# Exposure analysis methods impact associations between maternal physical activity and cesarean delivery

Marit L. Bovbjerg, Anna Maria Siega-Riz, Kelly R. Evenson, William Goodnight

Key Words: exercise; childbirth; cohort study; prospective; exposure coding

Bovbjerg (marit.bovbjerg@oregonstate.edu) is with the College of Public Health and Human Sciences, Oregon State University, Corvallis, OR. Siega-Riz and Evenson are with the Dept of Epidemiology, University of North Carolina, Chapel Hill, NC. Goodnight is with the Dept of Obstetrics and Gynecology, University of North Carolina, Chapel Hill, NC.

**Abstract**

*Background:* Previous studies report conflicting results regarding a possible association between maternal physical activity (PA) and cesarean delivery. *Methods:* 7-day PA recalls were collected by telephone from n=1205 pregnant women from North Carolina, without prior cesarean, during two time windows: 17-22 weeks and 27-30 weeks completed gestation. PA was treated as a continuous, non-linear variable in binomial regressions (log-link function); models controlled for primiparity, maternal contraindications to exercise, pre-eclampsia, pre-gravid BMI, and percent poverty. We examined both total PA and moderate-to-vigorous PA (MVPA) at each time. Outcomes data came from medical records. *Results:* The dose-response curves between PA or MVPA and cesarean risk at 17-22 weeks followed an inverse J-shape, but at 27-30 weeks the curves reversed and were J-shaped. However, only (total) PA at 27-30 weeks was strongly associated with cesarean risk; this association was attenuated when women reporting large volumes of PA (>97.5$^{th}$ percentile) were excluded. *Conclusion:* We did not find evidence of an association between physical activity and cesarean birth. We did, however, find evidence that associations between PA and risk of cesarean may be non-linear and dependent on gestational age at time of exposure, limiting the accuracy of analyses that collapse maternal PA into categories.

Cesarean delivery rates have risen dramatically in the US over the last few decades, and are currently nearly 33%.[1,2] Cesareans, though potentially life-saving procedures, are nonetheless not risk-free; most stakeholders agree that the US rate is substantially higher than optimal based on the risk:benefit ratio.[3–5] Interventions which reduce the cesarean rate could improve both neonatal and maternal outcomes as well as help to control health care costs.[6–8]

One proposed intervention has been physical activity (PA) during pregnancy, because theoretically an active woman's body might be better able to withstand the rigors of labor and birth.[9] Twenty-four previous studies have examined the association between PA or exercise during pregnancy and risk or odds of cesarean.[10–33] Reported effect estimates are not consistent across studies, with the slightly more than half reporting a decreased risk[19–32] of cesarean with higher levels of PA or exercise, but with a sizeable minority reporting no effect[10,12–14], an increased risk[15–17,19], decreased risk in one subgroup only[11], or decreased risk of elective/planned cesareans but increased risk of urgent/emergent surgeries.[18]

Several methodological issues arise when examining the body of work on this issue, as has similarly been observed in other studies of PA during pregnancy.[34] These methodological limitations include small samples, inconsistent exposure definitions, incomplete or simplistic exposure ascertainment, questionable generalizeablility, and inadequate statistical methods. For instance, among the 24 studies discussed here, only four conducted multivariable analysis[11,22,30,32], half had sample sizes of ≤100[10,12,14,15,20,21,24,28–31], and all treated the PA exposure variable as categorical, rather than continuous, as is preferred with data that are theoretically continuous.[35–37]

Additionally, for many intrauterine exposures (e.g., teratogens), timing is critical[38,39]; it is certainly possible that PA might affect pregnancy outcome differentially depending on gestational age when the exposure took place. Previously, our findings using data from the Pregnancy Risk Assessment Monitoring System (PRAMS) indicated that reporting more bouts of PA was associated with reduced risk of cesarean among women who delivered preterm, but not among those who delivered after 37 weeks.[11] However, in that study we could not discern whether the important facet of exposure was gestational age at the time of the reported PA exposure, or gestational age at birth: the PRAMS questionnaire asks about PA during the last 3 months of pregnancy, so for women delivering preterm this period falls earlier in gestation than for women delivering at term. Nonetheless, this preliminary study adds some weight to the possibility that controlling for gestational age at time of exposure might be important when considering maternal physical activity and birth outcomes.

The current study had two objectives. The first was to explore the associations between maternal PA and cesarean risk, using methods that, though relatively commonplace in epidemiology and clinical research, have not yet been applied to maternal physical activity: specifically, to use a continuous exposure variable, to pay particular attention to the shape of a possible dose-response curve, and to assess the effects of timing of PA (in relation to gestational age) on the estimated measure of effect.

