Population Research and Policy Review **22:** 41–64, 2003.

Race/ethnicity, Apgar and infant mortality

JAMIE MIHOKO DOYLE[1], SAMUEL ECHEVARRIA[2] & W. PARKER FRISBIE[3]
[1]*The University of Pennsylvania;* [2]*Ohio State University;* [3]*The University of Texas at Austin*

Abstract. *Objective.* Our general objective of this study is to further assess the predictive validity of Apgar scores on infant mortality using a national-level data set allowing for race/ethnic-specific variation. *Method.* This analysis is based on the 1989–1991 NCHS Linked Birth/Infant Death files. Multivariate, multinomial logistic regression models were constructed adjusting for maternal behavioral and health risks, socioeconomic and demographic factors, and infant characteristics. *Results.* After adjusting for maternal risk factors, infants with low Apgar scores experienced 86 times the odds of neonatal death relative to high Apgar score infants while infants with medium Apgar scores experienced 10 times the odds of neonatal death relative to the high Apgar referent. After controlling for maternal risk factors, birth weight and gestational age, the effect of low Apgar scores on neonatal mortality risk relative to high Apgar scores is reduced (from OR = 86.1 to OR = 31.2), but still remains a strong predictor. Fully adjusted, race/ethnic-specific models show that the predictive value of low Apgar on neonatal mortality risk is strongest for Mexican Americans (OR = 44.2) versus Non-Hispanic Whites and Blacks (OR = 36.2 and 20.4 respectively). *Conclusion* Our results support the conclusion that Apgar scores continue to be strong predictors of infant survival, independent of birth weight, gestational age, and a large number of maternal risk factors. They are also shown to be powerful predictors within different race/ethnic groups, solidifying their importance as both a diagnostic tool and descriptive health characteristic across various U.S. populations.

Keywords: Apgar scores, Infant mortality, Race/ethnicity, Risk factors

Introduction

Reasons for investigating disparities in infant mortality are well known (Bird 1995; Frisbie et al. 1997), including the need to reduce, and eventually eliminate, racial/ethnic inequalities in pregnancy outcomes (Healthy People 2000; Healthy People 2010). However, the explanation of race/ethnic inequalities in the risk of infant death remains far from complete. The infant mortality rate for Non-Hispanic Blacks (henceforth "Blacks") has been about twice that of Non-Hispanic Whites (or "Anglos") for many years, a disparity that has increased over the past decade (Guyer et al. 1998).

Pregnancy outcome differentials among Blacks and Anglos have spurred an ongoing interest in uncovering explanations for disparities in infant mortality. Since traditional measures of SES and other risk factors fail to adequately account for these disparities, it is surprising that very little research has been conducted using what appears to be one of the strongest predictors of infant survivorship, namely Apgar scores (Casey et al. 2001). The general objective of the present study is to shed additional light on the reasons for the disparities in the risk of infant death among Blacks, Anglos and Mexican Americans by focusing on attention on the predictive power of Apgar scores, which are routinely recorded at both one and five minutes following delivery, in combination with other risk factors that are more conventionally used in research of this nature.

In 1953, Virginia Apgar developed what became known as the "Apgar scoring system" as a method of evaluating an infant's need for medical attention at birth. Subsequent research has also shown it to be a reliable predictor of infant mortality (e.g., Casey et al. 2001). Each component of the Apgar scoring system is based on signs traditionally used by anesthesiologists to monitor a patient's condition during surgery (Papile 2001). These signs include heart rate, respiratory effort, reflex irritability, muscle tone, and color. Each component is given a score of 0 to 2 with a total maximum Apgar score of 10. A higher score implies a greater likelihood of survival. Perhaps the most significant of the five dimensions is heart rate, and the least important, and most highly debated, component is color. Newborn infants are often grouped into three categories according to their Apgar score: (1) severely depressed (0–3), (2) moderately depressed (4–6), and (3) good to excellent condition (7–10) – frequently denoted as Low, Medium, and High, both in previous research and in the present study. A score under seven typically indicates an infant's need for medical intervention such as cardiopulmonary resuscitation (CPR), artificial respiration, catheterization, and oxygen therapy as well as careful observation of the infant (Weinberger et al. 2000).

The Apgar scoring system has been used for approximately 50 years to assess the need for resuscitative care in newborn infants and as a predictor of infant mortality and morbidity (Weinberger et al. 2000). Jepson et al. (1991) report that 79% of full-term infants with low Apgar scores (0–3) died within one year of birth. More recent comprehensive reassessments of the utility of 1-minute and 5-minute Apgar scores yield overwhelming evidence supporting the conclusion that Apgar scores continue to be effective predictors of infant mortality and morbidity (Casey et al. 2001; Weinberger et al. 2000; Asakura et al. 2000; Papile 2001; Thorngren-Jerneck & Hurst 2001). Casey et al. (2001) report that Apgar has stronger predictive validity than umbilical artery pH, which some scholars have suggested as a more objective predictor. Birth weight and gestational age have long been considered to be the most powerful predictors of infant survival. However, at least one study concludes that Apgar scores were only "slightly less predictive than either of these variables alone" with the most powerful model being one that combines Apgar scores, birth weight and gestational age (Behnke et al. 1987: 121).

Although current research strongly supports the general predictive power of Apgar scores, research on the relationship between race/ethnicity and Apgar scores is quite limited. Petrikovsky et al. (1990) briefly comment that Black infants have lower Apgar scores than their White counterparts. A study by Heygi et al. (1998) also found that both 1- and 5-minute Apgar scores were significantly higher for White than for Black infants. Some authors have suggested that the depressed scores for Black infants are due to the "color" component of the Apgar score (Yama & Marx 1991). The argument is that there is a small likelihood of a Black infant appearing "pink", and that this at least partially curtails the potential of Black infants from receiving high Apgar scores. Evidence from the very small amount of comparative research suggests that Blacks infants often have a more depressed score than White infants (Heigi et al. 1998; Petrikovsky et al. 1990).

There is only a small amount of research comparing the Apgar scores of Blacks and Anglos and, as far as we can determine, virtually none investigating Hispanics Apgar scores. Thus it is important to identify and investigate other U.S. race/ethnic groups as a means of delineating the varied effects of social inequality on pregnancy outcomes.

The inherent heterogeneity of the Hispanic population has led researchers to conduct analyses on various Hispanic sub-groups, which demonstrate a wide range of pregnancy outcomes (Frisbie et al. 1998; Hummer et al. 1999). With respect to characteristics such as average education, prenatal care, and health insurance, Hispanics are at a greater disadvantage than Blacks (Bean & Tienda 1987; Frisbie et al. 1997). Due to their low socioeconomic status (SES) profile, many Hispanic groups experience a greater percentage of adverse pregnancy outcomes relative to their SES-advantaged Anglo counterparts. Although Mexican American mothers are characterized by a disadvantageous SES environment similar to that of Blacks, their infant mortality rates (IMR) are similar to, and sometimes superior to those for Anglos (Office of Minority Health 1998). This combination of comparatively favorable mortality rates (among adults, as well as infants) with a high-risk SES and sociodemographic profile has come to be known as an "epidemiological paradox" (Markides & Coreil 1986).

