A Proportional Hazards Model with Time-dependent Covariates and Time-varying Effects for Analysis of Fetal and Infant Death

Robert W. Platt1,2, K. S. Joseph3, Cande V. Ananth4, Justin Grondines1, Michal Abrahamowicz2, and Michael S. Kramer1,2

1 Department of Pediatrics, Faculty of Medicine, McGill University, Montreal, Quebec, Canada.
2 Department of Epidemiology and Biostatistics, Faculty of Medicine, McGill University, Montreal, Quebec, Canada.
3 Perinatal Epidemiology Research Unit, Department of Obstetrics and Gynecology and Pediatrics, Faculty of Medicine, Dalhousie University, Halifax, Nova Scotia, Canada.
4 Section of Epidemiology and Biostatistics, Department of Obstetrics, Gynecology and Reproductive Sciences, Robert Wood Johnson Medical School, University of Medicine and Dentistry of New Jersey, New Brunswick, NJ.

Received for publication June 30, 2003; accepted for publication November 17, 2003.

Birth-weight- and gestational-age-specific perinatal mortality curves intersect when compared by race and maternal smoking. The authors propose a new measure to replace fetal and infant mortality and an analytic strategy to assess the effects of risk factors on this outcome. They used 1998 data for US Blacks and Whites. Age-specific post–last menstrual period (LMP) mortality rate was defined as the proportion of deaths (stillbirth, perinatal death, or infant death) at a given age post-LMP. The authors used extended Cox regression with time-varying covariates and hazard ratios to model the effects of race and smoking on post-LMP mortality. Perinatal mortality rates (conventional calculation) for Blacks and Whites showed the expected crossover. However, analyses of post-LMP mortality showed no crossover. For the Black-White comparison, a hazard ratio of 1.72 (95% confidence interval: 1.67, 1.77) was obtained. The hazard was higher for smokers than for nonsmokers, but the hazard ratio increased from 1.09 (95% confidence interval: 0.98, 1.22) at 22 weeks to 1.82 (95% confidence interval: 1.72, 1.92) at 40 weeks. The hazard ratio associated with birth was also time dependent: higher than 1 for preterm gestation and lower than 1 for term gestation. The increasing adverse effect of smoking with gestational age suggests an accumulating effect of smoking on mortality. Modeling post-LMP mortality eliminates the crossover paradox for race and maternal smoking in a single statistical model.

birth weight; gestational age; infant mortality; proportional hazards models

Abbreviation: LMP, last menstrual period.

Editor’s note: A related article appears on page 207, two invited commentaries are published on pages 211 and 213, and a response by the authors of the first article to these commentaries is on page 215. In accordance with Journal policy, the author of the second article was asked whether he wanted to respond to these commentaries but chose not to do so.

Over 30 years ago, Yerushalmy et al. (1) identified a paradoxical relation between maternal smoking and birth-weight-specific neonatal mortality. Neonatal death rates for infants of smokers were lower than those for infants of nonsmokers at birth weights of 3,000 g or less; the reverse was true at higher birth weights. In the last three decades, this observation has been corroborated in many studies, including comparisons based on race, infant sex, and country (2–4), as well as other factors. Intersecting neonatal mortality curves present an inferential challenge. The argument that fetuses of women who...
smoke during pregnancy (or of a disadvantaged group such as twins or Blacks) are healthier than fetuses of nonsmokers (or an advantaged group) at some birth weights but not others lacks biologic plausibility and coherence. Sophisticated statistical approaches (5–9) have been proposed to address this phenomenon and include explanations based on “relative birth weight” (5–7) or “relative gestational age” (9). For instance, Wilcox and Russell (7) showed that examining perinatal mortality rates across categories of relative birth weight, that is, birth weight expressed in terms of the population mean and standard deviation, eliminates the crossover paradox.

However, the relative birth weight formulation has been criticized because it fails to distinguish the contributions of birth weight differences due to maturity (i.e., higher gestational age) versus fetal growth (represented by birth weight for gestational age) (10). Although use of relative birth weight (6) or relative gestational age (9) resolves the perinatal mortality crossover, these approaches require that birth weight or gestational age be treated differently from other predictors.

More importantly, relative birth weight or gestational age fails to account for the temporal nature of gestational age. Gestational ages are not exchangeable; that is, an infant born at 40 weeks was at risk of being born at 30 weeks, but an infant born at 30 weeks of gestation is not at risk of being born at 40 weeks. Thus, gestational age should be considered a time axis rather than an independent variable in the model. Research to date on perinatal outcomes has all but ignored the fact that gestational age is a time-to-event variable and has treated gestational age and birth weight as having the same potential causal effect on outcomes.