The second objective for this study was to conduct a rigorous multivariable analysis, using methods as determined by the first objective (i.e., perhaps dose-response associations are linear, in which case non-linear model terms would not be necessary). Because of the complexity of any causal model postulating an effect of PA on cesarean risk, and the highly-skewed nature of the exposure data, we also included a series of sensitivity analyses to assess robustness of the results.

**Methods**

The study objectives were addressed by merging two sources of data. The first the third Pregnancy, Infection, and Nutrition (PIN3) cohort, an ongoing study of pregnancy in central North Carolina that provided detailed PA exposure data as well as data on some covariables. The PIN3 Study recruited women between January 2001 and June 2005, from prenatal clinics affiliated with the University of North Carolina (UNC) Hospitals. Women were eligible if they presented for antenatal care before 20 weeks completed gestation, intended to deliver at a UNC hospital, were carrying a singleton fetus, were ≥16 years old, read and spoke English, and had access to a telephone. Details about the data collection protocols can be found at the PIN3 website (http://www.cpc.unc.edu/pin/design_pin3.html).

The PIN3 Study collected 7-day PA recalls by telephone interview during two time windows: 17-22 and 27-30 weeks completed gestation. These detailed interviews included information about occupational, recreational, indoor and outdoor household, care giving, and transportation physical activities during the immediate previous 7 days. Women were asked, for each domain, to list any specific activities, the frequency and average duration for each, and to rate the perceived intensity of the activity as "fairly light," "somewhat hard," or "hard or very hard." Expert review of selected taped interviews ensured consistency among interviewers. The entire questionnaire, along with evidence demonstrating reliability and validity in pregnant women, is available elsewhere.[40]

Based on the recall data, values for total hours/week of PA and hours per week of moderate-to-vigorous PA (MVPA—all bouts rated “somewhat hard” or “hard or very hard”) were calculated. These calculations were conducted separately for each recall (17-22 weeks, 27-

30 weeks). PA data were then examined for outliers. Data entry errors were corrected, and unreasonable/impossible values were set to missing if unconfirmed.[a]

The second data source, which provided outcome and co-variable data, was the Perinatal Database maintained by the UNC Hospitals Department of Obstetrics and Gynecology. Data are collected by labor and delivery (L&D) nurses, who review medical records for all admitted women and abstract information on demographics, obstetrical history, prenatal care, comorbidities, assessment on admission to L&D, the course of labor, and any complications arising during L&D. Monthly validity checks allow correction of impossible or inconsistent values.

The outcome for this paper was primary cesarean birth, covering both primary planned cesarean and primary emergent/urgent cesarean. Though we did not address reliability or validity of the outcome for this study, delivery mode is typically accurately and prominently recorded in medical records because of specialized patient care needs, liability concerns, and billing requirements.

These two data sources were merged on mother's medical record number and baby's date of birth. 3203 women were eligible for PIN3 based on patient logs at obstetrics clinics affiliated with UNC; of these 2006 agreed to participate (63%). Of the 2006, 2% became ineligible (4 multiple pregnancies, 43 pregnancy losses), 9% were lost to follow-up (126 did not complete any questionnaires or interviews; 48 asked to be dropped later in the study), and 121 (7%) were participating for the second or third time, leaving 1654 participants. Of these, 1488 (90%) were successfully merged with the Perinatal Database. For this analysis, all women with previous cesarean deliveries (n=282) were excluded because the repeat cesarean rate in the PIN3 Study

[a] One woman, for instance, had been on vacation at a large amusement park for some of the days covered by her recall. The large volume of walking she reported, though unusual, was nonetheless valid.

was over 95%, leaving little room for any possible effects of lifestyle behaviors. Finally, we excluded one woman with un-confirmed extreme PA values, leaving 1205 women. Both this analysis and the PIN3 Study protocols were approved by the Institutional Review Board (IRB) at UNC; this analysis was also approved by the IRB at Oregon State University. PIN3 participants provided written informed consent.

*Covariables*

Women in the PIN3 Study self-reported their race, marital status, education, and household information, including income, number of adults, and number of children living at the home. From these data we calculated the percent of the 2001 poverty level [41(p5)]: a score of 100 indicates a household living exactly at the poverty line.

