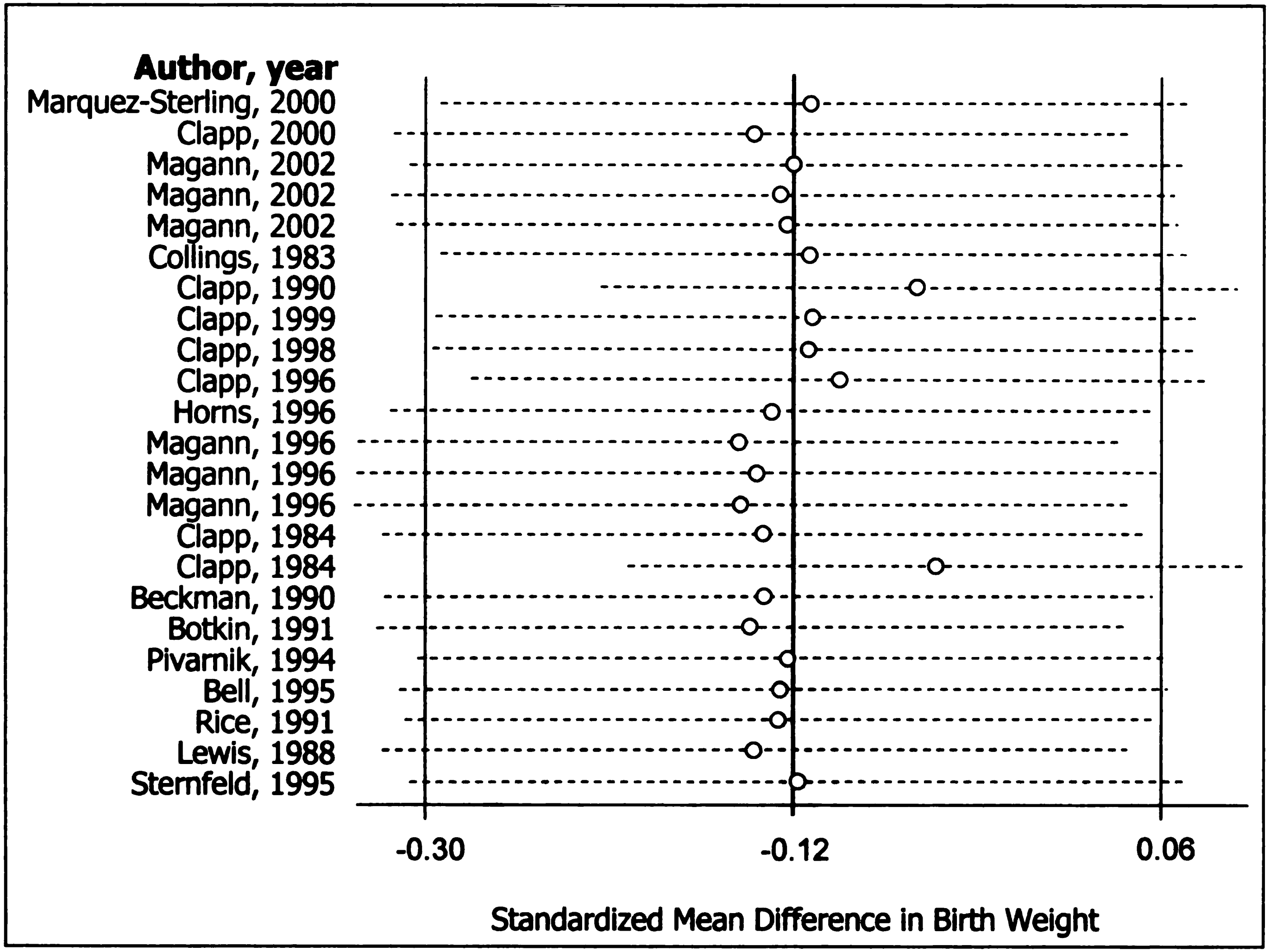
¥ P>0.05

Table 4. Sensitivity Analysis - standardized mean difference in birth weight (and 95% CI) computed by omitting one study in each turn.

Study omitted	Coefficient (95% CI)
Marquez-Sterling, 2000	-0.112 (-0.298 - 0.073)
Clapp, 2000	-0.140 (-0.328 - 0.047)
Magann, 2002	-0.120 (-0.313 - 0.072)
Magann, 2002	-0.127 (-0.322 - 0.069)
Magann, 2002	-0.125 (-0.320 - 0.071)
Collings, 1983	-0.113 (-0.298 - 0.072)
Clapp, 1990	-0.059 (-0.218 - 0.101)
Clapp, 1999	-0.111 (-0.299 - 0.077)
Clapp, 1998	-0.113 (-0.301 - 0.074)
Clapp, 1996	-0.098 (-0.282 - 0.086)
Horns, 1996	-0.131 (-0.323 - 0.060)
Magann, 1996	-0.148 (-0.338 - 0.042)
Magann, 1996	-0.139 (-0.338 - 0.059)
Magann, 1996	-0.147 (-0.339 - 0.045)
Clapp, 1984	-0.135 (-0.033 - 0.056)
Clapp, 1984	-0.050 (-0.204 - 0.105)
Beckman, 1990	-0.135 (-0.325 - 0.056)
Botkin, 1991	-0.142 (-0.329 - 0.044)
Pivarnik, 1994	-0.124 (-0.310 - 0.063)
Bell 1995	-0.127 (-0.318 - 0.064)
Rice, 1991	-0.128 (-0.315 - 0.058)
Lewis, 1988	-0.141 (-0.326 - 0.045)
Sternfeld, 1995	-0.119 (-0.312 - 0.074)
Combined	-0.121 (-0.304 - 0.062)

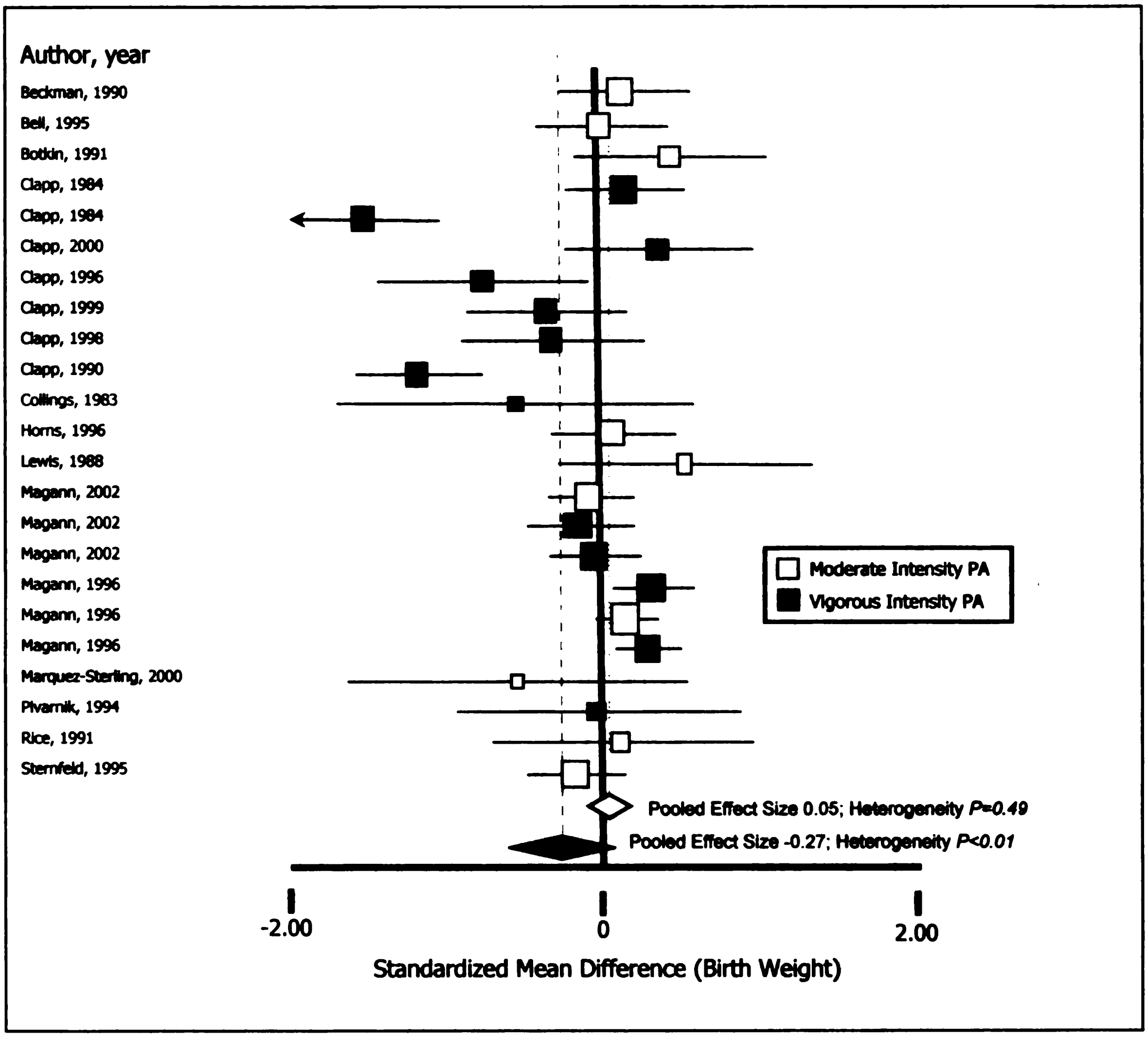
*all analyses were significantly heterogeneous

Figure 4. Sensitivity Analysis: Meta-analysis random-effects estimates with single study omitted



