J. biosoc. Sci. (2005) **37**, 193–208

DOI: 10.1017/S0021932004006595

MATERNAL, NEONATAL AND COMMUNITY FACTORS INFLUENCING NEONATAL MORTALITY IN BRAZIL

CARLA JORGE MACHADO* AND KENNETH HILL†

**Departamento de Demografia, CEDEPLAR/UFMG, Brazil and †Department of Population and Family Health Sciences, Johns Hopkins University, USA*

Summary. Child mortality (the mortality of children less than five years old) declined considerably in the developing world in the 1990s, but infant mortality declined less. The reductions in neonatal mortality were not impressive and, as a consequence, there is an increasing percentage of infant deaths in the neonatal period. Any further reduction in child mortality, therefore, requires an understanding of the determinants of neonatal mortality. 209,628 birth and 2581 neonatal death records for the 1998 birth cohort from the city of São Paulo, Brazil, were probabilistically matched. Data were from SINASC and SIM, Information Systems on Live Births and Deaths of Brazil. Logistic regression was used to find the association between neonatal mortality and the following risk factors: birth weight, gestational age, Apgar scores at 1 and 5 minutes, delivery mode, plurality, sex, maternal education, maternal age, number of prior losses, prenatal care, race, parity and community development. Infants of older mothers were less likely to die in the neonatal period. Caesarean delivery was not found to be associated with neonatal mortality. Low birth weight, pre-term birth and low Apgar scores were associated with neonatal death. Having a mother who lives in the highest developed community decreased the odds of neonatal death, suggesting that factors not measured in this study are behind such association. This result may also indicate that other factors over and above biological and more proximate factors could affect neonatal death.

Introduction

Determinants of infant and child mortality have been extensively studied in demographic and epidemiological research and an important conclusion is that in the 1990s there was a continued decrease in mortality among infants and children in most of the developing world (Rutstein, 2000). This sustained decline is mostly the result of child survival interventions, such as immunization and oral rehydration (Diamond, 2000). The average decline in child mortality in the 1990s was estimated to be most pronounced in ages 2–4 (10·5% decrease), and for children aged one year (4% decrease). Neonatal and postneonatal mortality rates declined less, with 3·0% and 2·5% decreases, respectively (Rutstein, 2000), and as a consequence there is an increasing proportion of infant deaths occurring in the neonatal period, which now accounts, worldwide, for two-thirds of deaths in children less than one year of age, and nearly two-fifths of all deaths in children less than five years of age (Moss *et al.*, 2002). Therefore, to further reduce child mortality, factors associated with neonatal mortality must be addressed.

This article studies the determinants described in the literature as most likely to exert an effect on infant survival, provided that they were available in the data. Proximate determinants of neonatal mortality are analysed (sex, plurality, mode of delivery, prior losses, Apgar scores at 1 and at 5 minutes, gestational age, prenatal care and birth weight), as well as less proximate determinants (parity and maternal age) and distal determinants (race, maternal education and community development) for the 1998 birth cohort from the city of São Paulo, Brazil. These determinants are related to the mother, the infant and the community and may give a more complete spectrum of risk factors for neonatal mortality in Brazil. Apgar scores, birth weight and gestational age are highly correlated with neonatal survival and in combination are a measure of newborn well-being, success of resuscitative efforts, newborn size and newborn maturity (De Hart, 1994). Twins or higher order births account for a disproportionately large percentage of neonatal deaths in all populations (Guo, 1993); male infants are known to be more likely to die in the first year of life (Wells, 2000). However, the association between mode of delivery and infant survival is controversial (Machado, 2002). A mother with a higher number of prior losses is also considered to bear infants with higher risks of death, probably due to a higher risk of malformation (Clark *et al.*, 2002), and mothers with a higher number of prenatal care visits are more likely to bear children with a lower risk of death in infancy, and it has been suggested that the effect is more pronounced in developing than in developed countries (Villar & Bergsjo, 1997). Race is considered in the analysis, since black infants usually experience a much worse health status than white infants (Barros *et al.*, 2001). Maternal education may lower neonatal mortality since educated mothers are less likely to accept fatalist explanations and are more likely to manipulate modern medical systems than less educated mothers (Hobcraft, 1993). Also, maternal education may be seen at least partially as a surrogate for household income, which negatively affects infant death. Higher parities (four or higher previous live births) and nulliparities are considered to exert extra risks for birth outcomes and they are likely to reflect deviance from normal uterine function. Also, high parity can be seen, in general, as a marker for low social status and income, which affect infant death (Haaga, 1989). In Brazil, a negative association between fertility level and social status and income has been found (Giffin, 1994; Potter *et al.*, 2002). Also, high parity is possibly capturing the effects of closely spaced births, which are known to be deleterious to infant survival (Curtis *et al.*, 1993). Regarding maternal age, various socioeconomic disadvantages and suboptimal health outcomes are associated with adolescent pregnancy, such as insufficient education, limited career and job opportunities and poor conditions for effective parenting (Cowden & Funkhouser, 2001). Women aged 35 and older are also at higher risk for chromosomal abnormalities and other factors associated with higher risks of neonatal mortality (Bianco *et al.*, 1996). Finally, the community where the mother resides may affect infant survival since it may influence attitudes and behaviour. Some characteristics may influence birth outcomes more directly, such as the availability of health services and environmental pollution (Mosley & Chen, 1984).