As suggested above, there has been no research either on Mexican American versus Anglo differentials in Apgar scores or on race/ethnic variation in predictive validity. However, there exists one analysis of Apgar scores and infant mortality within one specific Hispanic sub-population. Del Valle and Calzada (1994) investigated the relationship between Apgar scores and infant mortality in Puerto Rico, taking into account demographic and sociodemographic variables, but without race/ethnic comparisons. The lack of comparative analyses by race/ethnicity underscores the need for further investigation.

Conceptual framework

General approach

Before delineating the specific contributions offered by the present research, it may be useful to clarify the place of Apgar scores in infant mortality research. Apgar scores represent a risk factor inasmuch as they are an apparently very useful composite measure of the extent to which other risk factors have affected the survival chances of infants through various physiological dimensions. In this respect, an Apgar score is not conceptually different than either birth weight or gestational age, both of which are also infant characteristics which identify adverse births associated with deviations from ideal *in utero* development. Thus, a most important research direction involves analyses in which Apgar score is the outcome of interest – an extension that we are currently undertaking. However, before proceeding much farther with a research agenda that is comparative by race/ethnicity, we need additional information on at least two questions – information that is important in its own right and that will also provide necessary guidance for future research. First, does an infant's Apgar score add significant predictive power regarding risk of infant death, *beyond that which is provided by the powerful effects of birth weight and gestational age*, or are Apgar scores largely an indicator of risk engendered and/or captured by low weight and prematurity? If the latter is true, then the association of Apgar scores with infant mortality risk should largely disappear when models are statistically adjusted for birth weight and gestational age. Previous research, as cited above, suggests that Apgar retains significant predictive power with controls for weight and gestational age. However, we need to explore this question *with a national-level data set that allows examination of race/ethnic differentials* based on models that control for as wide a range of other risk factors as possible. Second, we need considerably more information on whether Apgar scores are *differentially effective* in their predictive power for different race/ethnic populations – again with a nationally representative data set that allows adjustment for other known risk factors.

Accordingly, the present research expands on previous work by focusing attention on the relationship between race/ethnic differences in Apgar scores and disparities in infant mortality among Mexican Americans, Blacks and Anglos. The two different issues requiring further exploration imply the need for two types of logistic regression models. First, we estimate *models with race/ethnicity as a covariate*, along with Apgar scores and a number of demographic, SES, behavioral and maternal health covariates. Second, we estimate *race/ethnic-specific models*. Models with race/ethnicity as a covariate will allow, apparently for the first time, a determination of the extent to which Apgar scores are independently predictive of infant mortality differences across three race/ethnic populations, net of a wide range of controls. However, such models cannot shed light on the question of whether Apgar scores are more powerful predictors for one or another of the race/ethnic groups. Race/ethnic-specific models directly address the latter issue.

Rationale for inclusion of specific risk factors

Infant mortality is associated with a substantial number of risk factors, including health history and medical care, as well as the more commonly included variables such as socioeconomic, demographic, and behavioral variables. For purposes of exposition, variables are grouped into three categories: socioeconomic/demographic factors, maternal health/behavioral factors, and characteristics of infant. The characteristics associated with the lowest risk of infant death are, in all instances, the reference in the models subsequently estimated.

Socioeconomic and demographic factors

The effects of maternal age on infant mortality are greater near the extremes of the fecund years. The risk of infant death is higher among teenage multiparas and older women giving birth for the first time. The Kleinman–Kessel index takes this joint effect into account (see also Hummer et al. 1995). There is also evidence relating very young maternal age to both low birth weight for gestational age and preterm delivery (Geronimus 1986; McCormick 1985).

Education exhibits a positive relationship with pregnancy outcomes (Cramer 1987, 1995; Frisbie et al. 1998). Because data on income are not available in vital statistics records, maternal education is frequently used as the socioeconomic status indicator in this type of study (Hummer et al. 1992: 1060).

Since infant mortality is higher for an infant whose mother is unwed, marital status is also included in this analysis. This higher risk likely results from the tendency of unmarried women to have inadequate familial, social, and economic resources (Eberstein et al. 1990). This, in turn, leads to lower birth weight and a subsequent increase in mortality risk (Frisbie et al. 1997; Kallan 1993).

Nativity offers further insight into race/ethnic differences. Infants born to immigrant women are more likely to have positive pregnancy outcomes, in terms of both birth outcomes and infant survival. The most common explanations of the advantage of the foreign birth are positive selection of migration and more healthy lifestyles of immigrants (Frisbie et al. 1998; Hummer et al. 1999; Scribner 1996). Previous studies conclude that immigrant women are more likely to have healthy diets (Guendelman & Abrams 1995) and are less likely to smoke or consume alcohol or drugs (Rumbaut and Weeks 1991, 1996). Although our interest here is not in assessing the validity of the two explanations, it is worth noting that, to the extent that the positive effect of foreign-birth on infant survival is maintained after controlling for behavioral differences, the hypothesis that women who migrate tend to be healthier than those who do not gains support.

Maternal behavioral and health factors

Smoking by a woman during pregnancy is included as a risk factor, partly based on the assumption that this behavior may indicate less concern with a healthy lifestyle. More important, smoking is known to directly and adversely affect birth outcomes, Apgar scores, and to increase the risk of mortality and morbidity (Garn et al. 1981; Thorngren-Jerneck & Herbst 2001).

There are three reasons for the inclusion of maternal weight gain in this analysis. The first is that, while maternal weight gain and nutritional status are not equivalent measures, their close relationship justifies weight gain as a proxy for nutrition (Chomitz et al. 1995). Second, a strong positive correlation exists between maternal weight gain and birth weight across race/ethnic groups (Frisbie et al. 1998). Third, Chomitz et al. (1995) found that most of the weight gain during pregnancy relates to fetal growth. The likelihood of fetal growth retardation lessens when pregnant women gain adequate weight; however, its relationship to prematurity is uncertain (Chomitz et al. 1995; Kramer et al. 1992).

Other maternal health factors associated with adverse pregnancy outcomes include the presence of labor and delivery complications and medical risks (Apgar 1953; Balcazar 1993; Kallan 1993; Rumbaut & Weeks 1991). Labor and delivery complications include abruptio placentae, breech presentation, oligohydramnios, labor induction, meconium in the amniotic fluid, and fetal hydrops. The presence of any maternal medical risks such as anemia, diabetes, hemoglobinopathy (a blood disorder associated with anemia, including sickle cell anemia), eclampsia, hypertension (both chronic and pregnancy induced), incompetent cervix, uterine bleeding, and other risks are indicated by a medical risks variable. Numerous studies reinforce the association of adequate prenatal care (PNC) with favorable pregnancy outcomes across race/ethnic categories (Eberstein et al. 1990; Institute of Medicine 1988; Frisbie et al. 2001), though recent studies have questioned whether the positive relationship with birth weight is casual or a product of self-selection (Alexander & Korenbrot 1995; Fiscella 1995). Whatever the nature of the effect of PNC on birth weight, there is substantial reason to believe that adequate PNC improves both infant survival chances and maternal health due to the "package of necessary services" (e.g., well-baby care and referrals) it provides (Shiono & Behrman 1995).

