Send correspondence to:
Timothy B. Gage
Department of Anthropology, AS 114
University at Albany – SUNY
Albany, NY 12222
Office 518-442-4704
Fax 518-442-5710

A Birth Weight by Gestational Age View of the Pediatric Paradox

Timothy B. Gage Ph.D.[1]
Fu Fang M.S.[2]
Howard Stratton Ph.D. [2]

Institutional affiliations: 1. Department of Anthropology and Department of Epidemiology, University at Albany-SUNY, and Department of Genetics, Southwest Foundation for Biomedical Research. 2) Department of Biometry and Statistics, University at Albany-SUNY.

Abstract

The pediatric paradox, lower African American infant mortality at low birth weights, is traditionally viewed from the marginal distribution of birth weight. Similar findings have not been reported based on the marginal distribution of gestational age. This paper examines the pediatric paradox in terms of birth weight by gestational age, using covariate density defined mixtures of logistic regressions, fitted to New York State birth cohorts 1985-88. The results indicate that the pediatric paradox is isolated in a subpopulation of births accounting for less then 20% of the total birth cohort. The phenomenon is not limited to low birth weight births as is usually reported. It also occurs at normal birth weights with low gestational ages. Finally, the results are consistent with the hypotheses that higher fetal losses may be responsible for the pediatric paradox and African American/European American infant mortality differentials are underestimated.

Introduction

It is commonly observed that low birth weight African American births tend to have better survivorship then low birth weight European American births (North and MacDonald 1977; Buehler, Kleinman et al. 1987; Johnson 1987; Sappenfield, Buehler et al. 1987; Wilcox and Russell 1990; Gage, Bauer et al. 2004; Platt, Ananth et al. 2004; Gage, Bauer et al. submitted). This phenomenon has been called the "pediatric paradox" based on the assumption that aggressive medical intervention is likely to be most effective for low birth weight infants, and that minority populations are less likely to have access to these medical services (Wilcox and Russell 1990). The recent development of surfactant treatment for respiratory distress syndrome is a case in point. It has proven more effective at reducing mortality among low birth weight European American births to a greater extent then among low birth weight African American births reducing or even reversing the traditional advantage of low birth weight African Americans with respect to respiratory distress syndrome. (Frisbie, Song et al. 2004). It is possible that this treatment has reduced the relative advantage of low birth weight African American infants due to all causes, however, examination of post surfactant treatment birth cohorts still clearly display the pediatric paradox (Platt, Ananth et al. 2004; Gage, Bauer et al. submitted). Thus the paradox remains and grows increasingly counter intuitive.

The paradox is not limited to African versus European American comparisons. It has been widely documented with respect to birth weight in other minority populations e.g. Hispanic Americans (Frisbie, Forbes et al. 1996), Asian Americans (Yip, Li et al. 1991), as well as, a variety of "stressed" versus "non-stressed" populations, smokers versus non-smokers, twin births versus singleton births, low versus high altitude births, first born versus higher order births, and low versus high social class (Buekens and Wilcox 1993). The ubiquity of the pediatric paradox among "stressed" versus "non-stressed" populations

suggests that the phenomenon may be a general result of stress. Gage and colleagues (Gage, Bauer et al. 2004; Gage, Bauer et al. submitted) have hypothesized that the paradox might be due to differential fetal loss, that is, that higher fetal loss in a "compromised" subpopulation selects an otherwise more robust birth cohort. The same mechanism could explain why surfactant treatment is more effective among European then African American infants, that is, the African American infants that would be helped by surfactant treatment may not survive to live birth. There is some evidence that fetal loss is higher in African Americans (Kallan 1993; Buck, Shelton et al. 1995). However, fetal loss is difficult to measure accurately. Despite the attention paid to the pediatric paradox over the years, the ultimate cause of the pediatric paradox remains unexplained.

Most recent studies of infant mortality condition births on gestational age as well as birth weight. While gestational age clearly influences birth weight, gestational age also appears to have an independent effect on mortality (Lubchenco and Koops 1987; Wilcox and Skjoerven 1992; Corry 1997). Interestingly, however, a pediatric paradox with respect to gestational ages is seldom reported (see (Platt, Ananth et al. 2004) for a rare example). No paper has reported on the pediatric paradox while modeling births by both birth weight and gestational age as continuous variables**.** The aim of this paper is to model the African versus European American pediatric paradox with respect birth weight and gestational age using covariate density defined (CDD) mixtures of logistic regression. This technique, a) differs from previous analyses in that birth weight and gestational age are modeled on a continuous scale and b) is capable of controlling for unobserved heterogeneity in infant mortality.

Data and Methods

The data used in this analysis consists of African and European American female and male birth cohorts from New York State, 1985-1988. Multiple births, births without matching parental racial identification, and births missing birth weight or gestational age were eliminated from the analysis. Gestational ages reported in the New York State data are truncated at 49 completed weeks. The characteristics of these birth cohorts are summarized in Table 1.

Table 1 about here

The model employed is a population based Covariate Density Defined (CDD) multivariate mixture of logistic regressions. This is a multivariate extension of the model proposed by Gage (Gage 2002)and Gage et al (Gage, Bauer et al. 2004; Gage, Bauer et al. 2004; Gage, Bauer et al. submitted). It differs substantially from the standard finite mixtures of logistic regressions (Wang 1994; McLachlan and Peel 2000), and from the generalized growth mixture regression models of Muthen (Muthen 2004) and others, because the CDD mixture is parametrically specified. The CDD multivariate Gaussian mixture of logistic regressions for the two subpopulation cases is defined as the joint density of the covariate vector (X), birth weight (x_1) by gestational age (x_2), and the occurrence of death (y):

$$f(X^t = [x_1, x_2], y; \beta, \theta) = f(y \mid X; \beta, \theta) \times f(X; \theta) \qquad 1$$

The birth weight by gestational age density, $f(X;\theta)$, is given by:

$$f(X;\theta = (\theta^{(1)}, \theta^{(2)}, \pi)) = \pi \times N(X, \theta^{(1)} = (\mu_1, \Sigma_1)) + (1-\pi) \times N(X; \theta^{(2)} = (\mu_2, \Sigma_2)) \qquad 2$$

with π, the mixing proportion, defined as the proportion of births belonging to the multivariate Gaussian density accounting for the larger proportion of the birth cohort. This subpopulation is labeled 1 and called the primary (P) subpopulation. The remaining subpopulation is labled 2 and referred to as the secondary (S) subpopulation. For the *i*=1 to 2 subpopulations, $N(X;\theta^{(i)} = (\mu_i, \Sigma_i))$is a multivariate Gaussian density function truncated at [0.0, 0.0] with mean vector

$\mu_i^t = [\mu_{i1}, \mu_{i2}]$ and variance-covariance matrix $\Sigma_i = \begin{bmatrix} \sigma_{i1}^2 & \rho_i \sigma_{i1} \sigma_{i2} \\ \rho_i \sigma_{i1} \sigma_{i2} & \sigma_{i2}^2 \end{bmatrix}$. The covariance terms represent the correlation between gestational age and birth weight. The probability of death conditioned on birth weight by gestational age is:

$$f(y=1 \mid X; \beta = (\beta^{(1)}, \beta^{(2)}), \theta) = q(X;\theta) \times P(X;\beta^{(1)}) + (1 - q(X;\theta)) \times P(X;\beta^{(2)}) \qquad 3$$