Women were also asked about previous pregnancies, including both live and stillbirths (after 20 weeks completed gestation), which were combined to define parity. Parity was collapsed into primiparous vs. multiparous, because there is a clear difference in labor pattern and cesarean risk between these two groups, but fewer differences are observed between higher order labors.[42(p121)] Maternal height was measured by study staff; pre-gravid weight was self-reported. Pre-gravid body mass index (BMI) was calculated from these values. Gestational age at birth was estimated using ultrasonography if the test was performed prior to 22 weeks (>90% of the PIN3 sample), and on date of last menstrual period otherwise. Birthweight was abstracted from the medical record.

Information about pregnancy complications came from the Perinatal Database. Complications considered as covariables were a global yes/no "contraindications to exercise during pregnancy" variable [as defined by the American College of Obstetricians and Gynecologists--includes incompetent cervix or cerclage, placenta previa or abruption, and undelivered premature labor[43]] and a global yes/no "severe hypertensive disorders of pregnancy" variable (included pre-eclampsia, eclampsia, and HELLP [hemolysis, elevated liver enzymes, low platelet count] syndrome).

*Data analysis, objective 1*

The first objective was to explore the associations between maternal PA and primary cesarean risk, particularly in regards to the shape of a possible dose-response curve and timing of activity in relation to gestational age. We used 4 different continuous exposure measures for this objective and throughout this paper: hours/week of *total* PA at both 17-22 weeks and 27-30 weeks; and hours/week MVPA at 17-22 weeks and 27-30 weeks. We analyzed both total PA and MVPA because while the current guidelines for exercise during pregnancy[43] explicitly prescribe *moderate* intensity activity, much evidence has surfaced in recent years about the value of light intensity activities accumulated over the course of a day.[44,45]

In unadjusted analyses using binomial regression with a log-link function, we either forced the exposure to be linear in the log risk or allowed it to depart from linearity via restricted cubic splines with 3 knots, placed at quantiles 0.10, 0.50, and 0.90.[36(p23)] Because we had a large sample size, we initially used 5 knots, and then 4, but both of these choices resulted in over-fitting at the lower end of PA where most of the data occurred (data not shown). Restricted cubic splines were chosen for the non-linear terms because they reduce the influence of data in the tails of a distribution, an important consideration with skewed data such as hours/week of physical activity.[36(p20)]

*Data analysis, objective 2*

The second objective was to conduct a multivariable analysis of the association between maternal PA and primary cesarean risk, basing exposure modeling assumptions on results from the first objective. We again used binomial regression with a log link function to account for covariables, which were chosen based on a directed acyclic graph (DAG)-style causal model.[46,47] Covariables thus chosen included percent poverty, contraindications to exercise during pregnancy, severe hypertensive disorders of pregnancy, primiparity, gestational age at time of exposure ascertainment (in days), and pre-gravid BMI. We included gestational age in days to further explore the issue of timing—we have exposure data from two time windows (17-22 weeks and 27-30 weeks); however each of these windows spans several weeks. It could be that PA at 17 weeks is associated with different outcomes than PA at 22 weeks, despite them being in the "same" time window according to the study design.

Models testing physical activity from the 27-30 week time window also included the level of physical activity from 17-22 weeks, to allow for isolation of PA effects at the second time window; these models dropped women who delivered prior to 27 weeks (n=9). Primiparity was initially included as a possible effect modifier because of the large differences between first labor and higher order labors[42(p121)]; however, no evidence of effect modification by parity surfaced for any of the exposures ($p > 0.5$ by analysis of deviance for all) so all interaction terms were dropped in the final analysis. Each of the 4 exposure variables (total PA at 17-22 weeks, total PA at 27-30 weeks, MVPA at 17-22, MVPA at 27-30) was, based on our findings from objective 1, entered into its respective model using a restricted cubic splines with 3 knots, though we anticipated from Objective 1 results that for MVPA exposures, the nonlinear term might not be strictly necessary.

*Sensitivity Analyses*

Because we were testing multiple exposures, on data that are self-reported and severely skewed, and for a causal relationship that would be quite complex, we conducted a set of sensitivity analyses to assess the robustness of our multivariable results. First, we re-ran the four models restricting the exposures to recreational PA only (rather than PA from all modes) at 17-22 weeks and 27-30 weeks. For these analyses using recreational PA as the exposure, we again controlled for percent poverty, contraindications to exercise during pregnancy, severe hypertensive disorders of pregnancy, primiparity, gestational age at time of exposure ascertainment (in days), and pre-gravid BMI we also controlled for PA from all other modes (i.e. total PA minus recreational PA). The rationale for limiting to recreational activity only was that the current American College of Obstetricians and Gynecologists recommendations for PA during pregnancy refer only to this type of activity.[43]

Next, because PA data were severely right-skewed (see data density functions on the X-axes and the vertical gray dashed lines denoting the 90$^{th}$ percentile, Figure 1), we ran a sensitivity analysis in which we excluded the top 2.5% of women for each of the 4 main exposures (i.e., total PA and MVPA, each at both time windows). Using restricted cubic splines helped to limit the influence of data at the extremes[36], but the upper tails in our data were so long that even with the spline terms, we were concerned about undue influence of women reporting large volumes of PA.