Following earlier research, previous pregnancy loss is also included in this analysis as a measurement of reproductive efficiency. Women who have had a history of miscarriages or stillbirths are at greater risk of adverse pregnancy outcomes (Eberstein et al. 1990; Kallan 1993; Frisbie et al. 1997). The NCHS data file for 1989–1991 also identifies women who have had a preterm or small-for-gestational-age (SGA) birth.

Infant characteristics

Birth weight and gestational age are typically viewed as the most powerful and proximate predictors of infant death or survival. However, one study that included 1- and 5-minute Apgar scores in models along with birth outcomes, reports that the predictive power of Apgar scores was slightly higher than for either birth weight or gestational age alone (Behnke 1987). Both birth weight and gestational age are controlled in our regression models, and the predictive utility of these factors compared to Apgar scores is evaluated. Finally, sex of infant is a necessary control because the risk of death is consistently greater for male babies (Powell-Griner 1988), although male infants tend to have a higher average birth weight.

Data/methods

General analytical approach

We report only results associated with 1-minute Apgar scores. Although the bulk of previous research has utilized 5-minute Apgar scores in lieu of 1-minute scores, for several reasons we feel it is important at this juncture to focus on the predictive power of 1-minute Apgar scores. To assess the predictive effects of Apgar scores relative to the more conventional demographic measures of birth weight and gestational age, we feel 1-minute Apgar scores most closely parallel birth weight and gestational age in that they are a more static measure in their relationship to infant mortality. In the case of low 1-minute Apgar score infants, there are usually various diagnostic reasons for attempting neonatal interventions in order to increase an infant's survival chances. Successful resuscitative efforts for infants with severely depressed Apgar scores at 1- minute reduces the variance in 5-minute Apgar scores and therefore dramatically increases the predictive relationship between low scores and infant mortality. In order to present a more conservative estimate of Apgar effects on infant mortality *and* avoid any possible interaction effects between race/ethnicity, low Apgar scores, and the availability, utilization and efficacy of neonatal technologies on infant survival, we only utilize 1-minute Apgar scores in our analyses.

To compare the relationship of Apgar scores to the risk of infant mortality for Anglos, Blacks, and Mexican Americans, logistic regression models were constructed with infant mortality regressed on the risk factors, including race/ethnicity. For reasons delineated earlier, race/ethnic- specific models were also estimated. Results are presented in the form of odds ratios.

Data

Our analysis is based on the1989–1991 NCHS Linked Birth/ Infant Death files for the United States (excluding Texas, California, and Oklahoma) with a total N of 6,544,004, including 5,100,942 Anglos, 1,252,870 Blacks, and 190,192 Mexican American mothers. Deficiency in the reporting of Apgar scores is the reason for exclusion of these three states. This is unfortunate since California and Texas are the two states with the largest Mexican American populations. Nevertheless, in the absence of any previous comparative analysis of Hispanics, we include Mexican Americans in our analysis. Given the large numbers of Mexican Americans resident in California and Texas, conclusions regarding the latter group must of course be regarded as tentative.

The NCHS 1989–1991 data set differs from its predecessors in important ways. It includes formerly unavailable items on maternal weight gain, smoking, maternal medical conditions, and previous pregnancy loss. Clinical measurements of gestational age were added allowing evaluation of, and substitution for, the sometimes-questionable estimates of gestational age calculated as the difference between date of birth and date of last normal menses. In the data on which our analysis is based, the only variable with a substantial amount of information missing is maternal weight gain during pregnancy. Emulating techniques previously used in handling this problem (Hummer et al. 1999; Singh & Yu 1996), a "missing" category for weight gain was created in the regression models.

The present study excludes plural births, infants with birth weights under 500 grams, and infants with less than 22 weeks of gestation. "Live births" under 500 grams and/or less than 22 weeks gestation are often excluded because they are thought to often reflect coding errors or misclassified stillbirths

(Buehler et al. 1987). Plural pregnancies result in smaller infants who also are more apt to have low Apgar scores. The relationship involving plurality is well known, and not part of the focus of this research.

Measurement

For the most part, risk factors are measured in a conventional manner as shown in Table 1. However, a few of the operationalizations warrant additional comment. The risk of infant death is higher among teenage multiparas and older women giving birth for the first time. We use the parity index developed by Kleinman & Kessel (1987) to take this joint effect into account. The index has three categories: (1) First Births; (2) Low Parity: second-order births to women 18 and older and third-order births to women 25 and older; (3) High Parity: second- or higher-order births to women under 18, third- or higher-order births to women under 25, and fourth- and higher-order births to women 25 and older (see also Hummer et al. 1995). Our diagnostics show no collinearity problem between this measure of parity and maternal age.

Prenatal care is measured by the Kotelchuck Adequacy of Prenatal Care Use (APNCU) index (Kotelchuck 1994a, b), which includes a high-risk "adequate plus" category of care that takes negative selection into account. This allows the identification of women who record more prenatal care visits than the standard recommended by the American College of Obstetricians and Gynecologists. Such women are likely to be negatively selected due to problem pregnancies. We use a three-category variant (adequate, less than adequate, and adequate plus) of the APNCU that takes into consideration timing of the first visit, number of visits, and length of pregnancy.

Birth outcomes (birth weight and gestational age) have most often been measured as dichotomies. Infants born at low birth weight ($<$2,500 grams) are distinguished from normal birth weight (2,500 grams or higher). Infants of less than 37 weeks gestation have been regarded as premature, while term infants are those born at 37 weeks and beyond. However, it has become increasingly clear that finer-grained measures afford more analytical leverage (Frisbie et al. 1996; Solis et al. 2000). Indeed, findings in past research which indicated substantial predictive power of Apgar scores, net of birth weight and gestational age effects, may have been due in part to lack of measurement precision. For present purposes, we employ a birth outcome classification, developed by Yerushalmy (1967), that has been widely used in public health research. This typology contains a separate category for very low birth weight ($<$1,500 grams) and makes most of the distinctions shown to be important by Frisbie et al. (1996), including a separate category for "heavy preemies" (infants born at weights $\geq$ 2,500 grams with gestational age $\leq$ 36 weeks; see Table 1 for the full classification).

A large proportion of research on the relationship between Apgar scores and infant mortality has focused on neonatal mortality (death at less than 28 days) (Drage & Berendes 1965; Casey et al. 2001), but some scholars have also examined the relationship in the postneonatal period (28 days to one year) (Jepson et al. 1991). Accordingly, we operationalize the outcome of interest as a trichotomy: neonatal mortality, postneonatal mortality, and infant survival past the first year of life.

Results

Descriptive results

Table 1 provides descriptive information, by race/ethnicity for the outcomes of interest (infant, neonatal, and postneonatal mortality rates per 1,000 live births) and for all risk factors (in terms of percentage distributions) included in the analysis. Consistent with all previous research, the overall infant mortality rate (IMR), as well as the neonatal mortality rate (NMR) and postneonatal mortality rate (PNMR), are all substantially higher among Blacks than among either Anglos or Mexican Americans. The results also support previous studies that report similar infant mortality rates for Anglos and Mexican Americans. Note that mortality rates in this analysis are lower than those reported in the vital statistics because of the exclusion of plural births and infants weighing less than 500 grams. Results from Table 1 also show that Anglos and Mexican Americans have greater percentages of infants with high 1-minute scores (both 92.4%) than do Blacks, among whom only 89.6% received a high score at one minute.