Methods

Information was obtained from the SINASC (Information System on Live Births) and SIM (Information System on Mortality) datasets for the 1998 birth cohort whose mother's place of residence in 1998 was the city of São Paulo, Brazil. Data were retrieved on live births and infant deaths and probabilistically matched; a full explanation of the results and methods are described in detail elsewhere (Machado, 2002). In brief, a one-to-one relationship was considered to hold between records in the two files. Birth date, place of maternal residence, birth weight, sex, delivery mode, plurality and current maternal age were used as identifiers in both files to match the records. Unfortunately, even though a large number of identifiers in the probabilistic matching were used, more than one birth record per death record still had to be accepted (i.e. a birth record for each of the 2581 neonatal death records), since there was no basis whatsoever to rule out the extra records. This happened due to a number of factors: first, there was a very large number of births that could potentially match each neonatal death; second, the large amount of missing information for some variables in the death records, such as birth weight (41·2%), delivery mode (43·1%), plurality (43·2%) and maternal age (41·8%), which made the record linkage somewhat imprecise. The criterion used was that if a death record, with a missing value for a given variable, matched a birth record with a missing value, and also matched a birth record with a value (non-missing), both pairs would be kept (or all the pairs obtained). Finally, another problem was that some identifiers, such as plurality, do not discriminate very well, since, in this population, 96·5% of all births were singletons.

There is no established way of dealing with pairs for which uncertainty does exist. The simplistic approach of using only data from pairs that were very confidently considered to be true matches is not recommended, since much additional information from the set of uncertain pairs might be thrown away, which is clearly a selection problem. Results could be severely biased because certain subsets of true matches may reside primarily in the set of potential matches (Scheuren & Winkler, 1993). In this study the set of uncertain matches is heavily weighted, for example, by infants with Apgar scores lower than seven at 1 or at 5 minutes and by low birth weight infants (Table 1).

Intuitively, the set of erroneous matches (included in the set of uncertain matches) may lead to an understatement of the true correlation between the independent variable and the dependent variable. For example, if a pair is mismatched, it means that a death record was brought together with a birth record of an infant who survived the neonatal period. The infant who survived is more likely to have a satisfactory birth weight, or a higher Apgar score, than the infant who died.

https://doi.org/10.1017/S0021932004006595 Published online by Cambridge University Press

Table 1. Univariate description of infants by matching status: 1998 birth cohort, São Paulo, Brazil

		Percentage		
Variable	Categories	Matched with one birth record only ($n=1907$)	Matched with more than one birth record ($n=674$)	p value
Infant's sex	Male	46·0	45·8	0·961
	Female	54·0	54·1	0·923
	Missing	0·0	0·1	1·000*
Plurality	Singleton	90·7	92·1	0·248
	Twins/higher order	8·1	6·8	0·245
	Missing	1·2	1·1	0·833*
Delivery mode	Normal	53·9	58·4	0·040
	Caesarean	41·6	39·2	0·261
	Other	2·9	1·1	0·002*
	Missing	1·6	1·3	0·444*
Prior losses	None	52·5	51·8	0·732
	One	10·6	11·3	0·648
	Two or more	3·8	4·2	0·704
	Missing	33·0	32·7	0·862
Apgar at 1 min	0–6	22·2	53·2	0·000
	7 and 8	43·5	26·3	0·000
	9 and 10	27·8	8·5	0·000
	Missing	6·6	12·0	0·000
Apgar at 5 min	0–6	9·4	31·5	0·000
	7 and 8	15·7	25·2	0·000
	9 and 10	68·5	31·0	0·000
	Missing	6·4	12·3	0·000
Weeks gestation	27 or less	6·2	19·3	0·000
	28–36	13·5	30·6	0·000
	37–41	72·2	41·1	0·000
	42 or more	1·0	0·6	0·285*
	Missing	7·4	8·3	0·452
Prenatal visits	None	3·8	6·8	0·004
	1–6	30·1	38·5	0·000
	7 or more	29·1	19·6	0·000
	Missing	37·0	35·2	0·413
Birth weight (g)	999 or less	9·7	25·2	0·000
	1000–1499	6·9	18·6	0·000
	1500–1999	4·1	11·2	0·000