We also explored models excluding women who reported no PA or no MVPA. At the 17-22 week recall, 7.1% of women reported zero hours/week of PA, and 34.5% reported zero hours/week of MVPA (9.0% and 36.8%, respectively, at 27-30 weeks). Again, we were concerned about potential undue influence of these participants on the effect estimates. All analyses were conducted using S-Plus version 8.1 for Windows (Tibco Spotfire, Inc., Palo Alto, CA), with the Hmisc and Design libraries enabled.[35,36]

**Results**

Demographics for our sample are shown in Table 1. Women in this study were largely Caucasian, married, and well-educated. Fourteen percent delivered preterm; 10% had a low birthweight baby. Women decreased total volume of PA slightly between 17-22 weeks and 27-30 weeks, and as expected, all physical activity data were severely right-skewed (see also Figure 1). Twenty-four percent had a cesarean birth (lower than the national rate of 32.9%[2] because women having repeat cesareans were excluded).

*Objective 1*

We analyzed the data with PA as a continuous exposure, but assuming linearity in the log risk; we then allowed the exposures to depart from linearity. These unadjusted results are shown together, with the linear effect estimate superimposed on the non-linear, in Figure 1.

Several trends are evident from this figure. First, PA was highly right-skewed, with the majority of participants reporting levels of PA within a fairly narrow range near the lower end of the spectrum (see data density function, the thin gray solid line at the bottom of each graph). This limits interpretation of these figures at higher levels of PA. Dashed gray vertical lines denote the 90$^{th}$ percentile of exposure; above these lines confidence limits are wide and estimates unstable. Throughout this paper, we therefore restrict our conclusions to women reporting levels of PA below the 90$^{th}$ percentile for any given exposure definition.

Second, for total hours/week of PA both at 17-22 weeks and 27-30 weeks (top two panels in Figure 1), the splined curve differs substantially from the curve estimated by assuming linearity in the log risk, suggesting that a linearity assumption would not be valid in these analyses. However, the linear approximation may be sufficient for exposures in this data set involving MVPA (bottom two panels).

Third, for both exposures at the 17-22 week time window (total PA, MVPA—left hand column in Figure 1), the association is an inverse J-shape, whereas the trend for exposures at the 27-30 week time window is the opposite. This reversing of direction supports the hypothesis that timing of exposure may be important when considering associations between maternal physical activity and birth outcomes.

Wald $X^2$ test statistic p-values for the unadjusted models shown in Figure 1 were all 0.25 or greater, with the exception of total PA at 27-30 weeks (top right panel, p = 0.027 overall; p = 0.007 non-linear). In unadjusted analyses, then, we did not find evidence of a consistent association between maternal physical activity and risk of cesarean delivery.

*Objective 2*

Graphical results from the final multivariable models for the four main exposures were nearly identical to the graphs presented in Figure 1, though the confidence bands were (as expected) slightly wider (figures not shown). Regression coefficients, standard errors, and test statistics from the final models for the four main exposures are shown in Table 2. Again, we did not find evidence of a consistent effect: the only exposure which was a strong predictor of cesarean risk was total PA at 27-30 weeks, the same single predictor identified in unadjusted analyses. This association of total PA at 27-30 weeks was weak when compared to the associations between the covariables and the outcome (see Table 2).

Two further results from our multivariable results are evident from Table 2. First, while large-scale timing of PA appears to be important (i.e., dose-response curve shapes again reversed between 17-22 weeks and 27-30 weeks, as in Figure 1), in no case did gestational age in days (i.e., precisely *when* during the 17-22 week window was the time 1 exposure assessed) add substantially to the fit of the model. Second, as suggested by results from Objective 1, for the two MVPA exposures the non-linear spline terms were unnecessary.

*Sensitivity Analyses*

First, we restricted the exposures to recreational PA only, controlling for all previous covariables plus PA from all other modes. These curves did not reverse direction at the 27-30 week time window when compared to the 17-22 week time point, nor did nonlinear terms add substantially to the model fit for any of the 4 exposures (data not shown). None of the recreational-only PA exposures was associated with cesarean risk.