Perinatal mortality has been criticized as an outcome because it is the combination of two etiologically heterogeneous events, stillbirth and early neonatal death (11, 12). In addition, an arbitrary restriction of early neonatal mortality to the first 7 days after birth seems questionable. It classifies an infant born after 250 days’ gestation and dying at 7 days of life (257 days post–last menstrual period (LMP)) as a perinatal death but an infant born after 249 days’ gestation and dying at 8 days of life (also 257 days post-LMP) as a perinatal survivor. Data on this infant would be part of the denominator but not the numerator when estimating the rate of perinatal mortality.

In this paper, we argue that the temporal nature and lack of exchangeability of gestational age is an important reason to consider it as a time axis in time-to-event analysis rather than as a conventional covariate. Accordingly, we develop a measure of mortality based on time since conception that does not result in the same problems as does perinatal mortality. We expand on a solution to the crossover paradox based on the number of fetuses at risk (12) by using survival analytic techniques, and we examine the risk of mortality post-LMP as a function of race and maternal smoking while investigating the possible time dependence of these associations.

MATERIALS AND METHODS

Data source

The national linked birth/infant-death data sets for the United States for 1998 were used for this study (13). These data were assembled by the National Center for Health Statistics of the Centers for Disease Control and Prevention in cooperation with the individual states and the District of Columbia.

Gestational age assignment is based predominantly on the date of the LMP. For a small proportion of records, for which the year and month of LMP was known but the day was unknown, the day was imputed by using a hot-deck procedure (14). For about 5 percent of births, a clinical estimate of gestation was used because of missing LMP date or month, but not year (13). In addition, the LMP-based estimate of gestation was replaced by the clinical estimate when the reported birth weight was grossly inconsistent with gestational age (based on LMP) (14). The National Center for Health Statistics performs these modifications to the gestational age data before releasing the data files.

We restricted our analysis to singleton births (n = 3,911,304) and excluded 57,919 (1.48 percent) births of infants whose birth weight was below 500 g and 156 (0.004 percent) whose birth weight was above 6,000 g. We further excluded 37,985 (0.97 percent) births for which gestational age was missing, 2,119 (0.05 percent) infants delivered before 22 weeks’ gestation, and 146,569 (3.75 percent) infants delivered at or after 43 weeks. As a sensitivity analysis, we also analyzed our data by including births up to 45 weeks’ gestation. Births occurring before 22 weeks were excluded because differential registration based on survival status biases mortality statistics (15, 16), whereas births at or after 43 weeks were excluded because of the high potential for gestational-age errors (17). Births to non-Black or non-White women (n = 197,277; 5.0 percent) were also excluded. Finally, we excluded an additional 13,625 (0.35 percent) pregnancies for which birth weight/gestational age combinations were implausible based on the algorithm of Alexander et al. (18), leaving 3,455,654 singleton births for analysis.

Statistical analyses

Age-specific post-LMP mortality was defined as death occurring at a specific number of days after the date of the mother’s LMP. Included were stillbirth, neonatal death, and postneonatal death. For example, a stillbirth at 175 days post-LMP and a birth at 170 days with early neonatal death 5 days later, that is, 175 days post-LMP, would both be considered post-LMP death at day 175.

We fit birth-weight- and gestational-age-specific mortality curves for smokers versus nonsmokers and for Blacks versus Whites by using smoothing splines with 5 degrees of freedom. We then analyzed the data by using the extended Cox proportional hazards regression model (19). Time zero was recorded as the date of the LMP as listed on the birth certificate or imputed by using the birth date and the clinical estimate of gestational age, and infants were censored at 47 weeks post-LMP (the last time when a neonatal death could occur in our data set; in the sensitivity
analysis, this time was changed to 49 weeks post-LMP). It is important to note that censoring occurs only at the end of the at-risk time period and not at 4 weeks postbirth for each infant. This latter approach would in effect be the study of neonatal mortality and would violate assumptions of the Cox regression model because censoring would not be independent of the outcome. To take into account and test for the differences in risk before and after birth, birth was entered into the analysis as a time-varying covariate coded 0 for before birth and 1 for after birth. The crossover pattern seen in classical analyses and in analysis in which the fetuses-at-risk approach is used (12) suggested that the hazard ratio may vary across time, so we took this into account by using an extended Cox model that allows nonproportional hazards (19).