In the general case, both birth weight specific infant mortality and gestational age specific infant mortality are assumed to be U shaped. The interaction term between birth weight and gestational age represents the impact on mortality of the combined effect of birth weight and gestational age, over and above the independent effects of birth weight and gestational age. Therefore, an infant of covariate vector *X* in the i^{th} subpopulation has probability of dying given in quadratic logistic form:

$$P(X;\beta^{(i)} = (a_i, B_i^t = [b_{i1}, b_{i2}], C_i = \begin{bmatrix} c_{i1} & \frac{1}{2}cv_i \\ \frac{1}{2}cv_i & c_{i2} \end{bmatrix})) = \frac{e^{a_i + B_i{}^t X + X^t C_i X}}{1 + e^{a_i + B_i{}^t X + X^t C_i X}} \qquad 4$$

Finally $q(X;\theta)$ is the probability that an infant of covariate vector X belongs to subpopulation 1. The birth weight by gestational age density submodel determines that

$$q(X;\theta) = \frac{\pi \times N(X;\theta^{(1)})}{\pi \times N(X;\theta^{(1)}) + (1-\pi) \times N(X;\theta^{(2)})} \qquad 5$$

The mixing proportion, π, has been transformed as $o = \log it(\pi)$ to remove the 0.0 to 1.0 bounds on π. All together there are 23 parameters, 11 defining the mixture, and 12 defining subpopulation-specific mortalities (Table 2).

Table 2 about here

The model is fitted to individual level data by the method of maximum likelihood. In particular, we use the ms() function in Splus to minimize the negative log likelihood of equation 1. Standard hierarchical methods and bias corrected confidence intervals are used to evaluate the significance of each

parameter and identify the parsimonious model. The hierarchical methods use the standard likelihood ratio criterion. Bias corrected confidence intervals are obtained by bootstrap methods. Two bootstraps of 100 repetitions each are generated from the original data and then fitted to the model (Equation 1). The first set of bootstraps is used to estimate the mean bias of the fitting procedure, while the second set of bootstraps defines the confidence interval. For additional statistical details, proof of identification and properties of the model see (Gage, Bauer et al. 2004; Fang, Stratton et al. submitted; Gage, Bauer et al. submitted).

Results

The mixture submodel separates the birth cohort into two subpopulations (Table 3 and Figure 1). These are each represented by the concave (bell-shaped) density surfaces presented in Figure 1a and b. In this case the full mixture submodel is parsimonious (Fang, Stratton et al. submitted) Gage MV. The primary subpopulation accounts for 82% and 84% of the African American female and male birth cohorts and 88% and 87% of the European American female and male birth cohorts. The primary subpopulation includes the majority of births by definition, however this component accounts for a significantly larger proportion of the birth cohort among European American compared to African American births (Table 1). Within each population, the primary means of birth weight and gestational age are higher compared to the secondary means of birth weight and gestational age. These means tend to be higher among European Americans versus African Americans of the same sex. Racial comparisons of the secondary subpopulation indicate similar patterns (Table 3). Within populations, the primary variance is small compared to the secondary variance, which is much larger. Across populations, the primary variances are similar, although the standard deviation of birth weight is slightly smaller and gestational age slightly larger among African Americans. Finally, within

populations the correlation between birth weight and gestational age is lower among primary births then among secondary births. African American births have significantly lower primary correlations but similar secondary correlations compared to European Americans on a sex specific basis (Table 3). As a result, in all populations, the primary subpopulation tends to be concentrated in the normal birth weight by gestational age range, while the secondary subpopulation accounts for most low birth weight, and/or short gestational age births. In addition the secondary subpopulation accounts for most large (macrosomic) births and births with excessively long gestational ages. Thus the secondary subpopulation accounts for all of the categories of birth traditionally considered to be "compromised", i.e. low birth weight, premature, IUGR, post-mature, macrosomic, and SGA. It is important to note, however, that some secondary births occur within the normal birth weight by gestational age range. The secondary subpopulation also accounts for implausibly heavy infants for gestational age, which occur at relatively high birth weights at gestational ages shorter then 30 completed weeks **().** The total birth weight by gestational age density of births is presented in Figure 1c.

Table 3 about here

Figure 1 about here

The mortality results for each subpopulation are presented in Table 4 and Figure 2. The results suggest that mortality is typically U-shaped with respect to both birth weight and gestational age. In most cases the interaction terms are not significant, while in a few cases, the squared terms are not significant and an L-shaped mortality curve is parsimonious (Table 4). The results are convex (bowl-shaped) mortality surfaces (Figure 2). In general, within populations, birth weight by gestational age specific secondary mortality is lower than primary mortality except in the normal birth weight by gestational age range where optimal mortality occurs and secondary exceeds primary mortality except among African American males. Among African American males secondary mortality is

always lower (Table 5). On the other hand, total secondary mortality (crude secondary mortality) is often an order of magnitude higher then total primary mortality (crude primary mortality) (Table 5). This is an example of Simpson's paradox. The total mortality surface (Figure 2c) is not a simple bowl-shape. In fact there is evidence of multiple optimal birth weight-gestational ages (Fang, Stratton et al. submitted). This could be due to the births with misstated gestational ages.

Table 4 about here

Figure 2 about here

Table 5 about here

Comparisons of African versus European mortality surfaces by sex indicate that the pediatric paradox is due entirely to the secondary subpopulation (Figure 3). Primary mortality is significantly higher among African American births compared to European Americans of the same sex except in the low birth weight-normal gestational age range (IUGR range), where there are very few primary births (Figure 3a). Similarly, secondary mortality is significantly lower among African American births throughout the area where most secondary births occur including within the normal birth weight and gestational age range (Figure 3b). Combined primary and secondary mortality (Figure 3c) indicate that within the normal birth weight by gestational age range, African American mortality is higher, but out side the normal birth weight by gestational age range, African American mortality tends to be lower. This suggests that the pediatric paradox includes not only low birth weight births but premature, normal birth weight births, and post mature normal and macrosomic births. The birth weight by gestational age extent of the pediatric paradox has been over looked in previous analyses based only on birth weight.

Figure 3 about here

While the pediatric paradox is visible from the marginal distribution of birth weight (Figure 4), it is not visible from the marginal distribution of

gestational age (Figure 5). Thus the presence of the pediatric paradox among normal birth weight infants with abnormally short and long gestational ages is obscured in both marginal views. The pediatric paradox tends to be strongest (that is African American mortality is lowest relative to European American mortality) at low birth weight at all gestational ages, but particularly at low gestational ages. On the other hand, there is no gestational age where the pediatric paradox occurs at all birth weights. Hence the pediatric paradox is not visible from the marginal view of gestational age. All of these results indicate that the pediatric paradox is due entirely to the secondary subpopulation. In fact these results suggest that the pediatric paradox also occurs among normal birth weight and gestational age secondary births (Figure 3b). This is obscured by the predominance of primary births in this region of the birth weight by gestational age distribution

Figure 4 about here

Figure 5 about here

Discussion.