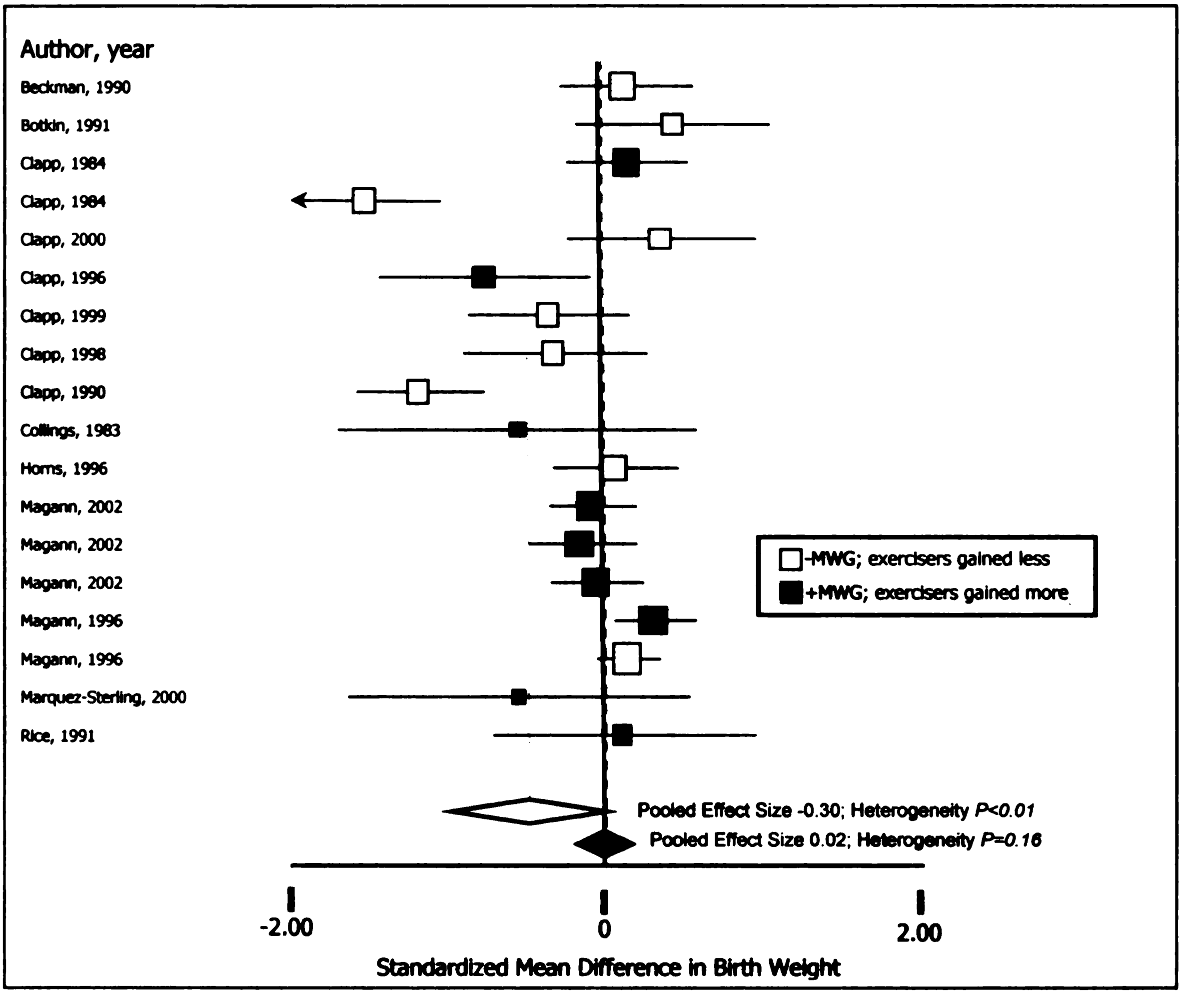
Note: Open circles (dashed lines = 95% CI) represent the pooled effect size of the meta-analysis excluding the study listed at the right. The center vertical line represents the pooled effect size of the original meta-analysis when all studies were included.

Figure 5. The Effect of Maternal Physical Activity on Birth Weight – Stratified by Physical Activity Intensity



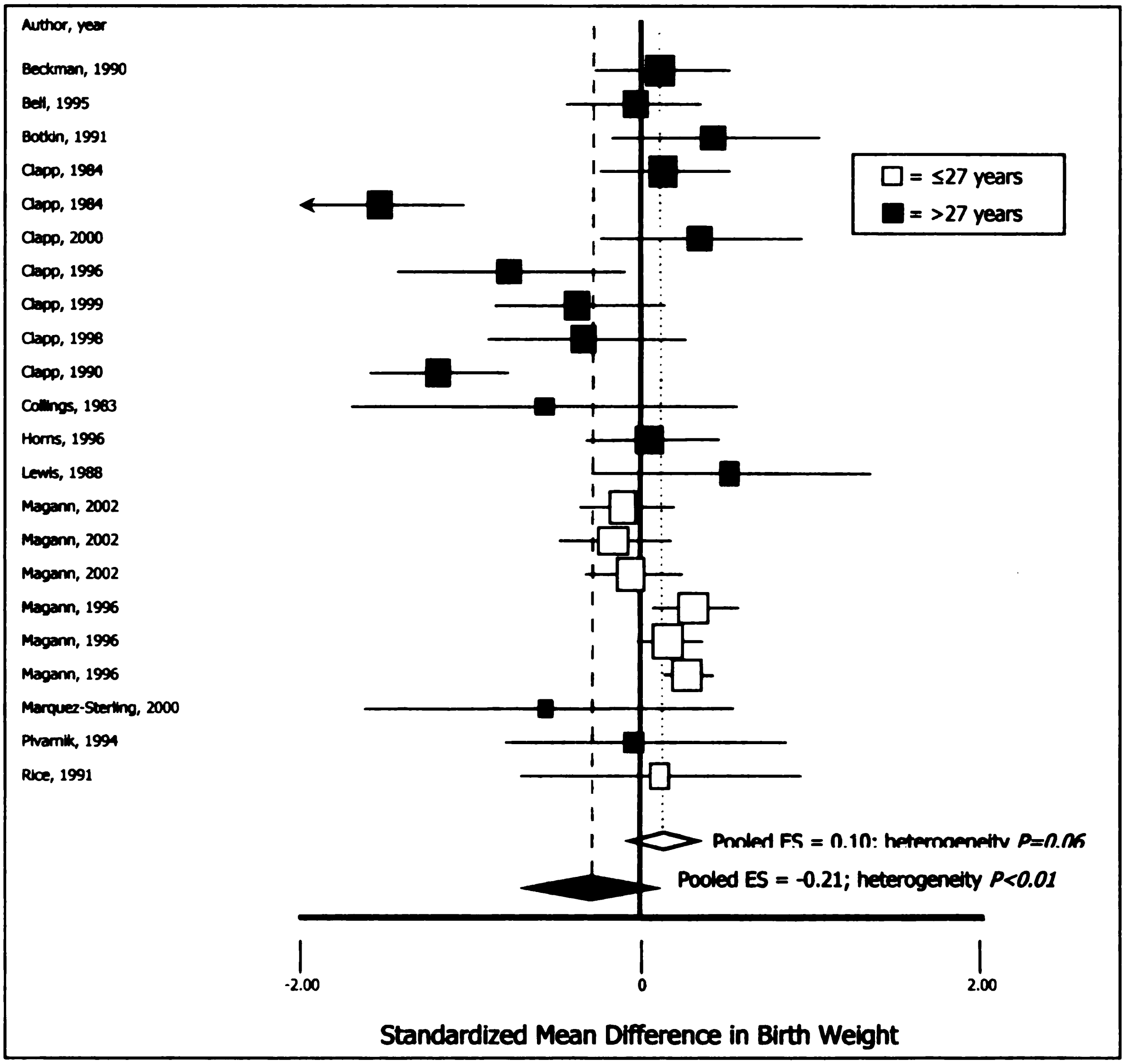
Note: The vertical line represents no difference in the birth weight of babies born to physically active and sedentary women.

Figure 6. The Effect of Maternal Physical Activity on Birth Weight – Stratified by Maternal Weight Gain.



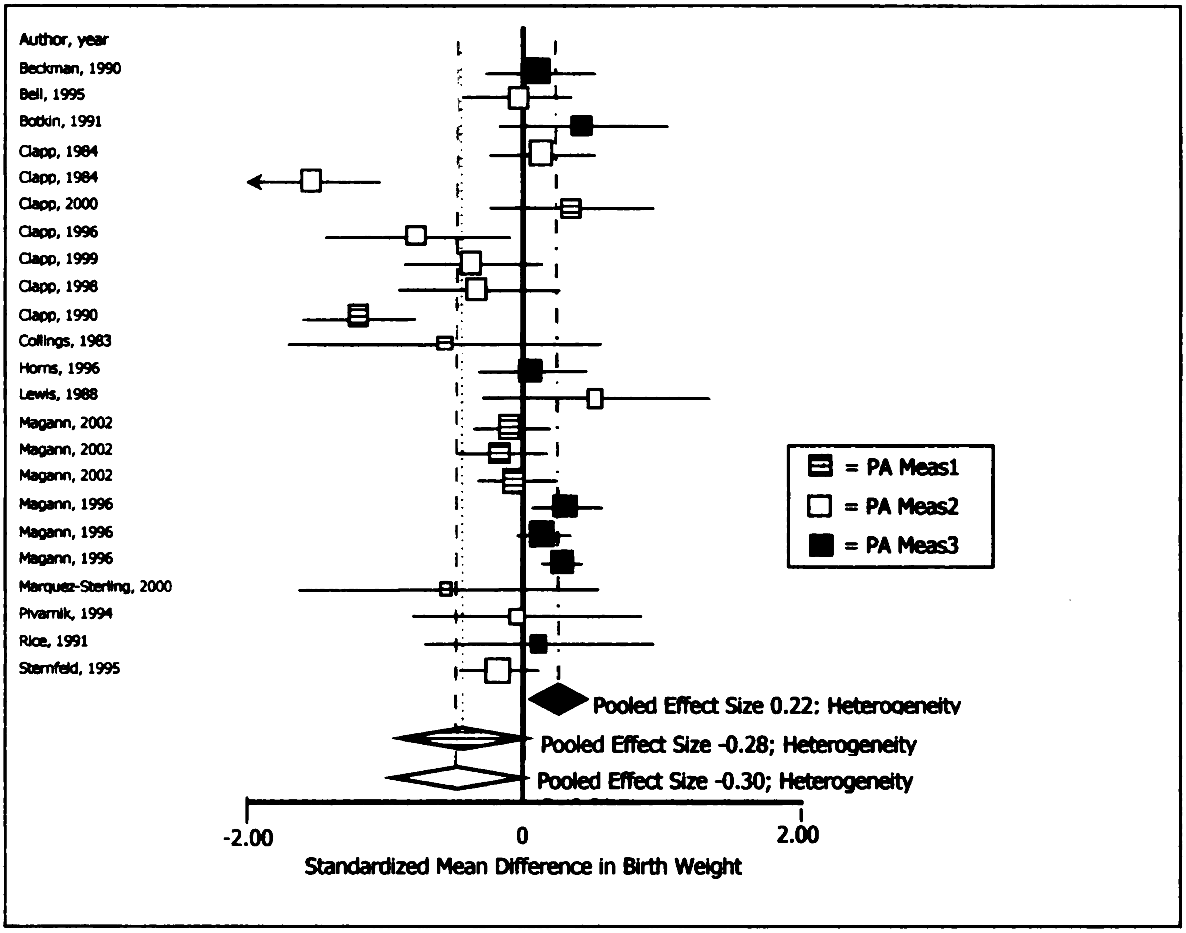
Note: Note: The vertical line represents no difference in the birth weight of babies born to physically active and sedentary women. -MWG = studies where the exercising group gained less weight during pregnancy than the nonexercising group; +MWG = studies where the exercising group gained more weight during pregnancy than the nonexercising group.