In line with nearly all prior research, Black mothers are disadvantaged relative to their Anglo and Mexican American counterparts with respect to weight gain, out-of-wedlock births, presence of medical risks, and especially in regard to the high proportions of low weight births. Indeed, while the normal birth outcome percentage for both Anglos and Mexican Americans is near or above 90%, the percentage normal among Blacks is only 78.3% (Table 1). Mexican American mothers are more disadvantaged than Blacks in terms of education and the percentage receiving adequate or better prenatal care. Mexican Americans smoke less and are less apt to have a history of previous preterm or small for gestational age births, or prior pregnancy loss. They are the least likely to present with medical risks or labor and delivery complications. If, as has often been suggested, adverse pregnancy outcomes are less common among foreign-born women, then Mexican American women have far and away the most advantageous nativity distribution (Hummer et al. 1999). Low parity births are most likely among Anglo women, who also

Table 1. Descriptive statistics of births by race/ethnicity

	Race/ethnicity			
	Anglo	Black	Mexican American	Total
Infant mortality	5.4	11.1	5.8	6.5
Neonatal mortality	2.8	5.5	3.1	3.3
Postneontal mortality	2.5	5.6	2.7	3.1
Apgar				
Low (0–3)	1.7	3.1	1.8	2.0
Medium (4–6)	5.9	7.3	5.8	6.2
High (7–10)	92.4	89.6	92.4	91.9
% Teen birth	10.0	24.7	19.0	13.1
% U.S.-born	96.7	95.4	44.8	95.0
% Unmarried	16.7	68.4	35.4	27.1
Parity				
Low	46.2	37.9	38.8	44.4
First birth	42.9	37.9	37.5	41.8
High	10.9	24.2	23.7	13.8
Education				
0–11 years	15.3	31.3	57.4	19.6
12 years	39.9	43.0	28.3	40.2
13+ years	44.8	25.7	14.3	40.2
Prenatal care				
< Adequate	25.4	43.8	50.2	29.7
Adequate	49.6	30.9	31.9	45.4
Adequate+	25.0	25.3	17.9	24.9
Weight gain				
Low	5.5	10.7	7.6	6.6
Normal	88.8	77.0	79.9	86.2
Missing	5.7	12.3	12.5	7.2
% Prenatal smoking	21.3	15.9	6.8	19.8
% Previous preterm/SGA	1.2	1.5	0.8	1.2
% Previous loss	25.6	26.6	20.1	25.6
% w/Med. Risk(s) present	18.7	23.5	18.2	19.6
% w/LDL complication	32.6	33.9	31.8	32.9
% Male	51.3	50.7	51.3	51.2
Birth outcomes				
(1) < 1,500 g	0.6	2.0	0.7	0.9
(2) 1,500 g–2,499 g, $\leq$ 36 ga	1.9	5.0	2.1	2.5
(3) 1,500 g–2,499 g, > 36 ga	1.9	4.2	2.1	2.3
(4) $\geq$ 2,500 g, $\leq$ 36 ga	4.8	10.5	6.7	6.0
(5) $\geq$ 2,500 g, > 36 ga	90.8	78.3	88.4	88.3
N	5,100,942	1,252,870	190,192	6,544,022

Source: NCHS, 1989–1991 linked birth/infant death files: U.S. residents, ≥500 g BWT, ≥22 weeks GA, singletons [excluding TX, OK and CA].

Table 2. Percent distribution of births by age of infant death, Apgar score and race/ethnicity

Age of infant death	Neonatal			Postneonatal			Survived		
Apgar	Low	Med.	High	Low	Med.	High	Low	Med.	High
Race/ethnicity									
Anglo	54.7	20.9	25.4	9.2	13.7	77.1	1.5	5.9	92.6
Black	59.3	19.0	21.7	12.8	15.4	71.7	2.7	7.2	90.1
Mexican American	56.0	20.4	23.6	9.4	16.6	74.1	1.6	5.7	92.7
N		21,885			20,484			6,501,635	

Source: NCHS, 1989–1991 linked birth/infant death files: U.S. residents, ≥500 g BWT, ≥22 weeks GA, singletons [excluding TX, OK and CA].

far exceed the other two race/ethnic groups in the percentage who receive more than a high school education. However, Anglo women are the most likely to smoke. All of these results are consistent with the results of earlier studies (Cramer 1987; Frisbie et al. 1998; Hummer et al. 1999).

Table 2 shows the Apgar distributions for Anglos, Blacks, and Mexican Americans for infants who died in the neonatal and postneonatal periods as well as those who survived the first year of life. As expected, the results are particularly striking for infants who died very early in life. Well over half of all neonatal mortality occurred among infants whose Apgar scores were low (0–3); with the Black percentage standing at 59.3%. About one-fifth of neonatal deaths are associated with medium scores. A much smaller, but still rather substantial proportion (roughly one-quarter) of postneonatal deaths occurred among infants who did not score high on Apgar.

If the relative distributions are calculated with Apgar scores as the initial subdivision (not shown), we find that, across all three race/ethnic groups, neonatal death is at least nine times more likely when the Apgar score is low than when the score is in the medium range (4–6). For all race/ethnic groups, under one percent of infants who scored high on Apgar failed to survive their first year, while among those with low scores, the proportion who died varies from 9.0% to 10.6%, depending on race/ethnicity. Thus, while Table 1 shows that low Apgar scores are fairly uncommon (two percent overall), Table 2 shows that a huge proportion of infant deaths are associated with low scores

Logistic models with race/ethnicity as a covariate

The values in the first column of Table 3 (designated Model 0) are the simple bivariate effects of each separate risk factor on risk of neonatal and post-neonatal mortality, compared to survival chances. Note that in this and all subsequent models, values are significant at least at the $p \leq 0.05$ level, unless specified as non-significant (ns) in the table. The findings from column 1 of

Table 3 are essentially the same as those that emerge from a descriptive analysis based on a mortality crosstabulation according to risk factor categories. Our focus on the predictive power of Apgar scores appears justified in that low scores on this indicator are associated with an extreme elevation in the risk of neonatal mortality (OR = 120.63), and with a substantial increase in risk of postneonatal death (OR = 7.32). Only the impact of very low birth weight ($\leq$ 1,500 grams) is greater. The remaining variables in column 1 are related to infant mortality as expected. We now turn to the multivariate models in which race/ethnicity is a covariate beginning with Model 2 of Table 3.

In Model 1 of Table 3, only race/ethnicity and Apgar scores are included. When adjusted for the latter, the odds of neonatal death among Blacks drop substantially (from 1.96 in the bivariate model to 1.37 in Model 1). On the other hand, risk of postneonatal mortality is only slightly attenuated. The magnitudes of Mexican Americans odds ratios are barely affected, but they become insignificant when compared to Anglos in Model 1. The magnitudes of the odds ratios associated with Apgar scores are not much affected by an adjustment for race/ethnicity.