Table 1. *Continued*

Variable	Categories	Percentage: Matched with one birth record only (n=1907)	Percentage: Matched with more than one birth record (n=674)	p value
Birth weight (g) *continued*	2000–2499	8·5	9·5	0·425
	2500–2999	19·2	12·8	0·000
	3000–3499	31·5	12·2	0·000
	3500–3999	15·8	6·1	0·000
	4000–4499	2·7	1·0	0·040*
	4500 or higher	0·3	0·4	1·000*
	Missing	1·2	2·9	0·009*
Parity	None	33·5	35·5	0·357
	One	25·3	20·9	0·017
	Two or three	15·6	15·6	0·976
	Four or more	3·6	5·0	0·162
	Missing	22·0	22·9	0·610
Maternal age	11–14	0·6	1·1	0·359*
	15–19	18·0	22·3	0·021
	20–24	31·9	26·2	0·005
	25–29	22·5	23·7	0·544
	30–34	17·0	15·2	0·255
	35–39	6·7	8·7	0·099
	40 or higher	2·3	1·8	0·421*
	Missing	1·0	0·9	0·815*
Race/colour	White	23·4	21·0	0·181
	Non-white	11·3	10·6	0·650
	Missing	65·3	68·4	0·140
Maternal education	No schooling	1·5	1·5	1·000*
	Elementary, incomplete	36·2	37·1	0·669
	Elementary, complete	17·4	16·7	0·643
	Secondary	15·0	14·7	0·849
	College	5·0	4·2	0·356
	Missing	24·9	25·7	0·674
Community development**	Low	13·7	16·0	0·184
	Medium	39·4	39·5	0·956
	High	46·9	44·5	0·309

Note: The frequencies were weighted according to the weight given to each matched pair.
*Fisher's Exact test was used, instead of Pearson's Chi-squared test.
***n* for matched with one birth record only is 1887; *n* for matched with more than one birth record is 571.
Source: DATASUS, 2000.

https://doi.org/10.1017/S0021932004006595 Published online by Cambridge University Press

Therefore, the true association between birth weight and mortality or Apgar score and mortality is likely to be understated, which is a very plausible effect. The aim was to have an adjusted estimator of the true correlation that can remove this understatement. In order to deal with this issue, a weight was given to each of the 4245 matched pairs obtained. The weight given to each pair represented the certainty with which the matched pair was correctly assigned for the same infant. It was assumed that certainty was inversely related to the number of matches for each infant who died. Death records with fewer matches received higher weights compared with the others. The weight was the inverse of the number of times the death record appeared in the matched file. Most death records were matched to one birth record only (1907 deaths or 74% of all deaths) and received a weight of 1·0. Regarding the 'multiple' matches (those not matched to one and only birth record), the majority were matched to two or three birth records and received a weight of at least 0·33 (19% or 499 death records). The remaining 175 deaths (7% of all deaths) received a weight of less than 0·33, i.e. matched to four or more death records. Therefore, even though each of the 674 multiple matches was matched to an average of 3·47 birth records, the distribution was very skewed with only about 6% being matched to more than four death records. When the birth record file was merged with the matched file, the birth records that were not involved in any matched pair were given a weight of one (205,383 birth records).