Figure 7. The Effect of Maternal Physical Activity on Birth Weight – Stratified by Maternal Age



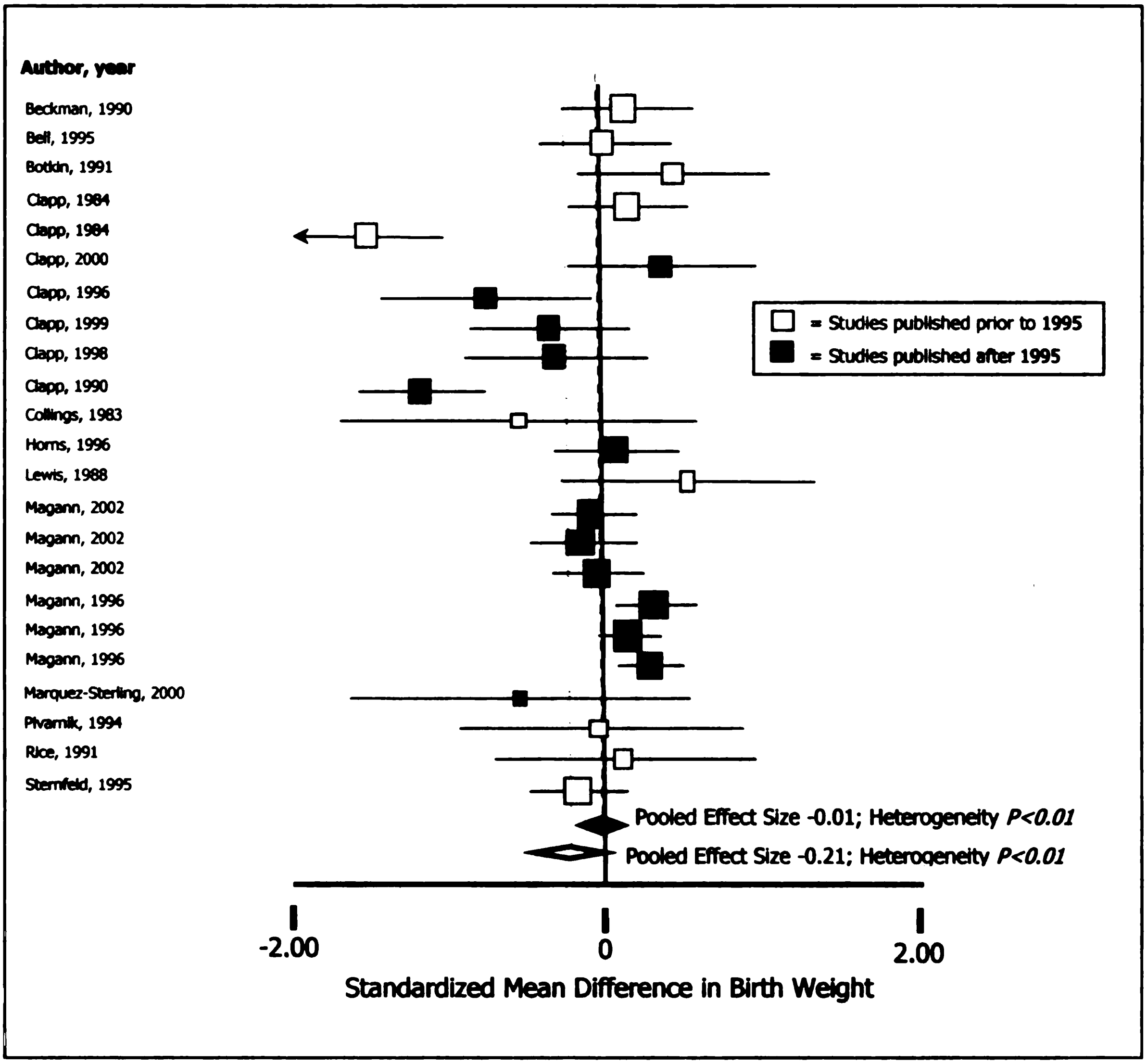
Note: The vertical line represents no difference in the birth weight of babies born to physically active and sedentary women.

Figure 8. The Effect of Maternal Physical Activity on Birth Weight – Stratified by Physical Activity Measurement Method.



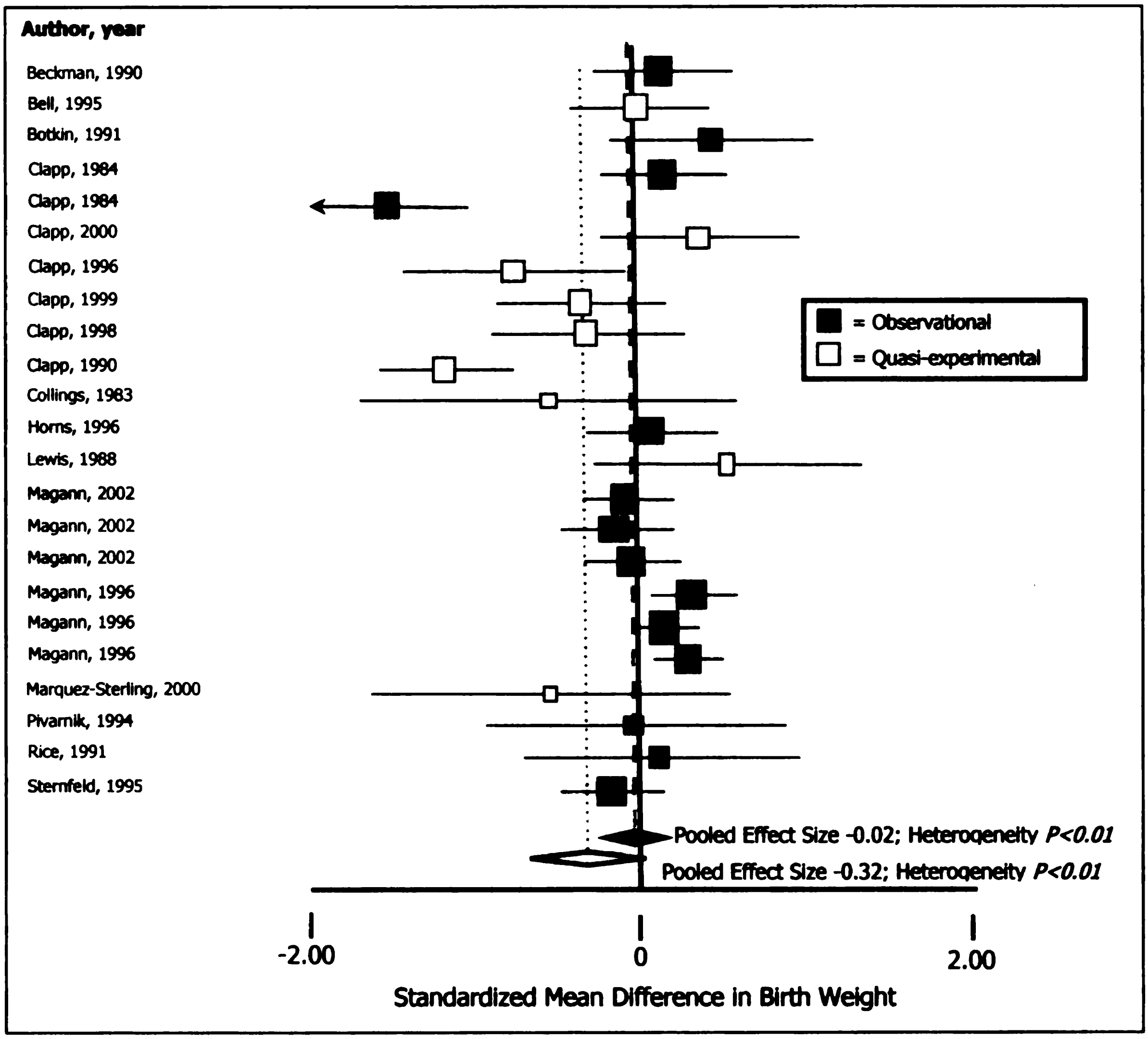
Note: The solid black vertical line represents no effect of physical activity on birth weight. PAmeas1 = studies where physical activity was directly monitored; PAmeas2 = studies where physical activity was measured via prospective self-report; PAmeas3 = studies where physical activity was measured via retrospective self-report.

Figure 9. The Effect of Maternal Physical Activity on Birth Weight – Stratified by Publication Year



Note: The vertical line represents no difference in the birth weight of babies born to physically active and sedentary women.

Figure 10. The Effect of Maternal Physical Activity on Birth Weight – Stratified by Study Design.



Note: The vertical line represents no difference in the birth weight of babies born to physically active and sedentary women.

CHAPTER 4

DISCUSSION

The main goal of a meta-analysis is to produce an estimate of the average effect seen in studies comparing the same exposures and providing some measure of the same outcome (Smith and Egger, 2001). The direction and magnitude of the average effect is intended to help guide public health decisions (Smith and Egger, 2001). The primary purpose of this meta-analysis was to determine whether physical activity during pregnancy affects birth weight. Additionally, a subsidiary purpose was to use pre-specified subgroups to explore the heterogeneity in birth weight.

The test for heterogeneity across all 23 comparisons for the effect of maternal physical activity on birth weight was statistically significant. Therefore, the hypothesis to research question one of this dissertation, that the pooled standardized mean difference in birth weight between exercising and sedentary mothers will be negative and not significantly heterogeneous, was rejected. Given the significant heterogeneity, the effect of maternal physical activity on birth weight was not clear. However, we suspected that the differences in birth weight between studies may be partially explained by other study- and subject characteristics.