One limitation of this multivariate CDD mixture of logistic regression model of infant mortality is the use of gestational age. We use mother's last menstrual period (lmp) to estimate gestational age. (Fang, Stratton et al. submitted). Gestational age is more often missing then birth weight (Table 1) and poorly measured (refs). In particular, there are a relatively large number of births at implausibly large birth weights at gestational ages less then 30 completed weeks. These could represent births to mothers mistaken about the date of their last menstrual period due forgetfulness or post conception bleeding etc. One advantage of the CDD mixtures of logistic regression method is that these births are isolated in the secondary subpopulation. In fact they can be identified using the same Gaussian mixture model used as the mixture submodel in CDD mixtures of logistic regression **(**Gage, Bauer et al. submitted**)**

The results presented above are consistent with most other analyses of birth weight using parametric mixture models (Fryer, Hunt et al. 1984; Gage and Therriault 1998; Gage 2000; Gage 2003) and CDD mixtures of logistic regression (Gage 2002; Gage, Bauer et al. 2004; Gage, Bauer et al. 2004; Fang, Stratton et al. submitted; Gage, Bauer et al. submitted). In general, the primary subpopulation accounts for 80% or more of births, and has a relatively high mean birth weight/gestational age, and relatively small variance in birth weight/gestational age. The secondary subpopulation is notable for its very large variance. Consequently, this subpopulation accounts for the majority of births who are classified as "compromised" using traditional methods, low birth weight, premature, IUGR, post-term, macrosomic, and SGA, although it also accounts for births in the normal birth weight/gestational age range. Clinitions have often argued that "compromised" births do occur in the normal birth weight by gestational age range but are understudied since they can not be identified using traditional criterion (). The implausably large births for gestational age may be completely "normal" if the error is due to simple forgetfulness or possibly "compromisded" if the error is due to post conceptual bleeding. As a result it is generally assumed that many if not most secondary births were "compromised" in some way during fetal development (Fryer, Hunt et al. 1984; Gage and Therriault 1998).

The results presented above are also consistent with other estimates of mortality based on CDD mixtures of logistic regressions of infant mortality (Gage 2002; Gage, Bauer et al. 2004; Gage, Bauer et al. 2004; Fang, Stratton et al. submitted; Gage, Bauer et al. submitted). In general, birth weight specific secondary mortality is lower than birth weight specific primary mortality at most birth weights and gestational ages, although secondary mortality is sometimes higher then primary mortality in the normal birth weight by gestational age range. However, crude secondary mortality is higher by an order of magnitude compared to crude primary mortality (Table 5). This is not true of gestational

age alone (Gage unpublished analyses). However, it appears to be true of birth weight by gestational age infant mortality within realistic birth weight and gestational age ranges. One hypothesis for the lower secondary birth weight by gestational age specific infant mortality is differential fetal loss. According to this hypothesis, higher fetal losses among the "compromised" subpopulation would result in a selected secondary cohort at birth, relatively more robust than their non-compromised birth weight and gestational age specific peers in the primary subpopulation.

In any event, the larger secondary subpopulation and generally lower secondary mortality of African American secondary births, compared to European Americans are responsible for the "pediatric paradox". The larger secondary, "compromised" subpopulation is consistent with a population under greater stress. Further if the lower secondary birth weight by gestational age specific infant mortality of African Americans is due to greater fetal loss in the compromised subpopulation, then this lower mortality among compromised births also reflects the population under greater stress. There is some independent evidence that fetal loss is higher among African Americans (Kallan 1993; Buck, Shelton et al. 1995), although fetal loss rates are very difficult to estimate. Finally, the higher mortality of African American primary births is also consistent with the population under greater stress. Thus the proximate cause of the pediatric paradox might be due to heterogeneity in the birth cohort and to differential fetal loss in the disadvantaged population. Under this hypothesis the "paradox" of the pediatric paradox is resolved. The cause of the heterogeneity, and the exogenous determinants that drive this system remain to be identified. If this interpretation is correct, racial health disparities may be considerably underestimated (Wilcox and Russell 1986; Gage, Bauer et al. 2004).

The results, interpretations, and methods presented here suggest a number of limitations of traditional analyses. First, the results suggest that there are a significant number of "compromised" births within the normal birth weight

by gestational age range. These births, assuming they exist, are under recognized and understudied because they can not be identified using the standard classification methods, e.g. low birth weight, short gestational age, or even small for gestational age (SGA) (Lubchenco and Koops 1987; Wilcox and Skjoerven 1992; Corry 1997). Second, the proximate cause of the "pediatric paradox" is entirely due to secondary, "compromised" births. The proximate cause of this low infant mortality may be fetal loss. However, the ultimate causes remain to be identified. However, CDD mixtures of logistic regression can be extended to include additional independent variables, and potential ultimate causes examined (Gage, Bauer et al. submitted). Third, the proximate cause of excess African American infant mortality is entirely due to the primary, "normal" births. The ultimate causes that influence primary mortality are understudied due to the traditional emphasis on studying low birth weight, premature and/or SGA births. These factors creating higher primary infant mortality are not necesisarly the same factors influencing secondary infant mortality. Further the same factors could have different effects in the normal and compromised subpopulations. None of these issues can be explored using traditional statistical techniques, which cannot distinguish between primary and secondary subpopulations.

These issues can, however, be explored using extensions of CDD mixtures of logistic regression, since it distinguishes between the primary and secondary subpopulations and treats them as separate regressions. Covariates representing exogenous factors may also be incorporated into the analysis on either or both subpopulations and may condition the birth weight by gestational age densities or the birth weight by gestational age mortality surfaces (Gage, Bauer et al. submitted). The CDD mixture of logistic regression method, can in theory identify the properties of the under studied, "compromised" births found in the normal birth weight by gestational age range, even it can not discriminate between "compromised" and "normal" individual births over much of the birth

weight/gestational age range. Nevertheless, the characteristics of these births can be statically explored using CDD mixtures of logistic regression. Further, since CDD mixtures of logistic regression treats the two subpopulations as separate regressions, the exogenous factors (additional covariates) influencing the low secondary mortality and the exogenous factors affecting excess primary African American infant mortality can be examined as statistically separate effects. Consequently, the CDD mixtures of logistic regressions methodology could greatly improve our understanding of racial and ethnic disparities in infant mortality.

Conclusions

1. The pediatric paradox is not be limited to low birth weight births. It is characteristics of all birth weight/gestational ages outside the normal range.
2. The pediatric paradox is due to the secondary "compromised" subpopulation and may be driven by fetal loss..
3. Excess African American infant mortality is due to the primary "normal" subpopulation. The role of the primary subpopulation is understudied due to our emphasis on compromised infants.
4. The African American/European American infant mortality differential may be significantly underestimated.
5. Finally, Covariate Density Defined mixtures of logistic regressions can be used to explore these issues.