A test for nonproportional hazards using the Schoenfeld residuals (19) was performed, and smooth estimates of hazard ratios were calculated using the method of Therneau and Grambsch (19). Thus, the statistical model for the analysis was \( h(t|X) = h_0(t) \exp[f_1(t)X_1(t) + f_2(t)X_2 + f_3(t)X_3] \), where \( h(t|X) \) is the instantaneous hazard for mortality at time \( t \) post-LMP and given covariates \( X \); \( h_0(t) \) is the baseline hazard when all covariates are zero; \( X_1(t) = 0 \) if \( t < t_b \), where \( t_b \) is the time of birth, and 1 if \( t \geq t_b \); \( X_2 \) is a binary indicator for race (1 for Black, 0 for White); and \( X_3 \) is a binary indicator for maternal smoking at LMP (1 for yes, 0 for no). Finally, \( f_1(t) \), \( f_2(t) \), and \( f_3(t) \) are smooth functions of time representing the time-dependent log hazard ratio for a unit increase in their respective covariates. If the proportional hazards hypothesis is rejected at the 0.05 significance level, the estimated shape of \( f_1(t) \) describes the pattern of the changes during time since LMP in the impact of \( X_1 \) on mortality. Otherwise, to reduce the risk of overfitting bias, the nonparametric estimate of \( f_1(t) \) is replaced by a constant \( \beta_j \), consistent with the proportional hazards assumption (20).

All analyses were carried out using S-PLUS version 6.0 software for Windows (Insightful Corporation, Seattle, Washington).

RESULTS

Table 1 contains summary statistics for the data set. The majority of births were to Whites; compared with Blacks, Whites had slightly higher average birth weights and gestational ages. The proportions of mothers who smoked, 9 percent of Blacks and 14 percent of Whites, was comparable to proportions found in U.S. population-based data for pregnant women.

Figures 1 and 2 show the crude gestational-age-specific perinatal mortality curves for Blacks and Whites and for smokers and nonsmokers, respectively. The crossing of the mortality curves is apparent, with the group for whom total perinatal mortality was higher (Blacks, smokers) having a lower risk at early gestational ages. A similar pattern occurred when birth weight was used as the x-axis (data not shown).

Table 2 contains the estimated hazard ratios for post-LMP mortality for each predictor, along with the p values for tests for nonproportionality of hazards. Both maternal smoking and maternal race had statistically significant effects, with infants of smokers and of Black women at higher risk for post-LMP mortality. Finally, newborn infants were at much higher average risk of death than were fetuses of the same post-LMP age.

However, the hazard ratio estimates shown in table 2 are valid only if the underlying proportional hazards assumption is correct (20). Figures 3, 4, and 5 show the estimated nonproportional log hazard ratios for maternal smoking, maternal race, and birth, respectively. Smoking was associated with higher mortality across all post-LMP weeks. The effect of smoking was statistically significantly different from a constant, however, with an estimated hazard ratio increasing from 1.09 (95 percent confidence interval: 0.98, 1.22) at 22 weeks to 1.82 (95 percent confidence interval: 1.72, 1.92) at 40 weeks. The effect of race was also statistically significantly different from a constant hazard ratio; however, the magnitude of this difference was smaller. Finally, the effect of birth was statistically significantly different from a constant and was a nonmonotonic function of time since LMP (figure 5). At preterm gestational ages, newborn infants had a much higher risk than fetuses remaining in the uterus, while they had a lower risk at term.

Including births at gestational ages 43–45 weeks had virtually no effect on the above findings (results not shown).

DISCUSSION

In this paper, we have shown that perinatal mortality curves comparing smokers with nonsmokers and Blacks with Whites intersect if gestational-age-specific perinatal mortality rates are based on the number of stillbirths and livebirths occurring at any particular birth weight or gestational age. We have provided further evidence that this crossover in perinatal mortality rates may be an artifact, and
it may be eliminated by representing gestational age as a time axis in survival analysis and by considering the appropriate denominator. We have proposed a new measure, post-LMP mortality, that combines stillbirths and neonatal deaths in a way that avoids problems inherent in the standard analysis of perinatal and infant mortality (11) while using a time-dependent indicator of birth to separate the two types of mortality.