Exploration of Heterogeneity

The purpose of a meta-analysis of a set of studies is quite different from the specific aims of any given individual study (Thompson, 2001). A single study designed to investigate the effect of physical activity on birth weight typically tests a single exposure, for a specific duration, to participants fulfilling certain eligibility criteria, using a particular definition of outcome measures. The purpose of a meta-analysis investigating the effect of physical activity on birth weight is broader, estimating the extent to which physical activity, achieved by a variety of means, generally influences birth weight. Because of the broader aims of a meta-analysis, the studies included usually encompass a substantial variety of exposure, types of participants, and outcomes (Thompson, 2001).

The heterogeneity exhibited in the overall analysis indicates the potential for several variables to modify the effect of physical activity on birth weight. Incompatibility in quantitative results (i.e., statistical heterogeneity) may be caused by known clinical or methodological differences between studies, or may be related to unknown or unrecorded trial characteristics (Thompson, 2001). The influence of these differences between studies on the overall results should be explored when heterogeneity exists. Exploration of heterogeneity was done by performing multiple subgroup analyses based on specific subject and study features.

In the case of maternal physical activity and birth weight, the most obvious causes of heterogeneity may be the trimester the physical activity was performed, the intensity of the physical activity performed, and the amount of weight gained during gestation. These variables have been shown to be influential such that birth weight is typically lower when physical activity is performed in late gestation (Clapp and Dickstein, 1984), and when gestational weight gain is less (Clapp et al., 1998; Clapp and Little, 1995). However, there are some studies that have shown the opposite effect, or even no association between trimester of physical activity (Magann, Evans, Weitz, and Newnham, 2002), intensity of physical activity (Magann, Evans, and Newnham, 1996; Kardel and Kase, 1998; Magann, Evans, Weitz, and Newnham, 2002), or maternal weight gain (Nuefeld, Haas, Grajeda, and Martorell, 2004) and birth weight. These individual studies prompted our initial subgroup analyses.

Subgroup Analysis: Physical Activity Intensity

Ten comparisons were available for the analysis of *moderate* physical activity on birth weight, and 13 comparisons were available for the analysis of *vigorous* physical activity on birth weight. In these analyses, there was nonsignificant heterogeneity (effect size estimates from the studies included did not differ greater than what was expected based on chance) among the moderate intensity studies and significant heterogeneity among the vigorous intensity studies. The overall effect of moderate intensity physical on birth weight was close to zero (ES=0.05), meaning that performance of moderate physical activity through the third trimester is unlikely to affect birth size. Alternatively, the overall effect of vigorous physical activity on birth weight was negative (ES=$^-$0.27). This translates to a biologically meaningful difference in birth weight of

~150 grams, and may be considered biologically meaningful, such that the decreased birth weight in these exercisers is comparable to the decreased birth weight associated with cigarette smoking during pregnancy (117-376 grams) (Fox, Koepsell, and Daling, 1984). However, the outcome of lower birth weight as a result of maternal physical activity may be different than the outcome of lower birth weight as a result of maternal cigarette smoking. Although a difference in birth weight of ~150 grams may be biologically meaningful, the analysis was significantly heterogeneous. This significant heterogeneity precludes further interpretation of the pooled mean difference. Therefore, because of the variability across study results, the impact of vigorous physical activity on birth weight remains unclear.

Visual examination of the forest plot of vigorous studies (Figure 5) raises some concern over potential outliers. It appears that there are two studies that may be potential outliers (Clapp and Dickstein, 1984; Clapp, 1990), showing a maximum difference in birth weight of ~850 grams between the exercisers and nonexercisers. The meta-analysis of vigorous physical activity, excluding these two studies remained significantly heterogeneous, while the pooled effect size moved close to zero (ES=-0.01; $P<0.01$). We were interested in what study characteristics may separate these 2 studies.

We carefully reexamined the 2 Clapp studies, comparing them to the remaining 11 that did not demonstrate such an extreme effect of physical activity on birth weight. In the study by Clapp and Dickstein (1984), the physically active group was a very fit sample of women who performed at a very rigorous intensity and long duration of high-impact, weight-bearing endurance exercise throughout pregnancy (>6 METs, ≥6 session/wk, ≥1 h). Even though the control group had slightly higher levels of smoking and were lower SES, two factors that are often associated with lower birth weight, the exercising group still had significantly smaller babies. The authors attribute these large differences to the differences in both gestational age and maternal weight gain. Women in the exercise group gained 2.4 kg less than the controls as of the last prenatal visit. Through statistical analysis, the authors determined that endurance exercise prior to pregnancy was not a significant determinant of weight gain; but rather, continued and sustained endurance exercise *during* pregnancy significantly reduced weight gain. These differences were only apparent in the latter half of pregnancy. In addition, the average gestational age of women in the exercise group was ~1 week shorter compared to the nonexercising group. Similar to the effect of weight gain, the significantly shorter gestation in the exercise group was shown to be affected by continued, sustained endurance exercise during pregnancy and not due to any endurance exercise prior to pregnancy. When groups were matched by demographics, differences in total weight gain and gestational age remained. Furthermore, the difference in birth weight between the exercisers and nonexercisers remained when gestational age was accounted for. The authors suggest that these findings differ from those in previous studies because of differences in design, group assignment, and exercise variables. The women were assigned to groups based on their actual exercise performance prior to and during pregnancy and the outcomes were examined on that basis. Statistically significant differences in the outcomes were limited to women who maintained their preconception exercise habits above minimum conditioning through the third trimester. The study design was advantageous is assessing actual exercise performance during pregnancy. The authors speculate that differing results in previous studies may be due to inconsistent changes in the actual performance of physical activity during pregnancy. Dale et al. (1982) found that by the 3rd trimester actual exercise performance of a group of runners had progressively fallen to 35% compared to level prior to pregnancy. Therefore, studies that do not account for changes in performance across the group over time, may be overestimating actual physical activity during pregnancy.

The second potential outlier among the vigorous studies was another study by Clapp (1990). In this study, the two groups of women were well matched for demographic, obstetric, morphometric, and life-style factors known to affect birth weight, and the exercising group had significantly smaller babies (~400 grams). There were no differences between groups with respect to weight gain and gestational age. A later study (Clapp and Capeless, 1990) assessing the neonatal morphometrics of the offspring in this cohort (Clapp, 1990) attributed the differences in birth weight to differences in fetal fat mass. The reduced fat mass of the infants born to exercisers did not create clinical difficulty for the babies in the neonatal period. Because very few studies have provided newborn body fatness measures, (Clapp and Capeless, 1990; Clapp, 1996; Clapp, Lopez, and Harcar-Sevcik, 1999) we were unable to perform stratified analyses based on this variable. However, in all of these studies, fetal fat mass and birth weight were lower in the exercising groups.

According to some current theories regarding birth weight, adiposity rebound, and subsequent predisposition to obesity-related diseases in adulthood (Eriksson et al., 2003), the lighter infant may be at greater risk of morbidity. However, if the leanness persists into adolescence and adulthood, lighter birth weight may be advantageous and lessen cardiovascular disease risk in later life (Clapp, 1990).

Subgroup Analysis: Maternal Weight Gain

It is biologically plausible that maternal weight gain acts both independently and combined with physical activity to affect birth weight. Women who are physically active during pregnancy often do not gain as much weight as their sedentary counterparts (Clapp, 1990). The relationship between maternal weight gain and birth weight is reported between $r={}^{-}0.37$ (Clapp and Dickstein, 1984) and $r={}^{-}0.42$ (Perkins et al., 2004), indicating that ~14-17% of the variance in birth weight can be explained by differences in maternal weight gain. In addition, Perkins et al., (2004) recently found that physical activity and gestational weight gain have both independent and combined effects on infant birth weight. The studies in this meta-analysis that provided data on the amount of weight gain during pregnancy (n=17) were divided into two groups, 1) exercisers who gained more than their sedentary counterparts, and 2) exercisers who gained less than their sedentary counterparts. Given the influence of maternal weight gain on birth weight in previous literature, we expected that by stratifying the women based on their weight gain, exercisers who did not gain as much weight as nonexercisers would also be those who had smaller babies.

There was nonsignificant heterogeneity in birth weights among the studies where exercisers gained as much or more than their nonexercising counterparts. The pooled effect size of birth weight in this group was zero, meaning that there was no difference in birth weight between exercisers and nonexercisers, when the exercisers gained more weight. We speculate that that some women who have active lifestyles are also more conscious of appropriate caloric intake. Women who compensate for the additional energy expenditure of physical activity by increasing caloric intake, and thus, gain comparable weight to a nonexerciser. This caloric compensation may offset the likelihood of having a smaller baby.

The studies where the exercising women gained less than the nonexercising women showed significant heterogeneity (*P=0.16*). Examination of this forest plot revealed that the exercisers who gained between 0.1 and 1.5 kg less than nonexercisers had babies of similar or heavier birth weight. While the exercisers who gained between 1.6 and 5.1 kg less than nonexercisers had smaller babies. This trend was validated by the results of the meta-regression of maternal weight gain on birth weight. The analysis showed that 17% of the variability in birth weight differences between exercisers and nonexercisers can be explained by the differences in maternal weight gain between these two groups. While prepregnancy size can contribute to maternal weight gain, there were not enough studies that included data on prepregnancy size (weight and height) in order to determine levels of appropriate weight gain among the participants. Future studies on birth weight should include both absolute maternal weight gain as well as gestational weight gain relative to prepregnancy body size.

Subgroup Analysis: Timing (Trimester) of Physical Activity

Previous research on the potential influence of trimester of physical activity on birth weight spurred the fourth research question of this dissertation. Initial ACOG guidelines for exercise during pregnancy suggested that women avoid certain intensities and types of physical activity in late gestation (ACOG, 1985). Clapp and Dickstein (1984) observed birth outcome in women who performed physical activity during pregnancy, but differed in the timing of physical activity. The authors found that those who continued exercising through mid- to late-gestation delivered lighter babies than women who either did not exercise in pregnancy or discontinued their exercise in early- to mid-gestation (Clapp and Dickstein, 1984). The majority of studies included in this meta-analysis provided data on women who exercised throughout the entire pregnancy. Consequently, we were unable to analyze differences in birth weight across studies stratified by trimester of physical activity. Similarly, while gestational length may be considered the strongest predictor of birth size, the studies included in this analysis varied little in gestational length. Therefore, we believe differences in gestational length would not likely explain the differences in birth weight found in these analyses.