Acknowledgments: This research was supported by NICHD grant HD37405.

References

Buck, G. M., J. A. Shelton, et al. (1995). "Racial variation in spontaneous fetal fetal deaths at 20 weeks or older in upstate New York, 1980-86." Public Health Rep. **110**: 587-592.

Buehler, J. W., J. C. Kleinman, et al. (1987). "Birth weight-specific infant mortality, United States, 1960 and 1980." Public Health Reports **102**(2): 151-161.

Buekens, P. and A. Wilcox (1993). "Why do small twins have a lower mortality rate than small singletons?" Am. Journal of Obstet Gynecol **168**: 937-941.

Corry, M. (1997). "Does gestational age in combination with birth weight provide better statistical adjustment of neonatal mortality rates than birth weight alone." Paediatric and Perinatal Epidemiology **11**(4): 385-391.

Fang, F., H. Stratton, et al. (submitted). "Infant mortality in heterogeneous populations: An applcation of mixed logistic regressions with covariate density defined components." Social Biology.

Frisbie, W. P., D. Forbes, et al. (1996). "Compromised birth outcomes and infant mortality among racial and ethnic groups." Demography **33**(4): 469-481.

Frisbie, W. P., S. Song, et al. (2004). "The increasing racial disparity in infant mortality." Demography **41**(4): 773-800.

Fryer, J. G., R. G. Hunt, et al. (1984). Biostatistical considerations: The case for using models. Prevention of Perinatal Mortality and Morbidity. F. Falkner. Basel, Karger. **Vol. 3:** 9-30.

Gage, T. B. (2000). "Variability of gestational age distributions by sex and ethnicity; An analysis using mixture models." American Journal of Human Biology **12**: 181-191.

Gage, T. B. (2002). "Birthweight-specific infant and neonatal mortality:The influence of heterogeneity in the birth cohort." Human Biology **74**: 165-184.

Gage, T. B. (2003). "Classification of births by birth weight and gestational age: An application of multivariate mixture models." Annals of Human Biology **30**(5): 589-604.

Gage, T. B., M. J. Bauer, et al. (2004). "Mixed logistic regressions with covariate density defined components: Applications to birth outcomes." CSDA Working Paper Series Working Paper #2004-1.

Gage, T. B., M. J. Bauer, et al. (submitted). "Covariate density defined mixtures of logistic regressions: Applications to birth outcomes." Sociological Methods and Research.

Gage, T. B., M. J. Bauer, et al. (2004). "The pediatric paradox: Heterogeneity in the birth cohort." Human Biology **76**: 327-342.

Gage, T. B. and G. Therriault (1998). "Variability of Birth-Weight Distributions by Sex and Ethnicity: Analysis Using Mixture Models." Human Biology **70**(3): 517-534.

Johnson, J. H. (1987). "U. S. differentials in infant mortality: Why do they persist?" Family Planning Prespectives **19**(5): 227-232.

Kallan, J. E. (1993). "Race intervening variables and two components of low birth weight." Demography **30**: 489-506.

Lubchenco, L. O. and B. L. Koops (1987). Assessment of weight and gestational age. Neonatology. G. B. Avery. Philadelphia, J. P. Lippincott Co.**:** 235-257.

McLachlan, G. J. and D. Peel (2000). Finite Mixture Models. New York, John Wiley and Sons INC.

Muthen, B. (2004). Latent variable analysis: Growth mixture modeling and related techniques for longitudinal data. Handbook of quantitative methodology for the social sciences. D. Kaplan. Newbury Park, CA, Sage Publications.

North, J., A. F. and H. M. MacDonald (1977). "Why are neonatal mortality rates lower in small black infants than in white infants in similar birth weight?" The Journal of Pediatrics **90**(5): 809-810.

Platt, R. W., C. V. Ananth, et al. (2004). "Analysis of neonatal mortality: Is standardizing for relative birth weight biased." BMC Pregnancy and Childhood **4**: 1-9.

Sappenfield, W. M., J. W. Buehler, et al. (1987). "Differences in neonatal and postneonatal mortality by race, birth weight, and gestational age." Public Health Reports **102**(2): 182-192.

Wang, P. (1994). Mixed Regression Models for Discrete Data. Vancouver, University of British Columbia.

Wilcox, A. and I. Russell (1990). "Why small black infants have a lower mortality rate than small white infants: The case for population-specific standards for birth weight." J. of Ped. **116**: 7-10.

Wilcox, A. J. and I. T. Russell (1986). "Birthweight and perinatal mortality: III. Towards a new method of analysis." International Journal of Epidemiology **15**(2): 188-196.

Wilcox, A. J. and R. Skjoerven (1992). "Birth weight and perinatal mortality: The effect of gestational age." American Journal of Public Health **82**(3): 378-382.

Yip, R., Z. Li, et al. (1991). "Race and birth weight: The Chinese example." Pediatrics **87**(5): 688-693.

Table 1 Descriptive statistics for the sample populations

Birth Cohort	# Births	Mean		# Deaths	CDR	Missing Data (%)		
		bwt	gest			bwt	gest	overall
Af. Am. F.	54,968	3119	39.1	690	12.55	0.13	2.11	2.22
Af. Am. M.	57,449	3240	39.1	848	14.76	0.16	2.03	2.19
Eu. Am. F.	255,516	3375	40.0	1294	5.06	0.20	2.04	2.22
Eu. Am. M.	270,189	3507	39.9	1734	6.42	0.21	2.05	2.24

Af. = African, Eu. = European, Am. = American, F. = Females, M. = Males

bwt = Birth weight (grams)

gest = Gestational age (weeks)

CDR = Crude death rate (deaths per 1000 births)

Table 2 Definitions of the model parameters

Symbol	Definition
	Mixture Submodel Parameters --- i^{th} Subpopulation (i = 1 to 2)
π	Mixing proportion (% primary subpopulation)
μ_{i1}	Mean birth weight for the i^{th} subpopulation
μ_{i2}	Mean gestational age for the i^{th} subpopulation
σ_{i1}	Standard deviation of birth weight for the i^{th} subpopulation
σ_{i2}	Standard deviation of gestational age for the i^{th} subpopulation
ρ_i	Correlation between birth weight and gestational age for the i^{th} subpopulation
	Mortality Submodel Parameters --- i^{th} Subpopulation (i = 1 to 2)
	(coefficients of a second degree bivariate polynomial)
a_i	Constant for the i^{th} subpopulation
b_{i1}	Linear term on birth weight for the i^{th} subpopulation
b_{i2}	Linear term on gestational age for the i^{th} subpopulation
c_{i1}	Square term on birth weight for the i^{th} subpopulation
c_{i2}	Square term on gestational age for the i^{th} subpopulation
cv_i	Interaction term on birth weight and gestational age for the i^{th} subpopulation

Table 3 Estimates and 95% confidence intervals for the mixture submodel parameters