Next, we dropped women in the upper 2.5% for each of the four main exposures, controlling for co-variables; this completely attenuated any associations between PA and cesarean (see Figure 2). We also dropped women reporting 0 hours/week total activity, or 0 hours/week MVPA. Excluding these women did not change the results, either with or without including the women in the top 2.5% (data not shown).

**Discussion**

Two dozen previous studies have published results regarding PA during pregnancy and cesarean birth[10–33]; however, no consensus has been reached in the literature about the magnitude or even the direction of the association. Our results suggest that some contributing factors to the lack of consensus could be use of cut points in the exposure, and lack of attention to gestational age at time of exposure. We also found undue influence exerted on the estimated effect measure by data points in the long right-hand tail (i.e., women reporting large volumes of PA).

To our knowledge, this study is the first on this topic to allow the exposure to be a continuous variable. Categorization schemes by definition do not capture all of the information available from a continuous variable, and can harbor residual confounding if categories are not sufficiently homogenous.[37(pp88–92)] Categorizing a continuous variable—or collecting what should be continuous data via categories in the first place—can therefore adversely affect a study's internal validity [36(p6)] and precision.[37(p244)] Furthermore, if the underlying association is non-linear, choice of cut point(s) will affect the estimated effect measure.[37(pp91–92)]

When comparing PA at mid-pregnancy (17-22 weeks) with PA at the start of the third trimester (27-30 weeks), we found marked differences in the shape and direction of the dose-response curve (Figure 1). Not only does this add further weight to the argument that continuous data should be kept continuous, lest choice of cutpoint drive a study's conclusions, but arguably one also cannot assume linearity in the log-risk (nor, presumably, in the log-odds if logistic models are used). In the top right panel of Figure 1, for instance, the predicted curve when assuming linearity is almost a perfect horizontal line—no effect. Yet the curve estimated when allowing the exposure to depart from linearity shows a clear J-shape. Were this continuous variable to be categorized for analytic purposes, the estimated risk ratios would be highly-dependant on chosen cutpoints. For instance, if the cutpoint chosen were 2 hours/week, then the risk ratio comparing women who reported more than 2 hours per week total PA at 27-30 weeks to those who reported 2 or fewer hours would be 0.81 (95% CL: 0.63, 1.04). However, if the cutpoint chosen were instead 17 hours/week, then the estimated RR would be 1.01 (0.65, 1.56); if the cutpoint were 25 hours/week, 1.23 (0.63, 2.39). One can observe from this example how categorizing a continuous variable, particularly if the variable is not linearly related to the log-risk of the outcome, can lead to a variety of conclusions merely by varying the cutpoint. Given that all 24 previous studies on this topic, including one of our own[11], used categorized exposure data, then these two methodological issues might help to explain the variation observed among published results.

Timing of exposure was an important determinant of the shape of the association between PA and cesarean when all women were included in the analysis (Figure 1). The curve reverses direction when comparing 17-22 weeks vs. 27-30 weeks; however, including exact gestational age in days at time of exposure ascertainment did not contribute substantially to model fit in multivariable analysis (Table 2). Thus, while 20 weeks vs. 30 weeks may be important as far as physiologic effects of PA, effects of gestational age are substantially smaller when considering a shorter time interval such as 27 weeks vs. 30 weeks. This is not necessarily surprising; by mid-pregnancy, major development of the fetus is not progressing as rapidly as in early pregnancy.[48] It could be that exact day of PA would be important for pregnancy outcomes following early exposure (as is the case with most teratogenic exposures); however, given the lifestyle nature of PA as an exposure, it is unlikely (though not impossible) that one woman's PA habits would vary dramatically over the course of a week or two. Her habits might (and much previous work suggests that they would[49–51]), though, vary over the long-term course of her pregnancy, as the major pregnancy-related mechanical and physiological changes occur.

In neither unadjusted nor adjusted analyses did we find evidence of a consistent association between PA and risk of all-cause primary cesarean delivery. We found strong effects for only one of the 4 exposures (total PA at 27-30 weeks, in both unadjusted and adjusted analysis); while this could be a 'true' result, it seems much more likely that it stems from either a type I error or residual confounding since this association did not remain during sensitivity analysis wherein all women reporting volumes of PA in the top 2.5% were dropped. Women who report large volumes of PA likely have other lifestyle characteristics which affect their birth outcomes, pointing to residual (or unmeasured) confounding as the explanation for the significant result seen for total PA at 27-30 weeks when all women are included in the model. On the other hand, there is some small fraction of women who accumulate large volumes of PA during pregnancy; though they are likely different from an "average" pregnant woman, these high-volume women nonetheless exist and should not be categorically excluded from studies of effects of PA on pregnancy. Determining relationships between participants with very high levels of PA and various health outcomes has historically been problematic for scientists[52]; it should come as no surprise that this issue extends into studying PA during pregnancy.