Subgroup Analysis: Maternal Age

Studies were also grouped based on the age of the participants. The decision to dichotomize age above and below 27 years was the result of visually examining the mean ages of the participants in each included study. The mean age of the women in each study (in both exercise and control groups) was either greater than 27 or less than 27 years of age. By this categorization we were given two fairly equal subgroups ($AGE_{\leq 27}$ n=7; $AGE_{>27}$ n=13). These analyses suggest that the younger women were likely to have slightly heavier babies, regardless of physical activity level. The analysis of the older participants was significantly heterogeneous, but showed a tendency towards lighter babies, regardless of physical activity level.

The literature on maternal age and birth weight shows a relationship between adolescent (aged 14-17 years) and young (aged ~≥20 years) mothers and increased risk for growth restriction when compared to older pregnant women (~≥25 years) (Strobino, 1995; Mondal, 2004). However, these risks often diminish when the mother's social environment (i.e., poverty and minority status) is considered (Strobino, 1995). Previous research on advanced maternal age is not as clear. For example, Aldous et al. (1993) found that increasing maternal age in US Caucasian primiparas is an independent risk factor for low birth weight and preterm delivery. Alternatively, Stotland et al. (2004) found that maternal age between 30 and 40 years old was a risk factor for macrosomia (large birth weight).

The studies included in this meta-analysis were somewhat homogeneous with respect to maternal age (range 26.9-31.8 years) and therefore limited our analysis of the affect of maternal age on birth weight. Additionally, our meta-analysis did not contain studies where the outcome of birth weight was clinically low (<2500 grams) or clinically high (macrosomia; >4500). Although there was nonsignificant statistical heterogeneity in the studies of younger aged women, this result should be interpreted cautiously. This subgroup analysis included seven comparisons (k=7) that arose from only three studies (n=3) (Rice and Fort, 1991; Magann, Evans, and Newnham, 1996; Magann, Evans, Weitz, and Newnham, 2002), and 2 studies were published by the same research group (Magann, Evans, and Newnham, 1996; Magann, Evans, Weitz, and Newnham, 2002). Although statistically heterogeneous, the average mean difference in birth weight (-140 grams) in the studies whose participants were older than 27 years of age is biologically noteworthy. Age of the participant is typically a part of the study design (inclusion criteria) and is often a way in which experimental and control groups are matched. Future studies should examine the difference in birth outcomes among women in different age groups who are matched for intensity, frequency, and timing of physical activity, while controlling for maternal weight gain.

Subgroup Analysis: Publication Year

In hypothesizing characteristics that may contribute to heterogeneity among the included studies, we speculated that the publication of the updated ACOG guidelines in 1994 influenced medical practice and/or research and subsequently, birth weight. The dichotomization of studies published before and after 1995 allowed us to equally divide the included studies. We selected the year 1995 in assuming an average time to publish being one year after data is collected. However in doing so, we make the liberal assumption that these updated guidelines were applied in practical and research setting. The analyses of these subgroups ($YEAR_{<1995}$ and $YEAR_{>1995}$) were significantly heterogeneous, with no differences in the pooled effect size between the two strata.

Subgroup Analysis: Physical Activity Measurement

Since the methods available to quantify physical activity vary in accuracy, we speculated that the different methods employed in these studies might explain some of the heterogeneity in birth weight. The 18 studies included in this meta-analysis quantified physical activity by one of three methods: directly monitored (PAmeas-1), prospective questionnaire/self-report (PAmeas-2), or retrospective questionnaire/self-report (PAmeas-3). We hypothesized that studies where the subject's physical activity was directly monitored, produced more accurate and reliable measures of the exposure. Therefore, we expected nonsignificant heterogeneity in the subgroup of studies where physical activity was monitored.

There was significant heterogeneity among the seven studies where physical activity was monitored directly (Figure 8). Likewise, studies (n=9) where physical activity was self-reported prospectively through questionnaire or diary was also significantly heterogeneous (Figure 8). However, the studies where physical activity was retrospectively self-reported questionnaire or diary (n=7) were nonsignificantly heterogeneous with a pooled effect size of ES=0.22. This result indicates that babies of exercising women were on average 120 grams heavier than the babies of nonexercising women, when physical activity was obtained retrospectively. These results do not support the theory that direct observation of exercise session would produce similar results across several studies. From these results it is unclear whether one method of measuring physical activity is more accurate than another.

Subgroup Analysis: Study Design

Finally, the studies were stratified based on the study design, quasi-experimental or observational. The majority of research on maternal physical activity and birth outcome is primarily observational and quasi-experimental in study design. Although a randomized controlled intervention might theoretically result in less confounded results, it is often viewed as unethical and impractical to randomize pregnant women to exercising and nonexercising groups. In our literature search we did not exclude studies based on study design. However all studies in this meta-analysis were observational or quasi-experimental. Within the context of this research topic these two study designs differed primarily in the degree to which the subjects were "expected" to exercise. For example, in the observational studies, the women were able to freely choose whether or not they would exercise and to what extent they would exercise throughout pregnancy. In the quasi-experimental studies, women were often selected based on previous exercise history as well as their intention to exercise through pregnancy. By design, these women may have been more regimented about performing physical activity because they were placed in a group where exercise was the expectation. Therefore, we hypothesized that there may be a difference in women who were "expected" to exercise compared to women who were free to make lifestyle choices with respect to physical activity during pregnancy.

Both strata, observational and quasi-experimental studies, were varied on how physical activity data were collected, and contained both levels of physical activity intensity. The observational studies (n=14) were significantly heterogeneous. Inspection of the forest plot (Figure 10) indicates that the Clapp and Dickstein (1984) study was a potential outlier. When this study was removed from the analysis, the remaining studies were nonsignificantly heterogeneous, with a pooled effect size close to zero (ES=0.06). This analysis indicates that women who are free to engage in physically active or sedentary lifestyles do not differ with respect to the birth weight of their babies. In contrast, women who were expected to participate in an exercise regimen were likely to have lighter babies than controls, but these results were significantly heterogeneous.

Limitations

Clinical trials offer more comparable methodological procedures, often with very distinct measures of the exposure of interest, as compared to observational studies. While these studies may be more easily pooled, the clinical trial approach is not appropriate for exercising pregnant women. Although random allocation of pregnant women to exercise and control groups has been utilized (Clapp et al., 2000; Marquez-Sterling et al., 2000) this method may be considered unethical, and more likely, impractical. Compliance to a rigorously defined physical activity program may have to be curtailed due to maternal or fetal complications that arise during gestation. Therefore, while the clinical trial approach might be beneficial in helping to determine absolute levels of physical activity that may be harmful, it could also be argued that this approach does not simulate "real life". In observational and quasi-experimental studies the exposure of interest more closely reflects the exposure outside of the research setting.

The exposure variable of physical activity was quantified and reported in many different ways in the included studies. A typical general descriptive statement provided in some studies was 'the physically active subjects performed aerobics, running, or a walking regimen for at least 20 minutes at least 3 days per week, at an intensity of approximately 60% of their maximum heart rate'. From these descriptions we had to determine if the level of physical activity was considered moderate or vigorous intensity. In the process of categorizing the physical activity into these two levels we have potentially ignored details of the physical activity exposure that could affect the outcome of interest. For example, due to lack of sufficient detail being provided in the papers, we were unable to explore the various modes of physical activity and birth outcome. Additionally, we were often unable to determine what percentage of the exercising women performed each mode of physical activity and if these modes changed during pregnancy.

In order to further explore heterogeneity, we were interested in determining whether some known or suspected variables would account for the differences among studies. Our analyses were guided by previous research as well as our own hypotheses driven by scientific reasoning. Because of the complexity of the data reported across all the studies, there were numerous ways that each potentially modifying variable could be explored. For example our subgroup analysis based on maternal age was performed by dichotomizing the mean age of the study women at 27 years. By providing individual data on subject age we may better be able to investigate the effect of maternal age and birth weight. Given that we only had access to the mean age of the women in each study group, our analyses were restricted to some level of dichotomization. This subgroup analysis of age may have been performed at multiple different cut points. The idea that one variable may be examined multiple ways can be applied to most of our subgroup analysis categorization criteria.

This meta-analysis was a quantitative systematic review of group data. The studies included in this meta-analysis contained two to three groups of women based on their physical activity level during pregnancy. The majority of these studies matched, or showed no significant differences, between groups with respect to a variety of subject characteristics that may confound the relationship between maternal physical activity and birth weight (i.e., parity, smoking status, SES). Given that there was not a lot of variability in these factors, these potential confounders were controlled. It is not possible to statistically control for such confounders when group data is used in the meta-analysis. Controlling for suspected confounders requires the systematic review of individual-level data from the selected studies.

Selection bias may have affected the results of the individual studies where women self-select into exercising and nonexercising groups. Additionally, selection bias was a potential factor associated with the location and selected of studies for this meta-analysis, such that only published studies indexed in *NLM Medline* and *Embase* or cited in relevant articles, were eligible to be included in this meta-analysis.

Conclusion

The results from these analyses are encouraging for women who wish to participate in physical activity, at both moderate and/or vigorous intensities, during pregnancy. Quantification of all studies done on physical activity and birth weight, by our inclusion and exclusion criteria, do not show that women who are physically active during pregnancy deliver smaller babies. However, when considering vigorous exercise studies alone, the trend shows that birth weight might be slightly less, although these results were too variable to know for sure.

The CDC/ACSM recommendations for moderate exercise were recently adopted by ACOG for use during pregnancy. The results of this meta-analysis indicate that women who participate in physical activity at an intensity less than ~5 METs for at least 5 days a week (moderate) deliver babies similar in birth weight to women who remain sedentary throughout pregnancy. It may be expected that women can obtain physical activity benefits during pregnancy (i.e., enhanced psychological well-being, maintenance of fitness, reduced cardiovascular stress, prevention of low back pain, reduced labor duration, quicker recovery from labor, and prevention of excessive weight gain) without affecting the birth weights of their babies. With the anticipated update to the CDC/ACSM guidelines to reemphasize the recommendations for *vigorous* physical activity, it was worthwhile to determine the effects of this intensity on birth weight in the event that ACOG may wish to amend their guidelines to include this physical activity intensity. It is noteworthy that the majority of studies where women exercised at an intensity exceeding 6 METs for at least 3 days per week delivered babies lighter than women who did not perform any physical activity. However, not every study indicated an effect size in the negative direction, which gave rise to the statistical heterogeneity in this subgroup analysis. We cannot conclude from this meta-analysis that women who wish to perform physical activity in accordance with the recommendations for vigorous activity will have lighter babies compared to those who remain sedentary.