Parameter	Estimate	LCI	UCI	Estimate	LCI	UCI
	African American Females			African American Males		
π	0.826	0.836	0.816	0.844	0.852	0.838
μ_{11}	3219	3214	3225	3345	3339	3350
μ_{12}	39.75	39.73	39.76	39.68	39.66	39.70
σ_{11}	450	445	455	462	458	466
σ_{12}	1.71	1.67	1.74	1.77	1.75	1.79
ρ_1	0.247	0.236	0.257	0.256	0.245	0.266
μ_{21}	2653	2621	2682	2691	2657	2718
μ_{22}	36.35	36.13	36.59	36.06	35.84	36.23
σ_{21}	952	929	971	985	967	1008
σ_{22}	5.36	5.26	5.48	5.51	5.41	5.61
ρ_2	0.49	0.47	0.51	0.50	0.48	0.52
	European American Females			European American Males		
π	0.876	0.880	0.872	0.874	0.879	0.869
μ_{11}	3423	3420	3425	3566	3564	3568
μ_{12}	40.19	40.18	40.20	40.08	40.08	40.09
σ_{11}	441	439	443	466	464	467
σ_{12}	1.47	1.46	1.48	1.52	1.50	1.53
ρ_1	0.288	0.284	0.293	0.318	0.314	0.322
μ_{21}	3048	3035	3062	3112	3096	3128
μ_{22}	38.95	38.87	39.03	38.62	38.55	38.71
σ_{21}	837	824	853	872	858	884
σ_{22}	5.17	5.07	5.27	5.22	5.11	5.31
ρ_2	0.469	0.457	0.480	0.498	0.486	0.509

LCI = Lower 95% confidence interval

UCI = Upper 95% confidence interval

Table 4 Estimates and bias-corrected 95% confidence intervals for the mortality submodel parameters

Parameter	Estimate	LCI	UCI	Estimate	LCI	UCI
	African American Females			African American Males		
a_1	6.64E+01	1.32E+00	1.24E+02	7.61E+01	3.90E+01	1.06E+02
b_{11}	2.73E-04	-6.38E-03	7.92E-03	-5.52E-03	-1.09E-02	-4.89E-04
c_{11}	9.68E-07	4.68E-07	1.40E-06	9.43E-07	5.77E-07	1.23E-06
b_{12}	-3.56E+00	-6.21E+00	-5.90E-01	-3.57E+00	-5.07E+00	-1.69E+00
c_{12}	5.23E-02	1.61E-02	8.47E-02	4.64E-02	2.06E-02	6.65E-02
cv_1	-1.89E-04	-3.76E-04	-1.44E-05	-3.81E-05 [&+]	-1.87E-04	1.31E-04
a_2	9.25E+00	6.56E+00	1.17E+01	9.24E+00	6.02E+00	1.32E+01
b_{21}	-3.83E-03	-4.72E-03	-3.08E-03	-3.25E-03	-3.95E-03	-2.47E-03
c_{21}	6.58E-07	5.17E-07	7.92E-07	6.92E-07	5.54E-07	8.00E-07
b_{22}	-4.28E-01	-5.74E-01	-2.62E-01	-3.94E-01	-6.51E-01	-1.86E-01
c_{22}	6.63E-03	4.01E-03	8.90E-03	6.33E-03	3.15E-03	1.01E-02
cv_2	-2.28E-05 [&]	-4.51E-05	3.84E-06	-5.33E-05	-7.73E-05	-2.50E-05
	European American Females			European American Males		
a_1	5.74E+01	-3.51E+00	1.15E+02	2.57E+01	-3.12E+01	8.39E+01
b_{11}	8.17E-04	-5.46E-03	7.66E-03	-1.66E-03	-6.72E-03	2.99E-03
c_{11}	1.28E-06	9.80E-07	1.51E-06	9.50E-07	6.30E-07	1.25E-06
b_{12}	-3.23E+00	-5.99E+00	-4.10E-01	-1.33E+00	-4.24E+00	1.55E+00
c_{12}	5.19E-02	1.67E-02	8.60E-02	2.24E-02 [+]	-1.52E-02	5.98E-02
cv_1	-2.61E-04	-4.30E-04	-9.71E-05	-1.46E-04 [+]	-2.76E-04	-7.94E-07
a_2	9.09E+00	7.62E+00	1.06E+01	8.74E+00	7.39E+00	1.01E+01
b_{21}	-3.04E-03	-3.41E-03	-2.65E-03	-3.50E-03	-3.96E-03	-3.09E-03
c_{21}	3.85E-07	3.19E-07	4.48E-07	3.99E-07	3.31E-07	4.69E-07
b_{22}	-4.19E-01	-5.09E-01	-3.34E-01	-3.38E-01	-4.23E-01	-2.51E-01
c_{22}	5.87E-03	4.50E-03	7.28E-03	4.15E-03	2.61E-03	5.59E-03
cv_2	-1.09E-05 [&+]	-2.45E-05	2.36E-06	1.23E-06 [&+]	-1.33E-05	1.67E-05

LCI = Lower 95% confidence interval

UCI = Upper 95% confidence interval

& = not significantly different from zero based on bias-corrected 95% confidence interval

+ = not significantly different from zero based on hierarchical analysis

Table 5 Optimum Birth weights and Gestational Ages Based on the Full Model

Birth Cohort	Birth Weight (g)		Gestational Age (wk)		Mortality* at Optimum Birth	Crude Death Rate*
	Mean	Optimum	Mean	Optimum		
	Total Birth Cohort					
Af. Am. F.	3119	3839	39.1	40.9	2.35	12.00
Af. Am. M.	3240	3748	39.1	40.0	3.37	14.09
Eu. Am. F.	3375	3826	40.0	40.7	1.12	4.75
Eu. Am. M.	3507	4096	39.9	42.0	1.46	6.02
	Primary Subpopulation					
Af. Am. F.	3219	3869	39.7	41.0	2.31	5.49
Af. Am. M.	3345	3738	39.7	40.0	3.43	6.69
Eu. Am. F.	3423	3825	40.2	40.7	1.08	2.12
Eu. Am. M.	3566	4211	40.1	43.4	1.33	2.64
	Secondary Subpopulation					
Af. Am. F.	2653	3580	36.4	38.5	2.88	43.17
Af. Am. M.	2691	4233	36.1	49.0	0.70	54.63
Eu. Am. F.	3048	4503	38.9	39.9	2.23	23.46
Eu. Am. M.	3112	4329	38.6	40.1	3.58	29.49