Our study has limitations. First, the PIN3 Study sample was wealthier, better educated, and more likely to be white and married then other US childbearing women; they also by definition received early antenatal care, which potentially limits generalizablility. Second, two of our four exposures included activities reported by the women as feeling "fairly light." However, the 7-day PA recall interview text asked women to report activities that "caused an increase in breathing or heart rate"; therefore, light intensity activities were likely under-reported. If reporting light intensity activities was differential by any predictor of cesarean birth, then confounding could result. Third, we asked about PA during two 7-day windows during pregnancy. To the extent that these two weeks were not representative of participants' usual PA patterns during pregnancy, our results would be affected in unpredictable ways.

Fourth, our exposure data come from self-report; self-reported lifestyle behaviors should always be treated with some degree of skepticism. However, the data collection instrument used was designed specifically for pregnant women, and evidence of reliability and validity in this population is presented elsewhere.[40] Additionally, we used immediate past week 7-day recalls;

generally speaking, short-term recall such as this is better for self-reported physical activity measures.[52,53]

Finally, as did nearly all previous studies, we treated cesarean birth as a dichotomous outcome. Narendren[18] and Magann[16,17] each separated urgent/emergent from planned/elective cesareans, but these are still heterogeneous groups; a pregnant woman might have a cesarean birth for any one of a large number of indications (e.g., umbilical cord prolapse, twins, previous cesarean, fetal distress, etc.). If PA *does* affect cesarean risk, it is unlikely that all such pathways are involved. Lumping all cesareans into one global, all-cause outcome variable could mask a true association, if one exists. Our outcomes data come from medical records, a known limitation of which is that data are selectively recorded to ensure adequate clinical care, without thought to future research projects. Thus, absence of a given condition does not necessarily imply that it was not present, merely that it was not recorded. Such misclassification errors would make results of any "indication for cesarean" analysis somewhat suspect in data sets derived from medical records.

*Conclusion*

In this study we did not find evidence of an overall association between PA during pregnancy and primary, all-cause cesarean birth. It is possible that there could be an association for a subgroup of women, or that PA is acting through one of the many pathways to cesarean (and thus our dichotomous outcome is masking the true association). Our results confirm that for physical activity as an exposure, researchers should employ continuous, non-linear exposure measures and consider gestational age at time of exposure as a covariable.

**Acknowledgements:** The third Pregnancy, Infection, and Nutrition (PIN3) Study and the UNC-OB/GYN perinatal database are joint efforts of many investigators and staff members whose work is gratefully acknowledged. We also thank Andy Olshan, Derek Hales, Ushma Mehta, and Viktor Bovbjerg for their insightful comments on earlier versions of this manuscript.

**Funding Sources:** Funding for this study was provided by National Institutes of Health (NIH: Bethesda, Maryland) / National Institute of Child Health and Human Development (#HD37584, #HD052468-01A2), NIH / National Cancer Institute (#CA109804-01), NIH/National Institute of Diabetes and Digestive and Kidney Diseases (#DK061981-05), and NIH General Clinical Research Center (#RR00046). Funding was also provided by the University of North Carolina at Chapel Hill Department of Obstetrics and Gynecology (UNC-OB/GYN: Chapel Hill, North Carolina). The content is solely the responsibility of the authors and does not necessarily represent the official views of the NIH or of the UNC-OB/GYN.

# References

1. Menacker F, Hamilton BE. Recent trends in cesarean delivery in the United States. *NCHS Data Brief*. 2010;(35):1–8.

2. Hamilton BE, Martin JA, Ventura SJ. Births: Preliminary Data for 2011. *National Vital Statistics Reports*. 2012;61(5):1–20.

3. Baicker K, Buckles KS, Chandra A. Geographic variation in the appropriate use of cesarean delivery. *Health Aff (Millwood)*. 2006;25(5):w355–367.

4. Declercq E, Menacker F, Macdorman M. Maternal risk profiles and the primary cesarean rate in the United States, 1991-2002. *Am J Public Health*. 2006;96(5):867–872.