Given the trend of physically active women delivering lighter babies, future studies should explore this relationship with respect to fetal origins theory and later life morbidities. It is possible that the relationship between birth weight and childhood morbidities may no longer be apparent when physical activity of the mother is considered. Additionally, future studies should explore the potential association of infant body composition and maternal physical activity.

The medical community's attitudes and beliefs about exercising during pregnancy have changed over time. Where we once thought pregnancy to be a state of illness, such that the expectant mother was restricted to bed rest, experts now contend that exercising during pregnancy can be advantageous for both the mother and unborn child. These past studies were performed prior to the adoption of physical activity recommendations for this population (CDC/ACSM recommendations) and were performed primarily to examine the safety of physical activity during pregnancy. As the growing body of literature demonstrates that physical activity is not likely to have deleterious effects on the maternal-fetal unit, and with additional application of the CDC/ACSM recommendations for physical activity during pregnancy, we can expect to see more studies in this area in the future.

APPENDICES

APPENDIX A

List of studies included in the meta-analysis
(Appearing in alphabetical order and separated by physical activity level)

Moderate intensity (does not meet CDC/ACSM recommendations)

Beckmann, C.R. and C.A. Beckmann, Effect of a structured antepartum exercise program on pregnancy and labor outcome in primiparas. J Reprod Med, 1990. 35: 704-709.

Bell, R.J., S.M. Palma, and J.M. Lumley, The effect of vigorous exercise during pregnancy on birth-weight. Aust N Z J Obstet Gynaecol, 1995. 35: 46-51.

Botkin, C. and C.E. Driscoll, Maternal aerobic exercise: newborn effects. Fam Pract Res J,1991. 11: 387-93.

Horns, P.N., L.P. Ratcliffe, J.C. Leggett, and M.S. Swanson. Pregnancy outcomes among active and sedentary primiparous women. J Obstet Gynecol Neonatal Nurs, 1996. 25: 49-54.

Lewis, R.D., C.Y. Yates, and J.A. Driskell, Riboflavin and thiamin status and birth outcome as a function of maternal aerobic exercise. Am J Clin Nutr, 1988. 48: 110-116.

Marquez-Sterling, S., A.C. Perry, T.A. Kaplan, R.A. Halberstein, and J.F. Signorile. Physical and psychological changes with vigorous exercise in sedentary primigravidae. Med Sci Sports Exerc, 2000. 32: 58-62.

Rice, P.L. and I.L. Fort, The relationship of maternal exercise on labor, delivery and health of the newborn. J Sports Med Phys Fitness, 1991. 31: p. 95-99.

Sternfeld, B., Physical activity and pregnancy outcome. Review and recommendations. Sports Med, 1997. 23: 33-47.

Moderate intensity (meets CDC/ACSM recommendations)

Magann, E.F., S.F. Evans, and J.P. Newnham, Employment, exertion, and pregnancy outcome: assessment by kilocalories expended each day. Am J Obstet Gynecol, 1996. 175: 182-187.

Magann, E.F., S.F. Evans, B. Weitz, and J. Newnham. Antepartum, intrapartum, and neonatal significance of exercise on healthy low-risk pregnant working women. Obstet Gynecol, 2002. 99: p. 466-472.

Vigorous intensity (meets CDC/ACSM recommendations)

Clapp, J.F. III, and S. Dickstein, Endurance exercise and pregnancy outcome. Med Sci Sports Exerc, 1984. 16: p. 556-562.

Clapp, J.F. III, The course of labor after endurance exercise during pregnancy. Am J Obstet Gynecol, 1990. 163: 1799-1805.

Clapp, J.F. III, Morphometric and neurodevelopmental outcome at age five years of the offspring of women who continued to exercise regularly throughout pregnancy. J Pediatr, 1996. 129: 856-63.

Clapp, J.F. III, S. Simonian, B. Lopez, S. Appleby-Wineberg, and R. Harcar-Sevcik. The one-year morphometric and neurodevelopmental outcome of the offspring of women who continued to exercise regularly throughout pregnancy. Am J Obstet Gynecol, 1998. 178: 594-599.

Clapp, J.F., 3rd, B. Lopez, and R. Harcar-Sevcik, Neonatal behavioral profile of the offspring of women who continued to exercise regularly throughout pregnancy. Am J Obstet Gynecol, 1999. 180: 91-94.

Clapp, J.F. III, H. Kim, B. Burciu, and B. Lopez. Beginning regular exercise in early pregnancy: effect on fetoplacental growth. Am J Obstet Gynecol, 2000. 183: 1484-1488.

Collings, C.A., L.B. Curet, and J.P. Mullin, Maternal and fetal responses to a maternal aerobic exercise program. Am J Obstet Gynecol, 1983. 145: 702-707.

[¥]Magann, E.F., S.F. Evans, and J.P. Newnham, Employment, exertion, and pregnancy outcome: assessment by kilocalories expended each day. Am J Obstet Gynecol, 1996. 175: 182-187.

[¥]Magann, E.F., S.F. Evans, B. Weitz, and J. Newnham. Antepartum, intrapartum, and neonatal significance of exercise on healthy low-risk pregnant working women. Obstet Gynecol, 2002. 99: 466-472.

Pivarnik, J.M., M.B. Mauer, N.A. Ayres, B. Kirshon, G.A. Dildy, and D.B. Cotton. Effects of chronic exercise on blood volume expansion and hematologic indices during pregnancy. Obstet Gynecol, 1994. 83: 265-269.

[¥] References appearing in multiple categories of physical activity intensity contained more than one exercise group in the study. Statistical measures were taken to account for the repeated use of the control group in these studies.

APPENDIX B

List of studies not included in this analysis

Adair L.S., and Cole T.J. Rapid child growth raises blood pressure in adolescent boys who were thin at birth. Hypertension, 2003. 41: 451-456.

Alderman, B.W., et al., Maternal physical activity in pregnancy and infant size for gestational age. Ann Epidemiol, 1998. 8: 513-519.

Berkowitz, G.S., J.L. Kelsey, T.R. Holford, R.L. Berkowitz. Physical activity and the risk of spontaneous preterm delivery. J Reprod Med, 1983. 28: 581-588.

Bung, P., C. Bung, R. Artal, N. Khodiguian, F. Fallenstein, and L. Spatling. Therapeutic exercise for insulin-requiring gestational diabetics: effects on the fetus--results of a randomized prospective longitudinal study. J Perinat Med, 1993. 21: 125-137.

Bungum, T.J., D.L. Peaslee, A.W. Jackson, and M.A. Perez. Exercise during pregnancy and type of delivery in nulliparae. J Obstet Gynecol Neonatal Nurs, 2000. 29: 258-264.

Campbell, M.K. and M.F. Mottola, Recreational exercise and occupational activity during pregnancy and birth weight: a case-control study. Am J Obstet Gynecol, 2001. 184: 403-408.

Clapp, J.F. III, The effects of maternal exercise on early pregnancy outcome. Am J Obstet Gynecol, 1989. 161: 1453-1457.

Clapp, J.F. III, and E.L. Capeless, Neonatal morphometrics after endurance exercise during pregnancy. Am J Obstet Gynecol, 1990. 163: 1805-1811.

Clapp, J.F. III, R. Rokey, J.L. Treadway, M.W. Carpenter, R.M. Artal, and C. Warnes. Exercise in pregnancy. Med Sci Sports Exerc, 1992. 24(6 Suppl): S294-S300.

Clapp, J.F. III, and K.D. Little, Effect of recreational exercise on pregnancy weight gain and subcutaneous fat deposition. Med Sci Sports Exerc, 1995. 27: 170-177.

Clapp, J.F. III, H. Kim, B. Burciu, S. Schmidt, K. Petry, and B. Lopez. Continuing regular exercise during pregnancy: effect of exercise volume on fetoplacental growth. Am J Obstet Gynecol, 2002. 186: 142-147.

Clapp, J.F., 3rd, K.D. Little, and J.A. Widness, Effect of maternal exercise and fetoplacental growth rate on serum erythropoietin concentrations. Am J Obstet Gynecol, 2003. 188: 1021-1025.

Collings, C. and L.B. Curet, Fetal heart rate response to maternal exercise. Am J Obstet Gynecol, 1985. 151: 498-501.

Dale, E., K.M. Mullinax, and D.H. Bryan, Exercise during pregnancy: effects on the fetus. Can J Appl Sport Sci, 1982. 7: 98-103.

Hall, D.C. and D.A. Kaufmann, Effects of aerobic and strength conditioning on pregnancy outcomes. Am J Obstet Gynecol, 1987. 157: 1199-1203.

Hatch, M.C., X.O. Shu, D.E. McLean, B. Levin, M. Begg, L. Reuss, and M. Susser. Maternal exercise during pregnancy, physical fitness, and fetal growth. Am J Epidemiol, 1993. 137: 1105-1114.

Hatch, M., B. Levin, X.O. Shu, and M. Susser. Maternal leisure-time exercise and timely delivery. Am J Public Health, 1998. 88: 1528-1533.

Jarrett, J.C., 2nd and W.N. Spellacy, Jogging during pregnancy: an improved outcome? Obstet Gynecol, 1983. 61: 705-709.

Kardel, K.R. and T. Kase, Training in pregnant women: effects on fetal development and birth. Am J Obstet Gynecol, 1998. 178: 280-6.

Kennelly, M.M., N. McCaffrey, P. McLoughlin, S. Lyons, and P. McKenna. Fetal heart rate response to strenuous maternal exercise: not a predictor of fetal distress. Am J Obstet Gynecol, 2002. 187: 811-816.

Kulpa, P.J., B.M. White, and R. Visscher, Aerobic exercise in pregnancy. Am J Obstet Gynecol, 1987. 156: 1395-1403.

Leiferman, J.A. and K.R. Evenson, The effect of regular leisure physical activity on birth outcomes. Matern Child Health J, 2003. 7: 59-64.