Af. = African, Eu. = European, Am. = American, F. = Females, M. = Males

* = deaths per 1000 births

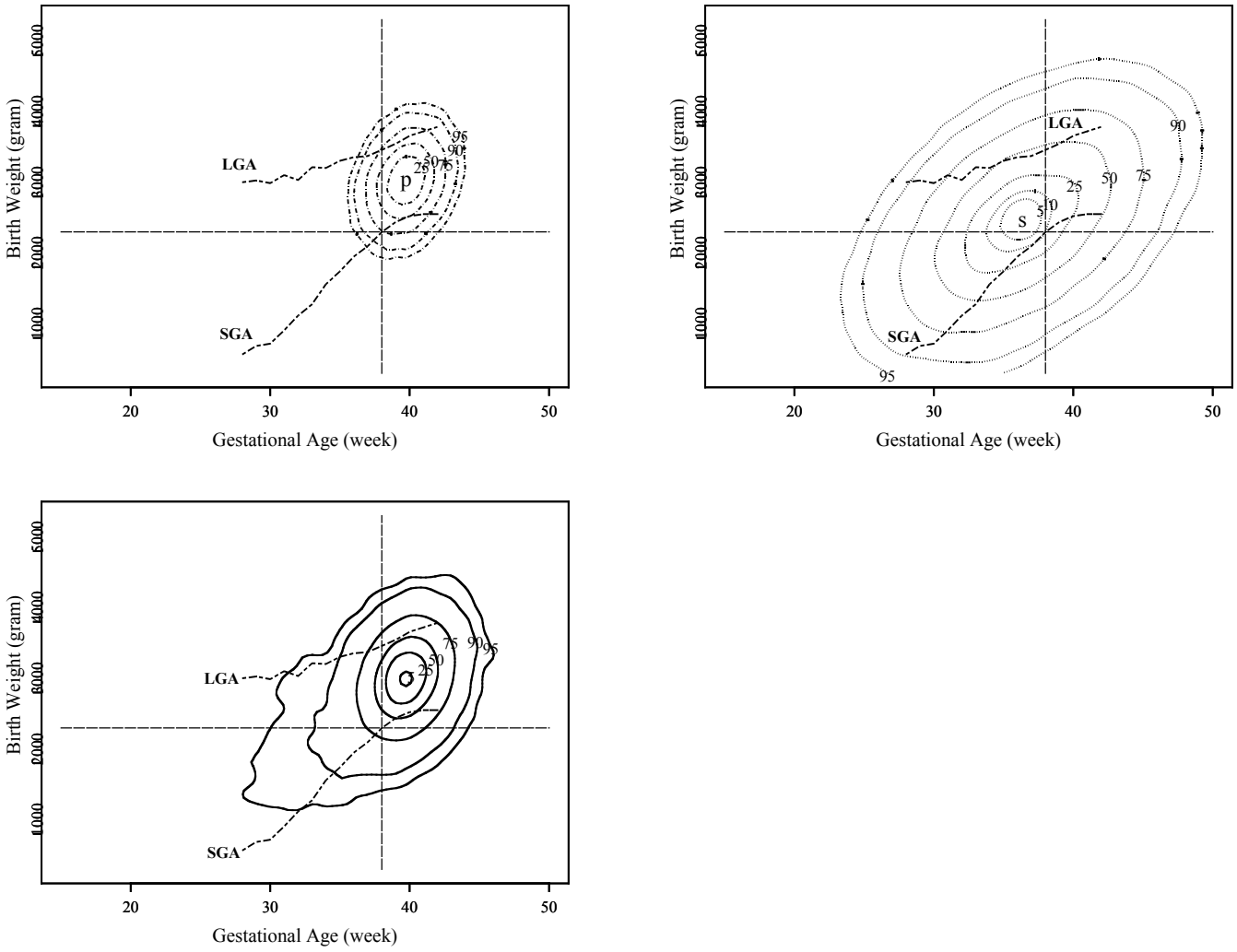

Figure 1. Contour plots of birth weight by gestational age multivariate Gaussian densities for NYS African American females. The plotted values represent the percent of the cohort inside the contour. The growth curves, SGA and LGA, represent the 10th and 90th percentiles based on the observed data, respectively. The cross-hair lines are the traditional cutoffs for low birth weight (<2500 grams) and premature (<38 completed weeks) infants. Panel a is the primary density designated by "p" and dot-dashed contour intervals. Panel b is the secondary distribution designated by "s" and dotted contour intervals. Panel c is a simulated density of the total birth cohort, solid lines.

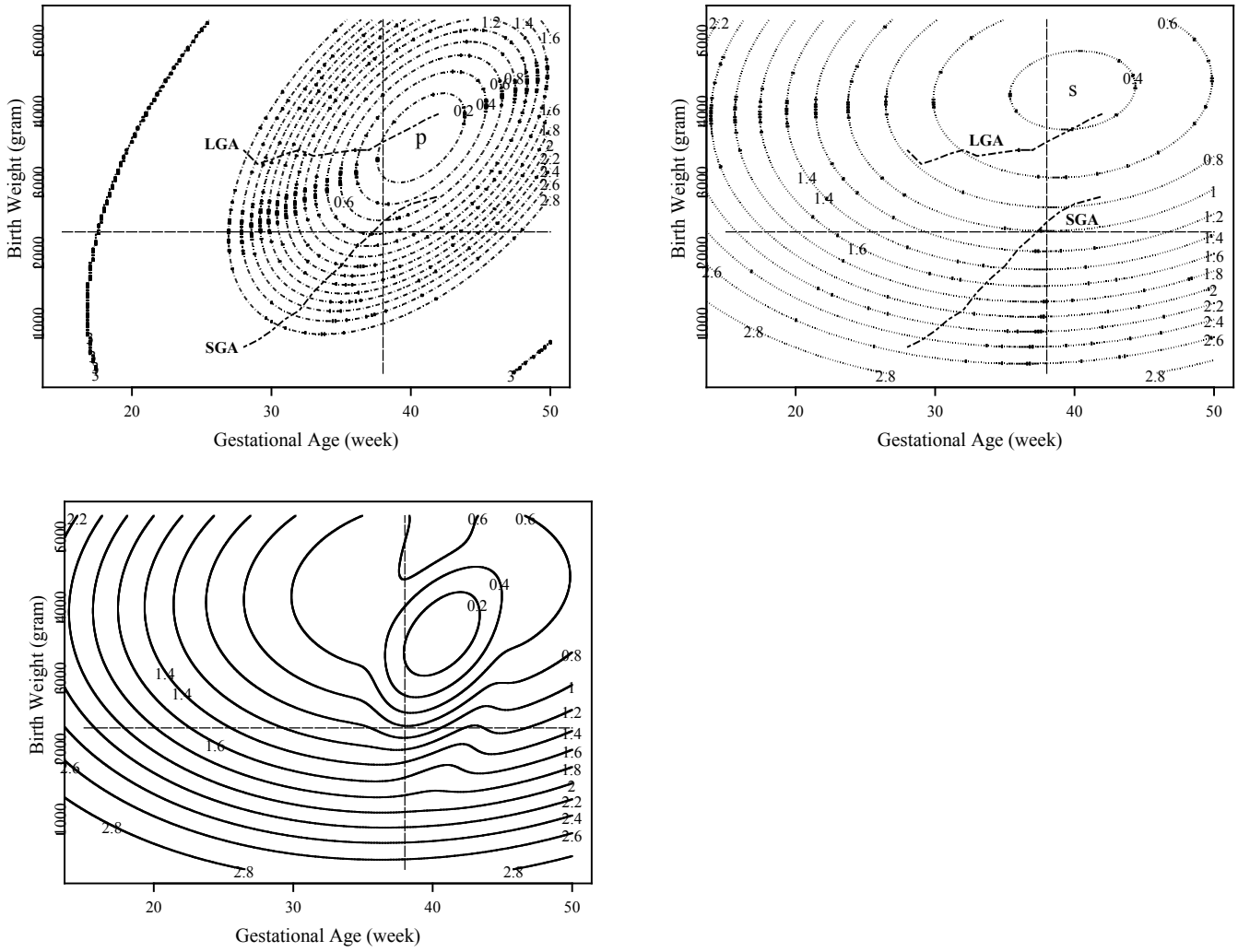

Figure 2. Contour plots of birth weight by gestational age specific mortality (deaths per 1000 births) using a log scale for NYS European American females. Thus a value of 3 indicates 100% mortality. The cross-hair lines are the traditional cutoffs for low birth weight (<2500 grams) and premature (<38 completed weeks) infants. Panel a is he primary mortality surface designated by "p" and dot-dashed contour intervals. Panel b is the secondary mortality surface designated by "s" and dotted contour intervals. Panel c is the model based total mortality surface..