5. Druzin ML, El-Sayed YY. Cesarean delivery on maternal request: wise use of finite resources? A view from the trenches. *Semin. Perinatol*. 2006;30(5):305–308.

6. Allen VM, O'Connell CM, Farrell SA, Baskett TF. Economic implications of method of delivery. *Am. J. Obstet. Gynecol*. 2005;193(1):192–197.

7. Halliday HL. Elective delivery at "term": implications for the newborn. *Acta Paediatr*. 1999;88(11):1180–1181.

8. MacDorman MF, Declercq E, Menacker F, Malloy MH. Infant and neonatal mortality for primary cesarean and vaginal births to women with "no indicated risk," United States, 1998-2001 birth cohorts. *Birth*. 2006;33(3):175–182.

9. Gaskin IM. *Ina May's Guide to Childbirth*. New York: Bantam Dell; 2003.

10. Botkin C, Driscoll CE. Maternal aerobic exercise: newborn effects. *Fam Pract Res J*. 1991;11(4):387–393.

11. Bovbjerg ML, Siega-Riz AM. Exercise during pregnancy and cesarean delivery: North Carolina PRAMS, 2004-2005. *Birth*. 2009;36(3):200–207.

12. Kardel KR, Kase T. Training in pregnant women: effects on fetal development and birth. *Am. J. Obstet. Gynecol*. 1998;178(2):280–286.

13. Kulpa PJ, White BM, Visscher R. Aerobic exercise in pregnancy. *Am. J. Obstet. Gynecol*. 1987;156(6):1395–1403.

14. Marquez-Sterling S, Perry AC, Kaplan TA, Halberstein RA, Signorile JF. Physical and psychological changes with vigorous exercise in sedentary primigravidae. *Med Sci Sports Exerc*. 2000;32(1):58–62.

15. Dale E, Mullinax KM, Bryan DH. Exercise during pregnancy: effects on the fetus. *Can J Appl Sport Sci*. 1982;7(2):98–103.

16. Magann EF, Evans SF, Newnham JP. Employment, exertion, and pregnancy outcome: assessment by kilocalories expended each day. *Am. J. Obstet. Gynecol*. 1996;175(1):182–187.

17. Magann EF, Evans SF, Weitz B, Newnham J. Antepartum, intrapartum, and neonatal significance of exercise on healthy low-risk pregnant working women. *Obstet Gynecol*. 2002;99(3):466–472.

18. Narendran S, Nagarathna R, Narendran V, Gunasheela S, Nagendra HRR. Efficacy of yoga on pregnancy outcome. *J Altern Complement Med*. 2005;11(2):237–244.

19. Zeanah M, Schlosser SP. Adherence to ACOG guidelines on exercise during pregnancy: effect on pregnancy outcome. *J Obstet Gynecol Neonatal Nurs*. 1993;22(4):329–335.

20. Baciuk EP, Pereira RI, Cecatti JG, Braga AF, Cavalcante SR. Water aerobics in pregnancy: Cardiovascular response, labor and neonatal outcomes. *Reprod Health*. 2008;5:10.

21. Beckmann CR, Beckmann CA. Effect of a structured antepartum exercise program on pregnancy and labor outcome in primiparas. *J Reprod Med*. 1990;35(7):704–709.

22. Bungum TJ, Peaslee DL, Jackson AW, Perez MA. Exercise during pregnancy and type of delivery in nulliparae. *J Obstet Gynecol Neonatal Nurs*. 2000;29(3):258–264.

23. Clapp JF. The course of labor after endurance exercise during pregnancy. *Am. J. Obstet. Gynecol*. 1990;163(6 Pt 1):1799–1805.

24. Collings CA, Curet LB, Mullin JP. Maternal and fetal responses to a maternal aerobic exercise program. *Am. J. Obstet. Gynecol*. 1983;145(6):702–707.

25. Erdelyi GJ. Gynecological survey of female athletes. *Journal of Sports Medicine and Physical Fitness*. 2:174–179.

26. Hall DC, Kaufmann DA. Effects of aerobic and strength conditioning on pregnancy outcomes. *Am. J. Obstet. Gynecol*. 1987;157(5):1199–1203.

27. Horns PN, Ratcliffe LP, Leggett JC, Swanson MS. Pregnancy outcomes among active and sedentary primiparous women. *J Obstet Gynecol Neonatal Nurs*. 1996;25(1):49–54.

28. Jarrett JC, Spellacy WN. Jogging during pregnancy: an improved outcome? *Obstet Gynecol*. 1983;61(6):705–709.

29. Lynch AM, McDonald S, Magann EF, et al. Effectiveness and safety of a structured swimming program in previously sedentary women during pregnancy. *J. Matern. Fetal. Neonatal. Med*. 2003;14(3):163–169.