Lynch, A.M., S. McDonald, E.F. Magann, S.F. Evans, P.L. Choy, B. Dawson, B.A. Blanksby, and J.P. Newnham. Effectiveness and safety of a structured swimming program in previously sedentary women during pregnancy. J Matern Fetal Neonatal Med, 2003. 14: 163-169.

Mullinax, K.M. and E. Dale, Some considerations of exercise during pregnancy. Clin Sports Med, 1986. 5: 559-70.

Rabkin, C.S., H.R. Anderson, J.M. Bland, O.G. Brooke, G. Chamberlain, and J.L. Peacock. Maternal activity and birth weight: a prospective, population-based study. Am J Epidemiol, 1990. 131: 522-531.

Rao, S., A. Kanade, B.M. Margetts, C.S. Yajnik, H. Lubree, S. Rege, B. Desai, A. Jackson, C.H. Fall. Maternal activity in relation to birth size in rural India. The Pune Maternal Nutrition Study. Eur J Clin Nutr, 2003. 57: 531-542.

Ruhling, R.O., J. Cameron, L. Sibley, C.L. Christensen, and T. Bolen. Maintaining Aerobic Fitness While Jogging Through A Pregnancy: A Case Study. Med Sci Sports Exerc 1981. 13: 93.

Rutkowska, E. and C. Lepecka-Klusek, The role of physical activity in preparing women for pregnancy and delivery in Poland. Health Care Women Int, 2002. 23: 919-923.

Sayer A.A., H.E. Syddall, E.M. Dennison, H.J. Gilbody, S.L. Duggleby, C. Cooper, D.J. Barker, and D.I. Phillips. Birth weight, weight at 1 y of age, and body composition in older men: findings from the Hertfordshire Cohort Study. Am J Clin Nutr 2004. 80: 199-203.

Sibley, L., et al., Swimming and physical fitness during pregnancy. J Nurse Midwifery, 1981. 26: 3-12.

Sorensen, T.K., M.A. Williams, I.M. Lee, E.E. Dashow, M.L. Thompson, and D.A. Luthy. Recreational physical activity during pregnancy and risk of preeclampsia. Hypertension, 2003. 41: 1273-1280.

Zhang, J. and D.A. Savitz, Exercise during pregnancy among US women. Ann Epidemiol, 1996. 6: 53-59.

REFERENCES

REFERENCES

Abrams, R., D. Caton, J. Clapp, and D.H. Barron. Thermal and metabolic features of life in utero. Clin Obstet Gynecol, 1970. 13: 549-564.

Adair L.S., and Cole T.J. Rapid child growth raises blood pressure in adolescent boys who were thin at birth. Hypertension, 2003. 41: 451-456.

Adamson, K.J., The role of thermal factor in fetal and neonatal life. Pediatr Clin North Am, 1966. 13: 599-619.

American College of Obstetricians and Gynecologists, ACOG home exercise programs – Exercise during pregnancy and the postnatal period. 1985, Division of Education of the American College of Obstetricians and Gynecologists.

American College of Obstetricians and Gynecologists, ACOG Technical Bulletin: Exercise During Pregnancy and the Postpartum Period. 1994, The American College of Obstetricians and Gynecologists: Washington, DC.

American College of Obstetricians and Gynecologists, ACOG committee opinion. Exercise during pregnancy and the postpartum period. Number 267, January 2002.

Artal, R. and M. O'Toole, Guidelines of the American College of Obstetricians and Gynecologists for exercise during pregnancy and the postpartum period. Br J Sports Med, 2003. 37: 6-12.

Barker D.J., C. Osmond, J. Golding, D. Kuh, and M.E. Wadsworth. Growth in utero, blood pressure in childhood and adult life, and mortality from cardiovascular disease. BMJ. 1989. 298 (6673): 564-567.

Barker, D.J., Fetal origins of coronary heart disease. BMJ, 1995. 311 (1998): 171-174.

Barker, D.J. and C.N. Martyn, The maternal and fetal origins of cardiovascular disease. J Epidemiol Community Health, 1992. 46: p. 8-11.

Bell, R., The effects of vigorous exercise during pregnancy on birth weight. J Sci Med Sport, 2002. 5: 32-36.

Brooks, G.A., T.D. Fahey, and T.P. White. Energetics and Human Movement, in Exercise Physiology: Human Bioenergetics and Its Applications, 3rd ed. 2000, McGraw-Hill Higher Education. p. 29-33.

Butterfield, G. and J. King, Nutritional Needs of Physically Active Pregnant Women, in Exercise in Pregnancy, R. Mittelmark, R. Wiswell, and B. Drinkwater, Editors. 1991, Williams & Wilkins. p. 109-122.

Caspersen, C.J., Physical activity epidemiology: concepts, methods, and applications to exercise science. Exerc Sport Sci Rev, 1989. 17: 423-473.

Caspersen, C.J., K.E. Powell, and G.M. Christenson, Physical activity, exercise, and physical fitness: definitions and distinctions for health-related research. Public Health Rep, 1985. 100: 126-31.

Centers for Disease Control and Prevention, Data 2010: The healthy people 2010 database. Healthy People 2010 - Leading health indicators at a glance. 2000, U.S. Department of Health and Human Services. Centers for Disease Control and Prevention. The National Center for Health Statistics. www.cdc.gov/nchs.

Cefalo, R. and A. Hellegers, The effect of maternal hyperthermia on maternal and fetal cardiovascular and respiratory function. Am J Obstet Gynecol, 1978. 1331: 687-394.

Clapp, J., Exercise during pregnancy., in Volume 9: perspectives in exercise science and sports medicine: exercise and the female - a life span approach., O. Bar-Or, D. Lamb, and P. Clarkson, Editors. 1996, Cooper Publishing Group: Carmel. p. 413-451.

Clapp, J.F. III, Acute exercise stress in the pregnant ewe. Am J Obstet Gynecol, 1980. 136: 489-94.

Clapp, J.F. III,. The changing thermal response to endurance exercise during pregnancy. Am J Obstet Gynecol, 1991. 165: 1684-1689.

Clapp, J.F. III, Morphometric and neurodevelopmental outcome at age five years of the offspring of women who continued to exercise regularly throughout pregnancy. J Pediatr, 1996. 129: 856-63.

Clapp, J.F. III, Maternal carbohydrate intake and pregnancy outcome. Proc Nutr Soc, 2002. 61: p. 45-50.

Clapp, J.F. III, and E.L. Capeless, Neonatal morphometrics after endurance exercise during pregnancy. Am J Obstet Gynecol, 1990. 163: 1805-11.

Clapp, J.F. III, and S. Dickstein, Endurance exercise and pregnancy outcome. Med Sci Sports Exerc, 1984. 16: 556-562.

Clapp, J.F. III, H. Kim, B. Burciu, and B. Lopez. Beginning regular exercise in early pregnancy: effect on fetoplacental growth. Am J Obstet Gynecol, 2000. 183: 1484-1488.

Clapp, J.F. III, H. Kim, B. Burciu, S. Schmidt, K. Petry, and B. Lopez. Continuing regular exercise during pregnancy: effect of exercise volume on fetoplacental growth. Am J Obstet Gynecol, 2002. 186: 142-147

Clapp, J.F. III, and K.D. Little, Effect of recreational exercise on pregnancy weight gain and subcutaneous fat deposition. Med Sci Sports Exerc, 1995. 27: 170-177.

Clapp, J.F. III, B. Lopez, and R. Harcar-Sevcik, Neonatal behavioral profile of the offspring of women who continued to exercise regularly throughout pregnancy. Am J Obstet Gynecol, 1999. 180: 91-94.

Clapp, J.F. III, S. Simonian, B. Lopez, S. Appleby-Wineberg, and R. Harcar-Sevcik. The one-year morphometric and neurodevelopmental outcome of the offspring of women who continued to exercise regularly throughout pregnancy. Am J Obstet Gynecol, 1998. 178: 594-599.

Clapp, J.F. III, M. Wesley, and R.H. Sleamaker, Thermoregulatory and metabolic responses to jogging prior to and during pregnancy. Med Sci Sports Exerc, 1987. 19: 124-30.

Collings, C.A., L.B. Curet, and J.P. Mullin, Maternal and fetal responses to a maternal aerobic exercise program. Am J Obstet Gynecol, 1983. 145: 702-707.

Counsell, C. Formulating questions and locating primary studies for inclusion in systematic reviews. Ann Intern Med, 1997. 127: 380-387.

Deeks, J., D. Altman, and M. Bradburn, Statistical methods for examining heterogeneity and combining results from several studies in meta-analysis, in Systematic Reviews in Health Care: Meta-analysis in context, M. Egger, D. Smith, and D. Altman, Editors. 2001, BMJ Publishing Group. p. 285-312.

DerSimonian, R. and N. Laird, Meta-analysis in clinical trials. Control Clin Trials, 1986. 7: 177-188.

Egger, M., G. Smith, and K. O'Rourke, Rationale, potentials, and promise of systematic reviews, in Systematic Reviews in Health Care: Meta-analysis in context, M. Egger, G. Smith, and D. Altman, Editors. 2001, BMJ Publishing Group: London. p. 3-22.

Enanoria, W.T., C. Ng, S.R. Saha, and J.M. Colford Jr. Treatment outcomes after highly active antiretroviral therapy: a meta-analysis of randomised controlled trials. Lancet Infect Dis, 2004. 4: 414-425.

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

Evenson, K.R., D.A. Savitz, and S.L. Huston. Leisure-time physical activity among pregnant women in the US. Paediatr Perinat Epidemiol 2004. 18: 400-407.

Glass, G., Primary, secondary, and meta-analysis of research. Educational Researcher, 1976. 5: 3-8.

Godfrey, K.M. and D.J. Barker, Fetal programming and adult health. Public Health Nutr, 2001. 4: 611-624.

Goodlin, R.C., Low birth weight infants. Obstet Gynecol, 1984. 64: 446-447.

Gorski, J., Exercise during pregnancy: maternal and fetal responses. A brief review. Med Sci Sports Exerc, 1985. 17: 407-416.

Hall, D.C. and D.A. Kaufmann, Effects of aerobic and strength conditioning on pregnancy outcomes. Am J Obstet Gynecol, 1987. 157: 1199-1203.

Hatch, M. and Z. Stein, Work and Exercise During Pregnancy: Epidemiological Studies, in Exercise in Pregnancy, R. Milttlemark Artal, R. Wiswell, and B. Drinkwater, Editors. 1991, Williams & Wilkins. p. 279-286.