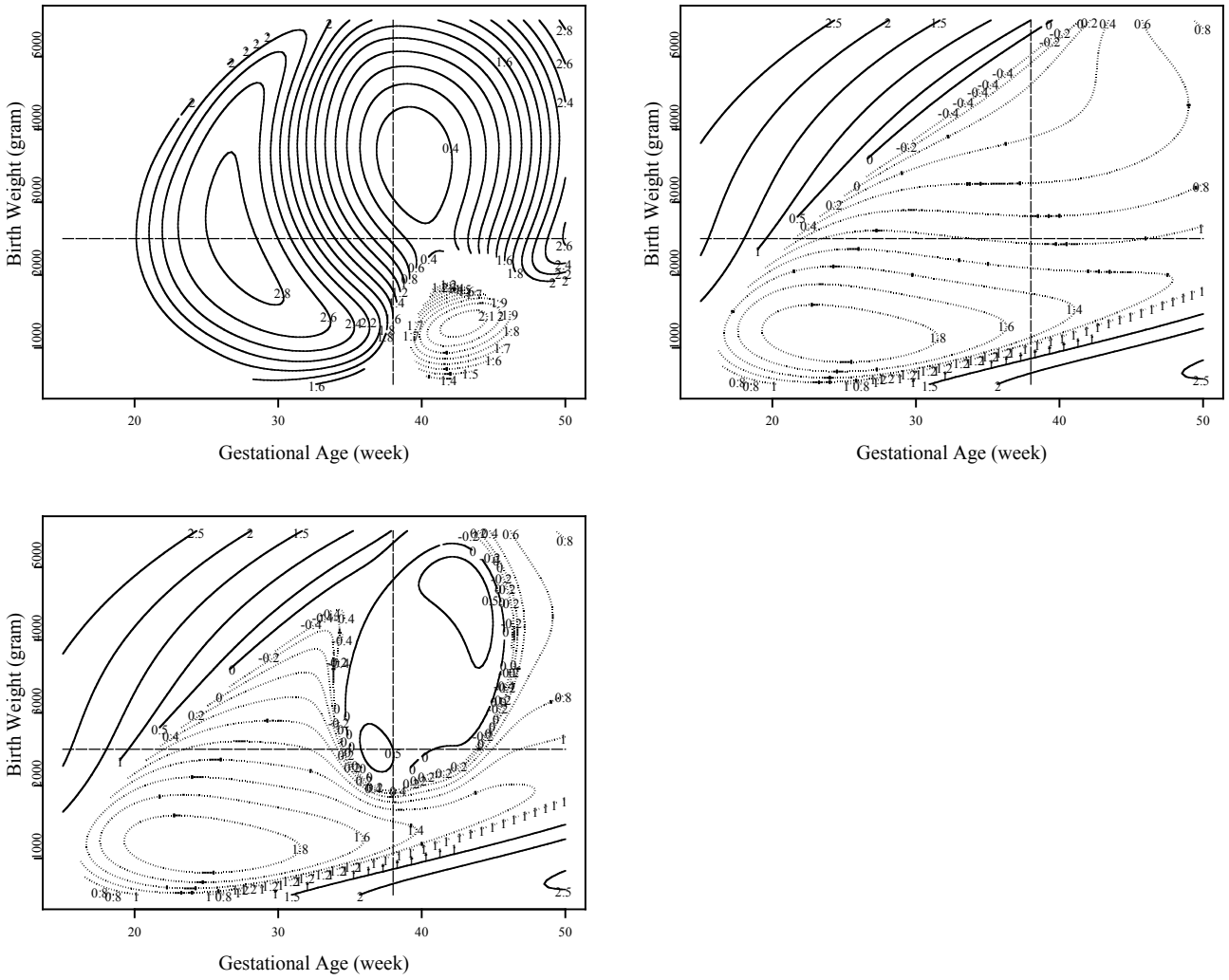

Figure 3. Contour plot of the differences of birth weight by gestational age specific mortality (log deaths per 1000 births) between NYS African and European American males. The cross-hair lines are the traditional cutoffs for low birth weight (<2500 grams) and premature (<38 completed weeks) infants. Panel a compares the primary subpopulations. Panel b compares the secondary subpopulations. Panel c compares total mortalities. Statistically significant differences in mortality are designated by contours intervals. The solid contour intervals represent areas where European American males have lower mortalities than African American males, the dotted contour intervals represent areas where African American males have lower mortalities than European American males.