Model 2 includes all covariates, except for the five-category birth outcome measure. From this model, it is obvious the Apgar score has by far the strongest effect on the odds of both neonatal and postneonatal mortality. In Model 2, the risk of Black neonatal mortality (OR = 1.04) is diminished to near equality with the Anglo risk (albeit still statistically significant). However, it should be noted that given the very large number of cases in our data set, differences of very small magnitude will usually attain statistical significance. Thus, while we continue to report *p*-values, we pay more attention to the direction and magnitude of estimates.

The postneonatal odds for Blacks are also rather substantially reduced. In addition, for Mexican American infants, the odds ratios for both neonatal and postneonatal mortality are now below those for Anglos. Note also that the magnitudes of the effects of the control variables are consistently reduced (some even losing statistical significance) as compared to the bivariate relationships (column 1, Model 0), which suggests that Apgar scores mediate the effects of many risk factors. A reasonable view is that Apgar scores are a very useful summary measure of the effects of determinants of infant mortality. Further, it is our hypothesis that, even with subjectivity in scoring, the Apgar index provides an indicator of the health status of neonates that is no doubt related to, but also distinct from, other crucial indicators, including birth weight and gestational age. However, any attempt to answer the primary questions underlying this analysis – namely (1) does evaluation of infants *via* Apgar provide information on perinatal health, *over and above that available from measuring birth weight and knowing gestational age*, that should mo-

Table 3. Multinomial logistic regression odds ratios of infant mortality by maternal and infant risk factors*

	Model 0		Model 1		Model 2		Model 3		Model 4	
	Neo.	Post.	Neo.	Post.	Neo.	Post.	Neo.	Post.	Neo.	Post.
Intercept	−5.69	−5.76	−7.10	−6.14	−7.86	−7.13	−7.28	−6.98	−7.48	−6.00
Race/ethnicity [Angle]										
African American	1.96	2.20	1.37	2.05	1.04	1.34	0.79	1.17	0.75	1.16
Mexican American	1.10	1.06	1.07^{ns}	1.06^{ns}	0.91	0.87	0.88	0.85	0.88	0.85
Apgar [high]										
Low	120.63	7.32	116.45	6.70	86.09	5.74			31.24	3.23
Medium	12.62	2.87	12.46	2.78	10.51	2.54			5.67	1.85
Age [20+ years]										
< 20 years	1.53	1.87			1.29	1.41	1.13	1.37	1.16	1.38
Nativity [foreign-born]										
U.S.-born	1.09	1.41			0.94^{ns}	1.20	1.01	1.20	0.95	1.19
Marital status [married]										
Unmarried	1.75	2.35			1.03^{ns}	1.32	0.98	1.23	0.94	1.27
Parity [low]										
First birth	1.20	0.85			0.90	0.69	0.87	0.68	0.85	0.68
High	1.48	2.10			1.13	1.44	1.10	1.41^{ns}	1.09	1.41^{ns}
Education [13+ years]										
0–11 years	1.74	3.14			1.12	1.58	1.07	1.54	1.06	1.54
12 years	1.30	1.65			1.08	1.22	1.07	1.22	1.05	1.21
Prenatal care [adequate]										
< Adequate	1.78	2.09			1.22	1.36	1.07	1.30	1.05	1.29
> Adequate	3.06	1.67			2.04	1.40	1.03	1.10	1.05	1.10

Table 3. Continued

	Model 0		Model 1		Model 2		Model 3		Model 4	
	Neo.	Post.	Neo.	Post.	Neo.	Post.	Neo.	Post.	Neo.	Post.
Previous loss [no]										
Yes	1.28	1.11			1.13	1.03^{ns}	1.02^{ns}	1.00^{ns}	0.99^{ns}	1.00^{ns}
Weight gain [normal]										
Low	5.05	2.21			3.14	1.53	1.57	1.19	1.42	1.18
Missing	3.02	1.70			2.18	1.31	1.52	1.14	1.36	1.12
Medical risks [none]										
Present	2.66	1.57			1.45	1.24	1.12	1.12	1.09	1.10
Labor/delivery [none]										
Present	3.68	2.15			1.60	1.23	0.84	0.85	0.87	0.86
Smoking status [no]										
Yes	1.23	2.11			0.96	1.62	0.79	1.47	0.82	1.48
Infant sex [female]										
Male	1.23	1.31			1.21	1.29	1.40	1.36	1.31	1.34
Birth outcomes [(5) normal]										
(1) ≤ 1,500 g	232.69	22.42					175.78	15.64	34.50	9.33
(2) 1,501 g–2,500 g, ≤ 36 ga	15.61	5.08					13.40	3.66	6.98	3.19
(3) 1,501 g–2,500 g, > 36 ga	9.71	4.64					9.57	3.45	6.57	3.24
(4) > 2,500 g, ≤ 36 ga	3.12	2.00					2.98	1.58	2.69	1.56
−2LL	570,223		482,522		465,848		462,855		433,894	
N	6,544,04		6,544,004		6,544,004		6,544,004		6,544,004	

Source: NCHS, 1989–1991 linked birth/infant death files: U.S. residents, ≥500 g BWT, ≥22 weeks GA, singletons [excluding TX, OK and CA].
* All coefficients significant at $p < 0.05$ unless denoted non-significant (ns).

tivate intensive medical intervention, and if so, (2) does this predictive power differ by race ethnicity – cannot be addressed until we inspect the results of additional models which include the birth outcome measure. At this juncture, we wish only to emphasize two other findings. First, even with adjustment for all variables, *except birth outcomes*, the likelihood of neonatal death is 86 times greater if the Apgar score is low and 10 times greater if the score is medium, than if the score is high. Second, from Model 2 onward, the risk of both neonatal and postneonatal death becomes significantly lower among Mexican Americans than among their Anglo counterparts.

Model 3 of Table 3 adds the five-category birth outcome measure (with normal birth outcome as the reference), but excludes Apgar scores. With only a few exceptions, the results are remarkably similar to those observed from Model 2. Birth outcome (especially very low birth weight) is the most powerful predictor, effects of other risk factors are attenuated, and the risk of Mexican American infant mortality is less than the Anglo risk. There are some notable differences between Model 2 and Model 3 results, however. First, with adjustment for birth outcome, the Black neonatal mortality becomes less than that of Anglos as evidenced by an odds ratio less than unity (OR = 0.79), and the odds of Black postneonatal mortality are also modestly lower (OR = 1.17 as compared to OR = 1.34 in Model 2). This is not particularly surprising since research over many years has consistently documented the fact that the low birth weight rates are much higher among Blacks than among Whites (NCHS 1998). In other words, it was virtually certain that controlling for this compositional disadvantage would lead to a sharp diminution of the risk for Black infants. Second, the odds associated with a previous preterm/SGA birth become less than unity in the case of both neonatal and postneonatal mortality. The same is true of smoking, though only in regard to neonatal risk. Obviously, this *cannot* be interpreted to mean that a history of problem pregnancies or smoking has a beneficial effect on infant survival. A partial explanation almost surely is that the impact of both of these factors is mediated by birth outcome, particularly low birth weight. This would account for a reduction to unity, but why do the estimates fall below unity? Our interpretation in the case of previous premature/SGA births is that greater precautions are apt to be taken by the woman and her family and that more intensive medical scrutiny (and perhaps earlier intervention) by a medical professional is more likely. Interpretation of the anomalous smoking effect is less certain, especially since it appears with respect to neonatal, but not postneonatal, mortality. It is possible, but not plausible in our view, that a differential amount of medical attention is paid to smokers vs. nonsmokers at parturition. The most likely explanation is instability of estimates, perhaps due to cell sizes so small that a shift of one or two infants between the outcome categories results in a "flip-flop" in the odds.