A second step was to verify whether there were any birth records that matched more than one death record. This is important since summing the weights for these records would be incorrect, since theoretically there would be a weight of more than 1·0 for a birth record for which there would be considerable uncertainty as to whether the neonate has died. Indeed, there were 256 matched pairs (which corresponded to 131 death records) that were involved with a birth record that matched more than one death record. It was decided to exclude those pairs from the analysis in order to avoid having an overestimated weight for the 142 birth records involved with more than one death record. Therefore, instead of having 4245 matched pairs, there were 3989 matched pairs to analyse (94% of the 4245 matched pairs).

STATA 6 software was used for the analysis of matched data.

Variable constructs

The outcome variable is whether or not an infant died in the neonatal period. An indicator variable of neonatal mortality (dying before 28 days of life) was created after merging the matched file and the birth record file. From this birth cohort of 209,628 infants, 2581 infants died in the neonatal period (12·3 per 1000 live births).

Explanatory variables were classified as proximate determinants (Apgar scores at 1 and at 5 minutes, birth weight, gestational age, prenatal care, sex, plurality, prior losses and mode of delivery), less proximate determinants (parity and maternal age) and distal determinants (race, maternal education and community development). Both Apgar scores were classified as: below 7; 7 and 8; 9 and 10; and unknown. Earlier results (Machado, 2002) showed that Apgar scores seemed to be biased towards high values and it was hypothesized that infants considered to be healthy were given Apgar scores of 9 or 10 without much adherence to specific score components. Hence, an infant with an Apgar score of 7 or 8 may be perceived to be less healthy than an infant of Apgar 9 or 10. Infants with Apgar scores of 7 or 8 may thus be at increased risk of dying in the neonatal or postneonatal period. Birth weight was categorized into nine groups of 500 g each: <1000 g, 1000–1499 g, 1500–1999 g, 2000–2499 g, 2500–2999 g, 3000–3499 g, 3500–3999 g, 4000–4499 g, >4500 g, and missing birth weight. Categories of gestational age were taken as they appeared in the database: '27 weeks or less', '28–36 weeks', '37–41 weeks', '42 weeks or more' and 'missing'. Infants were also categorized according to whether they were singleton births, non-singleton births or if this category was missing. Infants were also classified based on whether the delivery was normal; a Caesarean; other (primarily forceps); and of unknown mode of delivery. The number of prenatal care visits was classified based on the existing categories: no visits, 1–6 visits, 7 or more visits and unknown number of visits.

Finally, the number of stillbirths and abortions were considered together in a category called 'prior losses' and four categories were created: no prior loss, one prior loss, more than one prior loss, and missing information on prior loss.

Maternal age was categorized into consecutive five-year age groups from 15 upwards, as is done in most demographic research, with a first category of 11–14 years to estimate the effect of very young maternal age on infant mortality. The last category was 40 years and over, and a final additional category was for missing age. With regard to parity, it was categorized as zero parity; parity one; parity two or three; parity four or higher; and missing parity, as suggested in Bai *et al*. (2002).

In categorizing race, since whites are historically the most affluent group in Brazil, and because among non-whites 97% are mixed race or black, two categories were defined: 'white' and 'non-white' (and 'missing': 67% of all infants). For maternal education the categories were: illiterate; incomplete elementary education; complete elementary education; secondary school; college; and an additional category for missing. Finally, information at the community (or district) level exists for each of the 96 districts in São Paulo. To capture community effects, an indicator variable of development was created for each district and the method of principal components' analysis used to weight different district level indicators (Filmer & Pritchett, 1999, 2001; Mahy, 1999). Principal components' analysis uses linear combinations of variables to explain sets of observations on many variables and simplify information contained in a group of variables. The method applied here summarized the district level information on eighteen variables, including average educational attainment of the household head, average household income, average access to health services and other facilities, in each district, among others. The technique created a set of mutually uncorrelated components of the data. Intuitively, the first principal component is the linear index that captures the most common variation among the components obtained (Filmer & Pritchett, 1999). STATA's 6 factor command was used, and the principal components' option within the command was specified.