Hatch, M.C., X.O. Shu, D.E. McLean, B. Levin, M. Begg, L. Reuss, and M. Susser. Maternal exercise during pregnancy, physical fitness, and fetal growth. Am J Epidemiol, 1993. 137: 1105-1114.

Hytten, F., Nutrition., in Clinical physiology in obstetrics, F. Hytten and G. Chamberlain, Editors. 1980, Blackwell Scientific Publications: Oxford. p. 163-192.

Huxley, R., Neil A., and Collins R. Unraveling the fetal origins hypothesis: is there really an inverse association between birthweight and subsequent blood pressure? Lancet. 2002. 360(9334): 659-65.

Institute of Medicine. Historical Trends in Clinical Practice, Maternal Nutritional Status, and the Course and Outcome of Pregnancy, in Nutrition During Pregnancy: Part I. Weight Gain, Pt. II. Nutrient Supplements, I.o. Medicine, Editor. 1990, National Academy Press: Washington, DC. p. 37-62.

Institute of Medicine. Energy Requirements, Energy Intake, and Associated

Weight Gain During Pregnancy, in Nutrition During Pregnancy: Part I. Weight Gain, Pt. II. Nutrient Supplements, I.o. Medicine, Editor. 1990, National Academy Press: Washington, DC. p. 137-175.

Jadad, A.R., D. Moher, and T.P. Klassen, Guides for reading and interpreting systematic reviews: II. How did the authors find the studies and assess their quality? Arch Pediatr Adolesc Med, 1998. 152: 812-817.

Kardel, K.R. and T. Kase, Training in pregnant women: effects on fetal development and birth. Am J Obstet Gynecol, 1998. 178: 280-286.

Karn, M.N. and L.S. Penrose, Birth weight and gestation time in relation to maternal age, parity and infant survival. Ann Eugen, 1952. 16: 147-164.

Kohl, H. Mini-Symposium - CDC/ACSM recommendations for Physical Activity and Public Health: History, a Decade of Experience, and a Look to the Future. Kohl HW. Updating the Recommendations - Looking to the future. in American College of Sports Medicine National Conference. 2004. Indianapolis, IN.

Koos, B., G. Power, and L. Longo, Placental oxygen transfer with considerations for maternal exercise, in Exercise in Pregnancy, R. Artal Mittelmark, R. Wiswell, and B. Drinkwater, Editors. 1991, Williams & Wilkins. p. 71-86.

Kramer, M.S., Intrauterine growth and gestational duration determinants. Pediatrics, 1987. 80: 502-511.

Kramer, M.S., Regular aerobic exercise during pregnancy (Cochrane Review). The Cochrane Library, 2001(3): p. Issue 3, CD000180.

Landis, J.R. and G.G. Koch, The measurement of observer agreement for categorical data. Biometrics, 1977. 33: 159-174.

Law, C.M. and D.J. Barker, Fetal influences on blood pressure. J Hypertens, 1994. 12: 1329-1332.

Lewis, R.D., C.Y. Yates, and J.A. Driskell, Riboflavin and thiamin status and birth outcome as a function of maternal aerobic exercise. Am J Clin Nutr, 1988. 48: 110-116.

Lokey, E.A., Z.V. Tran, C.L. Wells, B.C. Myers, and A.C. Tran. Effects of physical exercise on pregnancy outcomes: a meta-analytic review. Med Sci Sports Exerc, 1991. 23: 1234-1239.

Lotgering, F.K., R.D. Gilbert, and L.D. Longo, Exercise responses in pregnant sheep: oxygen consumption, uterine blood flow, and blood volume. J Appl Physiol, 1983. 55: 834-841.

Lotgering, F.K., R.D. Gilbert, and L.D. Longo, Exercise responses in pregnant sheep: blood gases, temperatures, and fetal cardiovascular system. J Appl Physiol, 1983. 55: 842-850.

Magann, E.F., S.F. Evans, and J.P. Newnham, Employment, exertion, and pregnancy outcome: assessment by kilocalories expended each day. Am J Obstet Gynecol, 1996. 175: 182-187.

Magann, E.F., S.F. Evans, B. Weitz, and J. Newnham. Antepartum, intrapartum, and neonatal significance of exercise on healthy low-risk pregnant working women. Obstet Gynecol, 2002. 99: 466-472.

Marquez-Sterling, S., A.C. Perry, T.A. Kaplan, R.A. Halberstein, and J.F. Signorile. Physical and psychological changes with vigorous exercise in sedentary primigravidae. Med Sci Sports Exerc, 2000. 32: 58-62.

McCormick, M.C., The contribution of low birth weight to infant mortality and childhood morbidity. N Engl J Med, 1985. 312: 82-90.

McGinnis, J., The public health burden of a sedentary lifestyle. Med Sci Sports Exerc, 1992. 4 (supplement): S196-S200.

Meyer, B.A. and J.R. Daling, Activity level of mother's usual occupation and low infant birth weight. J Occup Med, 1985. 27: 841-847.

Miller, P., D. Smith, and T. Shepard, Maternal hyperthermia as a possible cause of anencephaly. Lancet, 1978. 1: 519.

Milunsky, A., M. Ulcickas, K.J. Rothman, W. Willett, S.S. Jick, and H. Jick. Maternal heat exposure and neural tube defects. JAMA, 1992. 268: 882-885.

Morishima, H.O., M.N. Yeh, W.H. Niemann, and L.S. James. Temperature gradient between fetus and mother as an index for assessing intrauterine fetal condition. Am J Obstet Gynecol, 1977. 129: 443-448.

Morton, M.J., M.S. Paul, and J. Metcalfe, Exercise during pregnancy. Med Clin North Am, 1985. 69: 97-108.

National Research Council, Recommended Dietary Allowances, 10th ed. Report of the Subcommittee on the Tenth Edition of the RDAs, Food and Nutrition Board, Commission on Life Sciences. 1989, Washington, DC: National Academy Press.

NFSHSA, 2001 Athletics Participation Totals. 2001, National Federation of State High School Associations.

Neufeld, L.M., J.D. Haas, R. Grajeda, and R. Martorell. Changes in Maternal Weight from the First to Second Trimester of Pregnancy Are Associated With Fetal Growth and Infant Length at Birth. Obstet Gynecol Surv., 2004. 59: 643-644.

Pate, R.R., M. Pratt, S.N. Blair, W.L. Haskell, C.A. Macera, C. Bouchard, D. Buchner, W. Ettinger, G.W. Heath, A.C. King, et al. Physical activity and public health. A recommendation from the Centers for Disease Control and Prevention and the American College of Sports Medicine. JAMA, 1995. 273: 402-407.

Perkins, C.D., J.M. Pivarnik, A.D. Stein, and J.M. Rivera. Recommended weight gain, maternal physical activity and fetal growth. Med Sci Sports Exerc 2004. 36 (supplement): S296.

Pleet, H., J. Graham, and M. Harvey, Patterns of malformations resulting from the teratogenic effects of trimester hyperthermia. Pediatr Res, 1980. 14: 587.

Power, G.G., H. Schroder, and R.D. Gilbert, Measurement of fetal heat production using differential calorimetry. J Appl Physiol, 1984. 57: 917-922.

Rabkin, C.S., H.R. Anderson, J.M. Bland, O.G. Brooke, G. Chamberlain, and J.L. Peacock. Maternal activity and birth weight: a prospective, population-based study. Am J Epidemiol, 1990. 131: 522-531.

Rice, P.L. and I.L. Fort, The relationship of maternal exercise on labor, delivery and health of the newborn. J Sports Med Phys Fitness, 1991. 31: 95-99.

Romen, Y., D. Maksaki, and R. Mittelmark, Physiological and Endocrine Adjustments to Pregnancy, in Exercise in Pregnancy, R. Milttlemark Artal, R. Wiswell, and B. Drinkwater, Editors. 1991, Williams & Wilkins. p. 9-29.

Rozycki, T., Oral and rectal temperature in runners. Phys Sportsmed, 1984. 12: 105-108.

Sady, S.P. and M.W. Carpenter, Aerobic exercise during pregnancy. Special considerations. Sports Med, 1989. 7: 357-375.

Sayer A.A., H.E. Syddall, E.M. Dennison, H.J. Gilbody, S.L. Duggleby, C. Cooper, D.J. Barker, and D.I. Phillips. Birth weight, weight at 1 y of age, and body composition in older men: findings from the Hertfordshire Cohort Study. Am J Clin Nutr 2004. 80: 199-203.

Smith, G. and M. Egger, Going beyond the grand mean: subgroup analysis in meta-analysis of randomised trials, in Systematic Reviews in Health Care: Meta-analysis in context, M. Egger, G. Smith, and D. Altman, Editors. 2001, BMJ Publishing Group: London. p. 143-156.

Snellen, J., Body temperature during exercise. Med Sci Sports Exerc, 1969. 1: 36-44.

Soultanakis, H.N., R. Artal, and R.A. Wiswell, Prolonged exercise in pregnancy: glucose homeostasis, ventilatory and cardiovascular responses. Semin Perinatol, 1996. 20: 315-327.

Tafari, N., R.L. Naeye, and A. Gobezie, Effects of maternal undernutrition and heavy physical work during pregnancy on birth weight. Br J Obstet Gynaecol, 1980. 87: 222-226.

Thompson, S., Why and how sources of heterogeneity should be investigated, in Systematic Reviews in Health Care: Meta-analysis in context, M. Egger, G. Smith, and D. Altman, Editors. 2001, BMJ Publishing Group: London.

US Department of Health and Human Services, Health promotion and disease prevention United States, 1985., in DHHS Publication No. 88-1591, National Center for Health Statistics. 1988, DHHS: Hyattsville, MD.

Wong, S.C. and D.C. McKenzie, Cardiorespiratory fitness during pregnancy and its effect on outcome. Int J Sports Med, 1987. 8: 79-83.

Woodward, S., How Does Strenuous Maternal Exercise Affect The Fetus? A Review. Birth and the Family Journal, 1981. 8: 17-24.

Zaidise, I., R. Mittelmark, and S. Bessman, Fuel Metabolism in Pregnancy - Theoretical Aspects, in Exercise in Pregnancy, R. Milttlemark Artal, R. Wiswell, and B. Drinkwater, Editors. 1991, Williams & Wilkins. p. 31-44.