The full model incorporates all risk factors so that the issue of whether Apgar scores have predictive utility independent of that provided by birth outcome can be directly addressed. The answer from Model 4 of Table 3 is a decided "Yes". The magnitude of the odds of neonatal mortality associated with a low Apgar score is greatly reduced (OR = 31.24 in Model 4 as compared to OR = 86.09 in Model 2). The effect of the birth outcome measure is reduced even more. For example, the odds ratio for the very low birth weight category in Model 3 stood at 175.78, but with Apgar scores in the equation, the relative risk was five times lower (OR = 34.50 in Model 4). The magnitude of the odds associated with medium Apgar scores and other compromised birth outcome categories are also diminished, though in a less dramatic fashion. The evidence seems relatively clear: Apgar scores retain very substantial predictive power, even after adjustment for a well-established measure of birth outcome based on birth weight and gestational age. The relationships of other risk factors cannot be used to qualify this conclusion inasmuch as their effects scarcely change between Models 3 and 4 of Table 3.

Race/ethnic specific models

The aim of determining whether Apgar scores predict differentially by race/ethnicity remains to be addressed. We address this question via the race/ethnic-specific models that appear as Table 4. *All* risk factors (except race/ethnicity, of course) are included in the logistic regressions on which this table is based. To conserve space, we show only the net effects of the primary variables of interest (i.e., Apgar scores and birth outcome) but the models in Table 4 include controls for all risk factors.

The results of Table 4 are rather striking. Apgar scores have great predictive utility for each of the race/ethnic groups, net of all controls. For both Anglos and Mexican Americans, the odds ratios relating low Apgar scores to neonatal infant mortality risk are of greater magnitude than those for very low birth weight. The predictive power for Mexican American infants is especially strong (OR = 44.24 for low scores). Very low birth weight is more strongly related to neonatal risk among Blacks than are low Apgar scores, but a large, independent association with Apgar remains. The estimates linking medium Apgar odds to neonatal mortality are of lesser magnitude, but they elevate risk by four to over seven times with the same race/ethnic pattern as that seen for low scores. As expected, the relationships with postneonatal mortality are of considerably smaller magnitude, but are nevertheless quite large relative to the effects of other risk factors (not shown, but available from the authors upon request). For later infant deaths, the edge in predictive

Table 4. Multinomial logistic regression odds ratios of infant mortality and race/ethnicity by maternal risk factors (not shown) and birth outcomes*

	Race/ethnicity					
	Anglo		Black		Mexican American	
	Neo.	Post.	Neo.	Post.	Neo.	Post.
Intercept	−7.64	−7.02	−7.32	−6.49	−7.47	−6.94
Apgar [high]						
Low	36.21	3.67	20.40	2.60	44.24	3.98
Medium	6.17	1.93	4.16	1.64	7.28	2.72
Birth outcomes [(5) normal]						
(1) ≤ 1,500 g	35.94	9.39	29.61	9.59	33.38	8.06
(2) 1,501 g–2,500 g, ≤ 36 ga	7.76	3.48	4.80	2.90	9.19	2.71
(3) 1,501 g–2,500 g, > 36 ga	7.60	3.68	3.87	2.60	7.95	3.99
(4) > 2,500 g, ≤ 36 ga	3.07	1.60	1.63	1.49	3.09	1.38
N	5,100,942		1,252,870		190,192	

Source: NCHS, 1989–1991 linked birth/infant death files: U.S. residents, ≥500 g BWT, ≥22 weeks GA, singletons [excluding TX, OK and CA].
* All coefficients significant at $p < 0.05$ unless denoted non-significant (ns).

power goes to birth outcome, but the adjusted odds associating Apgar scores with postneonatal risk are also fairly sizable.

Collinearity assessment

Apgar scores are known to be correlated with birth weight (Behnke et al. 1987) and gestational age (Catlin et al. 1986). If the two later explanatory variables are too highly correlated with Apgar scores, the models used in this study may suffer from a multicollinearity problem, particularly since these two factors compose the index for birth outcomes. We tested for collinearity in three ways: first by estimating a set of appropriate, nominal-level, bivariate measures of association, similar model used in this study; secondly, by estimating Pearson correlation coefficients between the full, continuous distributions of gestational age and birth weight with the full 10-point range of Apgar scores; and lastly by estimating similar regression models used in this study utilizing Ordinary Least Squares (OLS), looking specifically at the tolerance and variance inflation factors (VIF) for each of the explanatory variables. The method of observing the tolerance and VIFs has been previously suggested for multicollinearity diagnostics in logistic models (Allison 2000: 48–51), and for other regression models (Neter et al. 1996: 385–388; Belsley et al. 1980: 92–96). Our focus is on the correlation of the explanatory variables – not the parameter estimates using OLS. We have constructed bivariate, OLS models where the dependent variable is infant death (with survival as the reference category).

Results from the multicollinearity diagnostics

In general, bivariate measures of association show minimal correlation between Yerulshamy's birth outcome measure and 1-minute Apgar scores, with the Phi coefficient estimated at 0.28 and Cramer's V at 0.19. The correlation coefficients between Apgar score and gestational age (0.15) as well as birth weight (0.15) also show that multicollinearity should not present a major obstacle in the interpretation of our results.

The tolerance and VIFs show that the correlation between Apgar scores and birth outcomes (including both birth weight and gestational age separately) does not produce a collinearity problem. The tolerance for each of the variables for each model was no more than 0.97 while the VIF was no more than 1.29.

Discussion

Conclusions

The findings regarding the two substantive questions that gave rise to our analysis seem rather clear. In brief, our results support the conclusion that Apgar scores are strong predictors of infant survival, independent of birth weight, gestational age, and a large number of control variables. Apgar scores have high and independent predictive validity for all three race/ethnic groups considered, but show a stronger predictive power for Mexican Americans and Anglos, even after adjusting for birth outcome. This result was perhaps to be expected in light of the large compositional disadvantage of Blacks with respect to birth weight distributions. The so-called "color bias" may also have something to do with the fact that Apgar odds ratios were smaller for Blacks than for the other two groups. It may also be that Black infants survive better at severely depressed Apgar scores, just as previous research has found somewhat better survival chances among very low weight Blacks infant as compared to their White counterparts (Wilcox & Russell 1986, 1990). However, Apgar scores remain a strong indicator of the need for ameliorative medical action for infants with low and medium scores – be they Black, Mexican American, or Anglo.