All community indicator variables were based on the 1991 population census, 1996 population counts, and other district level information, such as the number of school enrolments and unmet need for services such as health care centres in each district. Most indicators are summarized in an index developed for each district called the 'Index of Social Exclusion/Inclusion' or IEX_IN (Sposati, 1996, 2000). The

Table 2. Factor loadings for first and second components

Indicator	Factor loadings First factor	Factor loadings Second factor
Percentage of households headed by females	0·828	0·402
Average monthly income	0·828	− 0·380
Number of new positions opened in the job market in the district	0·777	0·298
IEX_IN for investment	0·503	− 0·302
Average number of persons per bedroom	− 0·960	0·173
Average number of sleeping bedrooms per household	0·552	− 0·748
Average number of persons living in the household	− 0·931	− 0·267
Number of people living on streets	0·495	0·596
Index of disparity for precarious residence	− 0·257	− 0·030
IEX_IN for violence	0·050	− 0·013
IEX_IN for illiteracy of household head	0·938	0·030
IEX_IN for low level of education of household head	0·969	− 0·090
IEX_IN for high level of education of household head	0·631	0·417
IEX_IN for very high level of education of household head	0·892	− 0·330
IEX_IN for access to basic health care facilities	− 0·260	0·088
IEX_IN for access to day care centres	0·081	0·577
IEX_IN for unmet need for school for children	− 0·025	0·050

Source: Sposati (1996) and Sposati (2000).

interpretation of each index is that the higher its value, the higher the average degree of 'social inclusion' of its population into the society, i.e. the better off the population of the district is in comparison to all other districts. Therefore, the interpretation is quite intuitive: the higher, the better.

The first principal component explained 43·2% of the variation in these eighteen variables, which is a substantial percentage. The first component serves as a reasonable overall index and it correlated highly and positively with the IEX_IN for household head level of schooling (any), average household income, average rate of employment, average number of bedrooms per house, and very highly and negatively with number of persons living in the same household. Also, it correlated highly and positively with percentage of women that are household heads. The second component is less interpretable, as can be seen in Table 2. With the first component, therefore, an index has been obtained that summarizes dimensions of education, employment and income, and also household 'crowding', and to some extent, women's autonomy (Table 2). The higher the index, the better off is the district. After the index was constructed the infants were sorted by 'district development'. The 40% who scored lowest were categorized as living in a 'district with poor development', the middle 40% as living in a 'district with medium development' and the highest scoring 20% were categorized as living in a 'highly developed district' (Filmer & Pritchett, 1999).

Univariate analysis

The distribution for infants at risk of dying in the neonatal period from the 1998 birth cohort is presented in Table 3. Briefly, most infants had high Apgar scores, were males, singletons, weighed between 3000 and 3500 g at birth, were full-term, and were delivered vaginally (though a very high percentage of infants, 46·2%, were delivered by Caesarean section). Most infants who had race/colour recorded were white, but 66·3% of the infants did not have this information recorded. Most women did not have any prior losses, had more than seven prenatal care visits (but the percentage is not overwhelming since 32·4% of women had fewer than seven visits), delivered their first child, were in the age group 20–24 and had only incomplete elementary school education. The educational level of these mothers may be considered low: even though 79·1% of all mothers were at least 20 years old, an age by which every person, theoretically, could have finished secondary school, 37% of all mothers were illiterate or did not have an elementary school level of education.

Multivariate analysis

The odds ratio is a key measure in many epidemiologic analyses and represents a ratio between the probability that an event occurs and the probability the event does not occur. It measures the relative magnitude of two sets of odds occurring under differing conditions, and it varies from zero to infinity (Hosmer & Lemeshow, 1989). To obtain an adjusted odds ratio, a multivariate logistic regression was used to model the dichotomous outcomes under study. Because characteristics of mothers and infants from the same community were related, the standard errors for lack of independence between observations were corrected using the Huber/White Sandwich correction, which assumes that observations are independent across clusters but not within clusters (the community of mother's residence at the time of birth) (Liang & Zeger, 1993). STATA was used for the regression analysis and the cluster option within the logit command was selected. This approach does not change the coefficients but takes into account the clustering in the covariance matrix.

The ratio of two odds can vary by chance alone. Therefore, the p value indicates the probability that the two sets of odds occurring under differing conditions are equal to one, adjusted for all other covariates. The asterisks next to the odds ratio values presented in the tables indicate the significance of the Wald test of whether the odds ratio is equal to one (Hosmer & Lemeshow, 1989).