Perinatal mortality curves (as calculated conventionally) intersect because the steady decline in the ratio of deaths to births is reversed at different gestational ages in dissimilar populations. Using relative birth weight and relative gestational age eliminates these crossovers by comparing births at the same deviation from subgroup means. However, as we have demonstrated elsewhere, these approaches can cause other problems by assigning the same relative value to observations that have fundamentally different characteristics. In particular, at low relative birth weights, mean gestational age can differ substantially between groups (Platt et al., unpublished manuscript). In addition, use of birth weight (crude or relative) without gestational age confounds two processes—chronologic maturity and fetal growth—and cannot address risk factors that affect these two processes differentially (21–23).

Our results are substantially different from those based on traditional denominators, primarily because the different denominators are generated from very different risk sets. Traditional perinatal mortality curves cross because the “birth rate” (12) differs in the two compared groups; at early gestational ages, there are more observed births (relative to the total population) in the disadvantaged group, causing the proportion of deaths to births to be lower. However, when the denominator includes all fetuses and infants at risk, the mortality rate is higher for the disadvantaged group.

Our use of information on fetuses and infants at similar ages post-LMP in denominators for stillbirth is not new; this method of calculating stillbirth risk at different gestational ages was proposed several years ago (24). However, to our knowledge, our use of this approach to include neonatal and infant deaths is novel. Use of post-LMP mortality as the outcome in this work eliminates the problems due to use of perinatal mortality (11), and extended Cox regression with time-varying hazards (19) allows appropriate analysis. The use of time post-LMP is not new, either; in neonatal follow-up, it is standard to treat preterm infants on the basis of “corrected age” (time post-LMP) rather than chronologic age.

Our survival analytic approach to post-LMP mortality provides some important advantages over current methods. First, it respects the temporal nature of gestational age. Second, it enables simultaneous analysis of important risk factors in a single model specification. The relative-birth-weight approach and others based on birth weight or gestational age relative to the distribution for the relevant population subgroup may be biased (Platt et al., unpublished manuscript) and are limited in that one must specify the rele-
vant population subgroup. This approach is ideal for single comparisons (e.g., Blacks vs. Whites) but cannot easily be implemented for comparisons along multiple axes (e.g., Black smokers vs. White nonsmokers). The extended Cox proportional hazards approach that we propose allows adjustment for multiple covariates. Furthermore, comparisons in which relative birth weight is used are restricted to single categorical variables (e.g., race, smoking vs. nonsmoking). Examination of the effects of continuous predictors on mortality risk (e.g., number of cigarettes smoked per day) and the study of interactions can be easily implemented with our proposed method. Other methods (25) adjust relative birth weight for potential confounders in a multivariate model but retain other problems with the relative-birth-weight approach.

It is helpful to think of our approach in the context of a cohort of conceptions and to think about the relevant comparisons in terms of risk sets. For example, consider a liveborn infant born at exactly 40 completed weeks of gestation (280 days post-LMP) and having a birth weight of 2,800 g who dies after 20 days of life (300 days post-LMP). Conditioning on birth weight, the risk set would include all infants born at 2,800 g and exclude one born at the same time whose birth weight was 2,900 g. Conditioning on gestational age in completed weeks, the risk set would include all births, including stillbirths at 280–286 days post-LMP, but would ignore fetuses at risk of being born at subsequent gestational ages, that is, those who remain in utero through this gestational week. The fetuses-at-risk approach (12) is an improvement because it considers all infants at risk of being born but excludes those born 1 day earlier (in week 39), even though

![Figure 2](https://example.com/f2.png)  
**FIGURE 2.** Perinatal mortality rate for smokers (solid line) and nonsmokers (dashed line) as a function of gestational age, 1998 US National Center for Health Statistics data.

<table>
<thead>
<tr>
<th>Predictor</th>
<th>Hazard ratio</th>
<th>95% confidence interval</th>
<th>$\chi^2$ for nonproportionality</th>
<th>df</th>
<th>p value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Race (Black)</td>
<td>1.72</td>
<td>1.67, 1.77</td>
<td>75.56</td>
<td>1</td>
<td>&lt;0.0001</td>
</tr>
<tr>
<td>Smoking</td>
<td>1.69</td>
<td>1.64, 1.75</td>
<td>319.83</td>
<td>1</td>
<td>&lt;0.0001</td>
</tr>
<tr>
<td>Birth</td>
<td>24.64</td>
<td>23.59, 25.75</td>
<td>14,982.46</td>
<td>1</td>
<td>&lt;0.0001</td>
</tr>
</tbody>
</table>
FIGURE 3. Estimated hazard ratio (curved solid line) (with 95% confidence interval (the two dotted lines)) for Black race (relative to White race) as a function of time post–last menstrual period, 1998 US National Center for Health Statistics data. The horizontal dashed line represents the estimated fixed hazard ratio.