The predictive validity of the total Apgar score across all racial/ethnic groups may be due to the individual, strong effects of each component which contribute to the overall composite score. In essence, as birth weight and gestational age are readily available and somewhat easily measured general proxy measures for normal, non-deleterious fetal development and maturity, the component measures of the Apgar score taken at birth represent much more proximate indicators of said fetal development. To be sure, Apgar scores represent infant characteristics that are obviously related to birth weight and gestational age in that they both point to a certain objective state of fetal growth, but by being much more diagnostically complex, Apgar is situated much closer down the proximate determinant chain that we use to identify infant mortality risk. Heart rate, respiratory effort, reflex irritability, muscle tone, and color may even have independent risks on infant mortality, some more so than others. For example, "heart rate" certainly has the greatest diagnostic and prognostic value of the five signs, while "color" has the least (Apgar 1953). Despite their possible differential effects on infant mortality risk prediction, we hypothesize that the synergistic effect of the five components explains the predictive power of the total Apgar score. Decomposing the Apgar score into individual measurements and observing their independent effects on mortality risk is beyond the scope of this analysis, but may be an interesting subject for future research.

There are some obvious limitations to this analysis. Perhaps the most serious is the unknown generalizability of the results to the sizable Mexican American populations residing in Texas and California. Apgar scores constitute an important, but clearly imperfect, indicator. Inevitably, scoring involves some subjectivity. For example, studies show that obstetricians tend to assign higher scores than the delivery team, and independent observers tend to assign lower scores than those persons responsible for the delivery of the newborn (Del Valle & Calzada 1994; Clark 1988; Schmidt 1998). Such findings have led scholars, including Virginia Apgar, to suggest that an independent observer should assign Apgar scores to infants (Apgar 1953; Schmidt 1998; Jepson 1991). Nevertheless, very recent analyses, including this one, support the validity of Apgar scores.

We would be remiss if we did not reiterate that, while the Apgar method provides extremely important information regarding neonatal health and the need for interventions, apart from that provided by birth weight and gestational age, we do not question in any way the conclusion that the latter two variables are the most crucial for adequately explaining race/ethnic differences in risk of infant mortality. For example, in an ancillary model (not shown) we regressed our three-category dependent variable on race/ethnic categories and the five-category birth outcome measure. The result was that the risk of Black *neonatal* risk became lower than that of Anglos. Black postneonatal risk remained at least 60% higher and the odds for Mexican

Americans continued to be virtually the same as the Anglo odds. None of this diminishes the importance of the finding that Apgar scores possess powerful and independent predictive validity, and that the relationship of the index remains very strong across the three race/ethnic groups. Indeed, our results suggest the need for continued investigations, including research in which Apgar score is the outcome of interest.

Acknowledgements

The authors gratefully acknowledge the support provided for this research by the National Institute of Child Health and Human Development (Grant No. R01 HD36429). We also extend appreciation to Robert A. Hummer and Starling Pullum for their valuable comments and assistance.

References

Alexander, G. R. & Korenbrot C. C. (1995). The role of prenatal care in preventing low birth weight, *The Future of Children* 5: 103–120.

Allison, P. D. (2000). *Logistic Regression Using the SAS System*. Cary NC: SAS Institute.

Apgar, V. (1953). A proposal for a new method of evaluation of the newborn infant, *Current Researches in Anesthesia and Analgesia* 32: 260–267.

Asakura, H., Ichikawa, H., Nakabayashi, M., Ando, K., Kaneko, K., Kawabata, M., Tani, A., Satoh, M., Takahashi, K. & Sakamoto, S. (2000). Perinatal risk factors related to neurologic outcomes of term newborns with asphyxia at birth: A prospective study, *Journal of Obstetric and Gynecological Research* [Japan] 26: 313–324.

Atkinson, D. (1983). An evaluation of Apgar Scores as predictors of infant mortality, *North Carolina Medical Journal* 44: 45–54.

Balcazar, H. (1993). Mexican Americans' intrauterine growth retardation and maternal risk factors, *Ethnicity and Disease* 3: 169–175.

Bean, F. D.& Tienda, M. (1987). *The Hispanic Population of the United States*. New York: Russell Sage Foundation.

Behnke, M., Carter, R. L. & Hardt, N. S. (1987). The relationship of Apgar scores, gestational age, and birthweight to survival of low-birthweight infants, *American Journal of Perinatology* 4: 121–124.

Belsley, D. A., Kuh, E. & Roy, E. W. (1980). *Regression Diagnostics: Identifying Influential Data and Sources of Collinearity*. New York: John Wiley.

Bird, S. T. (1995). Separate Black and White infant mortality models: Differences in the importance of structural variables, *Social Science and Medicine* 41: 1507–1512.

Buehler, J. W., Kleinman, J. C., Hogue, C. J. R, Strauss, L. T. & Smith, J. C. (1987). Birth weight-specific infant mortality, United States, 1960 And 1980, *Public Health Reports* 102: 151–161.

Casey, B. M., McIntire, D. D. & Leveno, K. J. (2001). The continuing value of the Apgar score for the assessment of newborn infants, *New England Journal of Medicine* 344: 467–471.

Catlin, E. A., Marshall, W. C., Brann, B. S., Mayfield, S. R., Shaul, P. W., Goldstein, M. & Oh, W. (1986). The Apgar score revisited: Influence of gestational age, *The Journal of Pediatrics* 109(5): 865–868.

Chomitz, V. R., Cheung, L. W. Y. & Lieberman, E. (1995). The role of lifestyle in preventing low birth weight, *The Future of Children: Low Birth Weight* 5: 121–138.

Clark, D. A. & Hakanson, D. O. (1988). The inaccuracy of Apgar scoring, *Journal of Perinatology* 8(3): 203–205.

Cramer, J. C. (1987). Social factors and infant mortality: Identifying high-risk groups and proximate causes, *Demography* 24: 299–232.

Cramer, J. C. (1995). Racial and ethnic differences in birthweight: The role of income and financial assistance, *Demography* 24: 231–247.

Del Valle, M. Z. & Calzada, J. L. (1994). Apgar score and infant mortality in Puerto Rico, *Puerto Rico Health Science Journal* 13: 175–181.

Drage, J. S. & Berendes, H. (1965). Apgar scores and outcome of the newborn, *Pediatric Clinics of North America* 13: 635–643.

Eberstein, I. W., Nam,, C. & Hummer, R. A. (1990). Infant mortality by cause of death: main and interaction effects, *Demography* 27: 413–430.

Fiscella, K. (1995). Does prenatal care improve birth outcomes? A critical review, *Obstetrics and Gynecology* 85: 468–879.

Frisbie, W. P., Forbes, D. & Pullum, S. (1996). Comprised birth outcomes and infant mortality among racial and ethnic groups, *Demography* 33: 469–481.

Frisbie, W. P., Biegler, M., de Turk, P. B., Forbes, D. & Pullum, S. (1997). Determinants of intrauterine growth retardation and other compromised birth outcomes: A Comparison of Mexican Americans, African Americans, and non-Hispanic Whites, *American Journal of Public Health* 87(12): 1977–1983.

Frisbie, W. P., Forbes, D. & Hummer, R. A. (1998). Hispanic pregnancy outcomes: Additional evidence, *Social Science Quarterly* 79: 149–169.