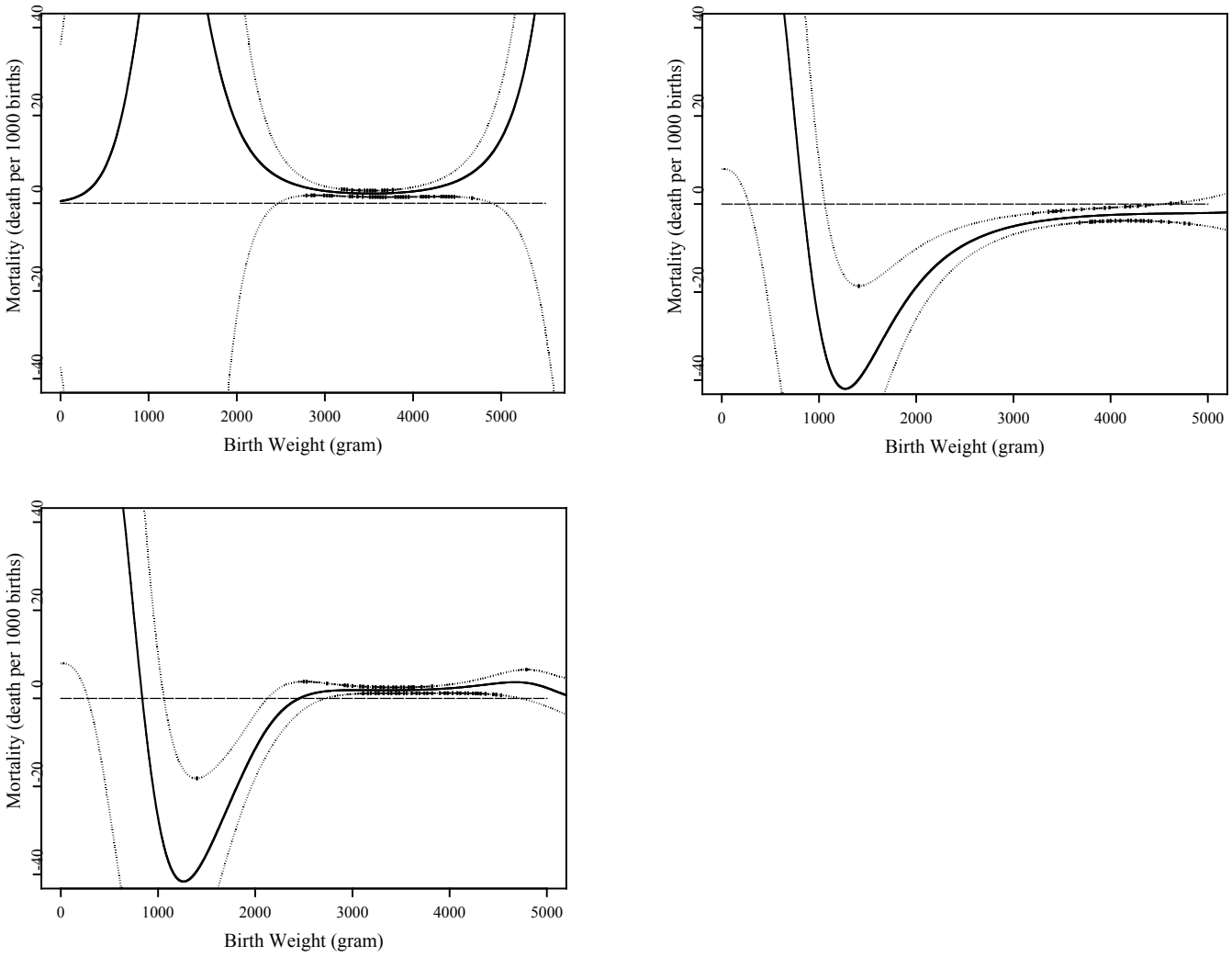

Figure 4. Difference of marginal birth weight specific mortality (deaths per 1000 births) between NYS African and European American males: Panel a primary subpopulation; Panel b secondary subpopulation; Panel c total birth cohort. The dotted lines are the upper and lower 95% confidence intervals of the differential. The cross-hair line, y = 0, represents no mortality difference between these two ethnic groups. Mortality values larger than 0 suggest that African American males have a higher birth weight specific death rate; and mortality values smaller than 0 suggest that European American males have a higher birth weight specific death rate.

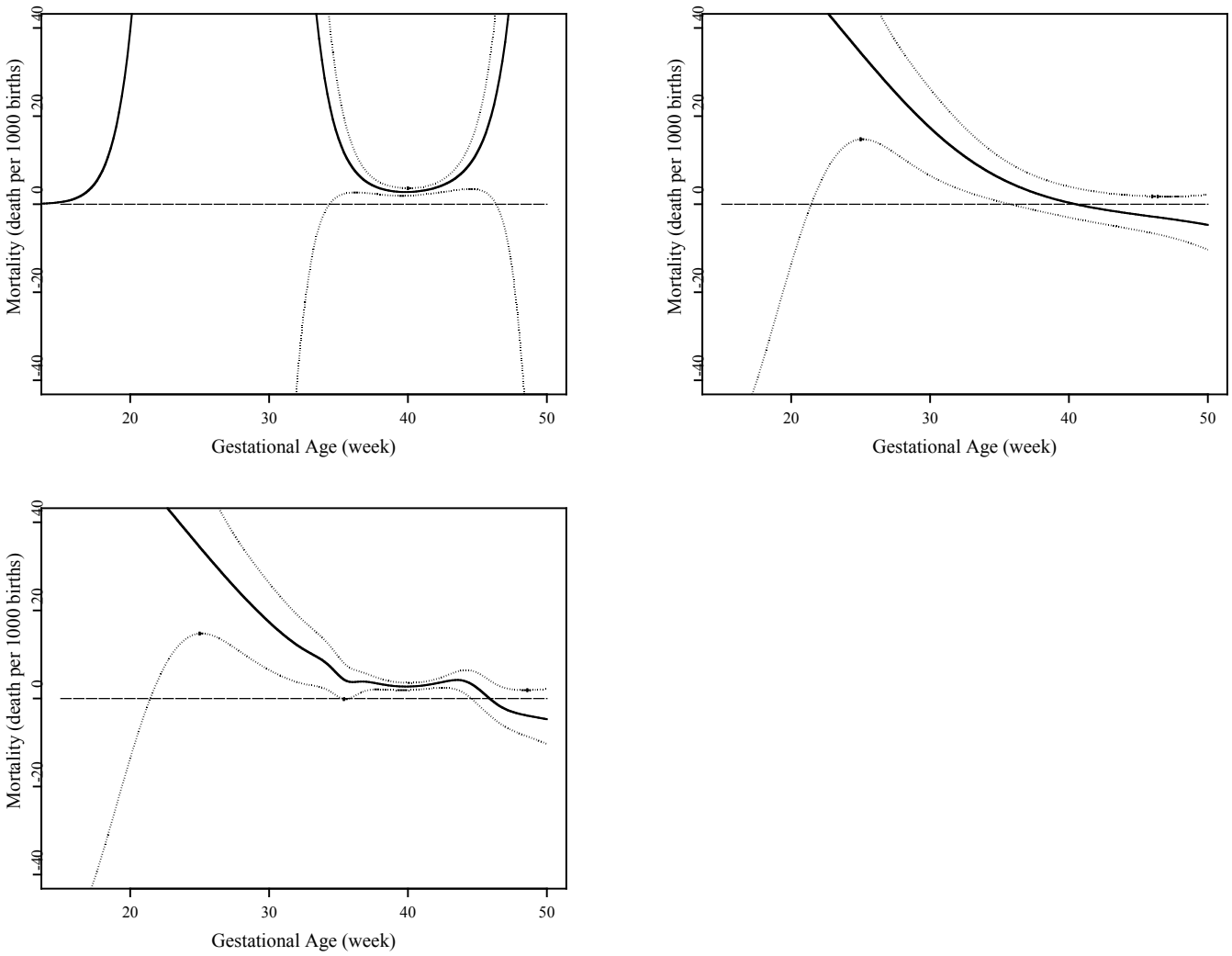

Figure 4. Difference of marginal gestational age specific mortality (deaths per 1000 births) between NYS African and European American males: Panel a primary subpopulation; Panel b secondary subpopulation; Panel c total birth cohort. The dotted lines are the upper and lower 95% confidence intervals of the differential. The cross-hair line, y = 0, represents no mortality difference between these two ethnic groups. Mortality value larger than 0 suggests that African American males have a higher birth weight specific death rate; and mortality value smaller than 0 suggests that European American males have a higher birth weight specific death rate.