FIGURE 4. Estimated hazard ratio (curved solid line) (with 95% confidence interval (the two dotted lines)) for maternal smoking at the time of the last menstrual period (LMP) (relative to mothers who did not smoke) as a function of time post-LMP, 1998 US National Center for Health Statistics data. The horizontal dashed line represents the estimated fixed hazard ratio.
these infants are very similar. The current approach considers all of the births discussed in this paragraph as part of the same risk set.

Several clinical and epidemiologic implications emerge from our findings. The mortality hazard ratio for Blacks versus Whites shows that fetuses and infants of Black mothers are at a significantly increased risk of death post-LMP at any gestational age. The findings for smoking are more strikingly time dependent, a result not apparent in birth-weight- and gestational-age-specific analyses. The effect of smoking is near null at the earliest viable gestational ages (22–25 weeks) but increases substantially by term. This finding is consistent with conceptualizing smoking as a risk factor that can reduce fetal growth in generally healthy infants, but with most fetal growth (and hence the adverse effect of smoking) occurring during the last trimester. It is also consistent with the argument that smoking is a risk factor for both mortality and reduced fetal growth and that these processes operate in parallel.

The effect of birth on mortality (figure 4) is worth special mention. The curve can be interpreted as follows: early in gestation (22–37 weeks post-LMP), those infants who are born are at higher risk than those still unborn; after 37 weeks post-LMP, the trend reverses. This finding is as expected: it is better to be born at term than either earlier or later.

With regard to epidemiologic and statistical modeling in perinatal epidemiologic research, the proposed conceptualization of the risk set mandates a reexamination of the role of gestational age. Logistic regression modeling of neonatal mortality with gestational age (or birth weight) as a determinant is based on the conventional gestational-age-specific denominators and disregards the temporal nature of gestational age. Crossover of mortality curves is implicit in such a model, and the effect of race (or maternal smoking, plurality, etc.) represents an average of effects observed across gestational ages.

Our study has the limitations typical of studies involving information obtained from large databases. Gestational age as measured by using the date of LMP is prone to substantial error (26). However, we used an error-correction technique that eliminates the largest errors (18). The gestational age of stillbirths typically represents the physician-determined gestational age at delivery rather than gestational age at death, which may occur days or even weeks earlier. The remaining errors are unlikely to have affected comparisons in a substantial way. If errors were nondifferential between groups, their effect would have been to create observations at early gestations that were in fact term births. This problem will underestimate the risk due to birth at early gestations; for these newborns, the observed risk will be lower than the risk expected at their nominal gestational age. Other errors, including inconsistencies due to data transcription or missing information, are relatively infrequent. This work is also limited by available data. It would be particularly inter-

**FIGURE 5.** Estimated hazard ratio (curved solid line) (with 95% confidence interval (the two dotted lines)) for birth as a function of time post–last menstrual period (LMP). This information represents the ratio of the hazard for liveborn infants relative to fetuses in utero at a given time post-LMP, 1998 US National Center for Health Statistics data. The horizontal dashed line represents the estimated fixed hazard ratio.
esting to be able to study other factors (e.g., bacterial vaginosis, psychosocial stress) that affect gestational duration rather than fetal growth; however, birth certificate data do not include such factors.

In summary, we have provided a setting for analysis of fetal and neonatal mortality that eliminates problems resulting in an artifact of intersecting perinatal mortality curves, enables modeling of multiple risk factors using gestational age as a time axis, and accounts for possible changes in their impact over time since LMP.

ACKNOWLEDGMENTS

This work was funded by Canadian Institutes of Health Research (formerly Medical Research Council of Canada) grant MA-14707.

Drs. Platt, Abrahamowicz, Joseph, and Kramer are career scientists of the Canadian Institutes of Health Research. Dr. Joseph is the recipient of a Clinical Research Scholar award from the Dalhousie University Faculty of Medicine and the Peter Lougheed/Canadian Institutes of Health Research (CIHR) New Investigator award from the Canadian Institutes of Health Research. Drs. Abrahamowicz and Kramer are James McGill Professors at McGill University. Dr. Ananth is supported in part through a grant (HD-38902) from the National Institutes of Health, Bethesda, Maryland.

REFERENCES