Frisbie, W. P., Echevarria, S. & Hummer, R. A. (2001). Prenatal care utilization among Non-Hispanic Whites, African Americans, and Mexican Americans, *Maternal and Child Health Journal* 5(1): 21–33.

Garn, S. M., Johnston, M., Ridella, S. A. & Petzoid, A. S. (1981). Effect of maternal cigarette smoking on Apgar scores, *American Journal of Diseases of Children* 135: 503–506.

Geronimus, A. T. (1987). On teenage childbearing and neonatal mortality in the United States, *Population and Development Review* 13: 245–271.

Guendelman, S. & Abrams, B. (1995). Dietary intake among Mexican-American women: Generational differences and a comparison with white non-Hispanic women, *American Journal of Public Health* 85: 20–25.

Guyer, B., MacDorman, M. F., Martin, J. A. Peters, K. D. & Strobino, D. M. (1998). Annual summary of vital statistics – 1997, *Pediatrics* 102: 1333–1349.

Hamvas, A., Wise, P. H., Yang, R. K., Wampler, N. S., Noguchi, A., Maurer, M. M., Walentik, C. A., Schramm, W. F. & Cole, F. S. (1996). The influence of the wider use of surfactant therapy on neonatal mortality among Blacks and Whites, *New England Journal of Medicine* 334: 1635–1640.

Heygi, T., Carbone, T., Anwar, M., Ostfeld, B., Hiatt, M., Koons, A., Pinto-Martin, J. & Paneth, N. (1998). The Apgar score and its components in the preterm infant, *Pediatrics* 101: 77–89.

Hummer, R. A., Eberstein, I. W. & Nam, C. B. (1992). Infant mortality differentials among Hispanic groups in Florida, *Social Forces* 70: 1055–1075.

Hummer, R. A., Schmertmann, C. P, Eberstein, I. W. & Kelly, S. ((1995). Retrospective reports of pregnancy wantedness and birth outcomes in the United States, *Social Science Quarterly* 76: 402–418.

Hummer, R. A., Biegler, M., de Turk, P. B., Forbes, D., Frisbie, W. P., Hong, Y. & Pullum, S. (1999). Race/ethnicity, nativity, and infant mortality in the United States, *Social Forces* 77: 1083–1118.

Institute of Medicine. (1988). *Prenatal Care: Reaching Mothers, Reaching Infants*. Washington, DC: National Academy Press.

Jepson, H. A., Talashek, M. L & Tichy, A. M. (1991). The Apgar score: Evolution, limitations, and scoring guidelines, *Birth* 18: 83–92.

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

Kleinman, J. E. & Kessel, S. S. (1987). Racial differences in low birthweight: Trends and risk factors, *New England Journal of Medicine* 317: 749–753.

Kotelchuck, M. (1994a). An evaluation of the Kessner adequacy of prenatal care index and a proposed adequacy of prenatal care utilization index, *American Journal of Public Health* 84: 1414–1420.

Kotelchuck, M. (1994b). The adequacy of prenatal care utilization index: Its US distribution and association with low birthweight, *American Journal of Public Health* 84: 1486–1489.

Kramer, M. S., McLean, F. H., Eason, E. L. & Usher, R. H. (1992). Maternal nutrition and spontaneous preterm Birth, *American Journal of Epidemiology* 136: 574–583.

Markides, K. S. & Coreil, J. (1986). The health of Hispanics in the southwestern United States: An epidemiologic paradox, *Public Health Reports* 101: 253–265.

McCormick, M. C. (1985). The contribution of low birth weight to infant mortality and childhood mortality, *New England Journal of Medicine* 312: 82–90.

National Center for Health Statistics. (1998). *Socioeconomic Status and Health Chartbook*. Hyattsville MD: National Center for Health Statistics.

Neter, J., Kutner, M. H., Nachtsheim, C. J. & Wasserman, W. (1996). *Applied Linear Regression Models*, 3rd edn. Chicago IL: Irwin.

Papile, L. (2001). The Apgar score in the 21st century [Editorial], *New England Journal of Medicine* 344: 519–520.

Petrikovsky, B. M., Diana, L. & Baker, D. A. (1990). Race and Apgar scores [Correspondence], *Anesthesia* 45: 988–999.

Powell-Griner, E. (1988). Differences in infant mortality among Texas Anglos, Hispanics, and Blacks, *Social Science Quarterly* 69: 452–467.

Rumbaut, R. & Weeks, J. R. (1991). Perinatal risks and outcomes among low income immigrants. *Final Report for the Maternal and Child Health Research Program*. Rockville: DHHS.

Rumbaut, R. & Weeks, J. R. (1996). Unraveling a public health enigma: Why do immigrants experience superior perinatal health outcomes, *Research in the Sociology of Health Care* 138: 337–391.

Schmidt, B., Kirpalani, H., Rosenbaum, P. & Cadman D. (1998). Strengths and limitations of the Apgar score: A critical appraisal, *Journal of Clinical Epidemiology* 41: 843–850.

Scribner, R. (1996). Editorial: Paradox as paradigm – the health outcomes of Mexican Americans, *American Journal of Public Health* 86: 303–305.

Shiono, P. H. & Behrman, R. E. (1995). Low birth weight: Analysis and recommendations, *The Future of Children* 5: 4–18.

Singh, G. & Yu, S. M. (1996). Adverse pregnancy outcomes: Difference between US and foreign-born women in major US racial and ethnic groups, *American Journal of Public Health* 86: 837–843.

Solis, P., Pullum, S. G. & Frisbie, W. P. (2000). Demographic models of birth outcomes and infant mortality: An alternative measurement approach, *Demography* 37: 489–498.

Thorngren-Jerneck, K. & Herbst, A. (2001). Low 5-minute Apgar score: A population based register study of 1 million term births, *Obstetrics and Gynecology* 98: 65–70.

U.S. Department of Health and Human Services (2000). *Healthy People 2010: National Health Promotion and Disease Prevention Objectives*. Washington, DC: U.S. GPO.

U.S. Department of Health and Human Services (1991). *Healthy People 2000: National Health Program and Disease Prevention Objectives*. Washington, DC: U.S. GPO.

Weinberger, B., Anwar, M., Heygi, T., Hiatt, M., Koons, A. & Paneth, N. (2000). Antecedents and neonatal consequences of low Apgar scores in preterm infants, *Archives of Pediatrics and Adolescent Medicine* 154: 294–312.

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

Wilcox, A. J. & Russell, I. T. (1990). Why small Black infants have a lower mortality rate than small White infants: The case for population-specific standards for birth weight, *Journal of Pediatrics* 116: 7–10.

Yerushalmy, J. (1967). The classification of newborn infants by birth weight and gestational age, *Journal of Pediatrics* 71: 164–172.

Yama, A. Z. & Marx, G. F. (1991). Race and Apgar scores, *Anesthesia* 46: 330–331.

Address for correspondence: Dr J. M. Doyle, University Pennsylvania, Population Studies Center, 3718 Locust Walk, Philadelphia PA, 19104-6298, U.S.A.
Phone: +512-619-8408; Fax: 512-471-4886;
E-mail: jmdoyle@sas.upenn.edu