30. Melzer K, Schutz Y, Soehnchen N, et al. Effects of recommended levels of physical activity on pregnancy outcomes. *Am. J. Obstet. Gynecol*. 2010;202(3):266.e1–6.

31. Pomerance JJ, Gluck L, Lynch VA. Physical fitness in pregnancy: its effect on pregnancy outcome. *Am. J. Obstet. Gynecol*. 1974;119(7):867–876.

32. Dumith SC, Domingues MR, Mendoza-Sassi RA, Cesar JA. Physical activity during pregnancy and its association with maternal and child health indicators. *Rev Saude Publica*. 2012;46(2):327–333.

33. Barakat R, Pelaez M, Lopez C, Montejo R, Coteron J. Exercise during pregnancy reduces the rate of cesarean and instrumental deliveries: results of a randomized controlled trial. *Journal of Maternal-Fetal and Neonatal Medicine*. 2012;25(11):2372–2376.

34. Gavard JA, Artal R. Effect of exercise on pregnancy outcome. *Clin Obstet Gynecol*. 2008;51(2):467–480.

35. Azola C, Harrell F. *An Introduction to S and the Hmisc and Design Libraries*. 2006. Available at: http://biostat.mc.vanderbilt.edu/twiki/pub/Main/RS/sintro.pdf. Accessed October 11, 2010.

36. Harrell FEJ. *Regression Modeling Strategies, with Applications to Linear Models, Logistic Regression, and Survival Analysis*. New York: Springer; 2001.

37. Selvin S. *Statistical Analysis of Epidemiologic Data*. 3rd ed. Oxford, England: Oxford University Press; 2004.

38. Conover E. Hazardous exposures during pregnancy. *J Obstet Gynecol Neonatal Nurs*. 1994;23(6):524–532.

39. Dencker L, Eriksson P. Susceptibility in utero and upon neonatal exposure. *Food Addit Contam*. 1998;15 Suppl:37–43.

40. Evenson KR, Wen F. Measuring physical activity among pregnant women using a structured one-week recall questionnaire: evidence for validity and reliability. *Int J Behav Nutr Phys Act*. 2010;7:21.

41. Proctor BD, Dalaker J. *Poverty in the United States: 2001*. Washington, D.C.: US Government Printing Office; 2002.

42. Oxorn H. *Human Labor & Birth*. 5th ed. New York: McGraw-Hill; 1986.

43. Anon. ACOG committee opinion. Exercise during pregnancy and the postpartum period. Number 267, January 2002. American College of Obstetricians and Gynecologists. *Int J Gynaecol Obstet*. 2002;77(1):79–81.

44. Powell KE, Paluch AE, Blair SN. Physical activity for health: What kind? How much? How intense? On top of what? *Annu Rev Public Health*. 2011;32:349–365.

45. Woodcock J, Franco OH, Orsini N, Roberts I. Non-vigorous physical activity and all-cause mortality: systematic review and meta-analysis of cohort studies. *Int J Epidemiol*. 2011;40(1):121–138.

46. Greenland S, Pearl J, Robins JM. Causal diagrams for epidemiologic research. *Epidemiology*. 1999;10(1):37–48.

47. Shrier I, Platt RW. Reducing bias through directed acyclic graphs. *BMC Med Res Methodol*. 2008;8:70.

48. Blackburn ST. *Maternal, Fetal, & Neonatal Physiology: A Clinical Perspective*. second. St. Louis, MO: Saunders; 2003.

49. Borodulin K, Evenson KR, Herring AH. Physical activity patterns during pregnancy through postpartum. *BMC Womens Health*. 2009;9:32.

50. Borodulin KM, Evenson KR, Wen F, Herring AH, Benson AM. Physical activity patterns during pregnancy. *Med Sci Sports Exerc*. 2008;40(11):1901–1908.

51. Evenson KR. Towards an Understanding of Change in Physical Activity from Pregnancy Through Postpartum. *Psychol Sport Exerc*. 2011;12(1):36–45.

52. Haskell WL. Physical activity by self-report: a brief history and future issues. *J Phys Act Health*. 2012;9 Suppl 1:S5–10.

53. Matthews CE, Moore SC, George SM, Sampson J, Bowles HR. Improving self-reports of active and sedentary behaviors in large epidemiologic studies. *Exerc Sport Sci Rev*. 2012;40(3):118–126.