Results

The results are presented in Table 4. As expected, female infants had a lower odds of dying in the neonatal period, as compared with male infants (OR = 0·85, $p<0{\cdot}01$). Non-singleton births had a significantly lower odds of death in the neonatal period (OR = 0·66, $p<0{\cdot}01$). It is important to remember that after controlling for birth weight and gestational age, twins or higher order births are expected to have a better chance of survival (Guo, 1993). Caesarean section did not exert any effect on mortality as compared with vaginal deliveries. Other method of delivery substantially reduced the odds of neonatal mortality (OR = 0·57, $p<0{\cdot}001$). As compared with

Table 3. Univariate description of infants: 1998 birth cohort, São Paulo, Brazil

Variable	Categories	Percentage
Infant's sex	Male	50·9
	Female	49·1
	Missing	0·0
Plurality	Singleton	96·6
	Twins/higher order	1·9
	Missing	1·5
Delivery mode	Normal	49·2
	Caesarean	45·4
	Other	3·9
	Missing	1·5
Prior losses	None	57·5
	One	9·4
	Two or more	2·9
	Missing	30·2
Apgar at 1 min	0–6	7·6
	7 and 8	49·9
	9 and 10	38·0
	Missing	4·5
Apgar at 5 min	0–6	1·2
	7 and 8	8·7
	9 and 10	85·8
	Missing	4·3
Weeks gestation	27 or less	0·5
	28–36	5·2
	37–41	87·4
	42 or more	1·1
	Missing	5·8
Prenatal visits	None	2·1
	1–6	29·9
	7 or more	34·9
	Missing	33·1
Birth weight (g)	999 or less	0·4
	1000–1499	0·8
	1500–1999	1·6
	2000–2499	6·0
	2500–2999	24·6
	3000–3499	41·2
	3500–3999	20·1
	4000–4499	3·7
	4500 or higher	0·5
	Missing	1·1

Table 3. *Continued*

Variable	Categories	Percentage
Parity	None	35·2
	One	24·6
	Two or three	16·7
	Four or more	4·1
	Missing	19·4
Maternal age	11–14	0·5
	15–19	16·6
	20–24	28·5
	25–29	25·7
	30–34	18·2
	35–39	8·0
	40 or higher	2·0
	Missing	0·5
Race/colour	White	23·1
	Non-white	10·4
	Missing	66·5
Maternal education	No schooling	1·4
	Elementary, incomplete	35·5
	Elementary, complete	17·3
	Secondary	17·4
	College	7·7
	Missing	20·7
Community development	Low	38·5
	Medium	41·1
	High	20·4

Note: The frequencies were weighted according to the weight given to each matched pair.
Source: DATASUS, 2000.

infants with Apgar scores of 7 and 8 at 1 minute, infants with lower Apgar scores at 1 minute had a higher odds of death (OR = 1·83, $p<0·001$). Infants with scores at 1 minute of 9 and 10 had odds of dying that were lower (OR = 0·77, $p<0·001$). Infants with low Apgar scores at 5 minutes, compared with infants with scores of 7 and 8, had odds of death in the neonatal period that were substantially higher (OR = 4·21, $p<0·001$). Infants with highest 5-minute scores had odds of death that were lower as compared with the reference group (OR = 0·56, $p<0·001$). Pre-term infants (28–36 weeks), and very pre-term infants (<28 weeks) had higher odds of death in the neonatal period (OR = 2·37, $p<0·001$ and OR = 1·37, $p<0·01$, respectively). Post-maturity, on the other hand, was not associated with neonatal death. Prenatal care did not show a significant association with neonatal death. Infants below 3000 g had a higher odds of dying than their 3000–3499 g infant counterparts. The weight for

Table 4. Adjusted odds ratios for neonatal death: 1998 birth cohort, São Paulo, Brazil

Variable	Categories	Adj. odds ratio
Infant's sex (ref.: male)	Female	0·845*
Plurality (ref.: singleton)	Twins/higher order	0·663*
	Missing	0·595
Delivery mode (ref.: normal)	Caesarean	0·997
	Other	0·565**
	Missing	0·523*
Prior losses (ref: no prior loss)	One	1·130
	Two or more	1·054
	Missing	1·054
Apgar at 1 min (ref.: 7 and 8)	0–6	1·832**
	9 and 10	0·766**
	Missing	0·822
Apgar at 5 min (ref.: 7 and 8)	0–6	4·212**
	9 and 10	0·562**
	Missing	1·780
Weeks gestation (ref.: 37–41)	27 or less	2·367**
	28–36	1·381*
	42 or more	1·236
	Missing	1·183
Prenatal visits (ref.: 7 or more)	None	1·057
	1–6 visits	1·153
	Missing	1·151
Birth weight (g) (ref.: 3000–3500)	999 or less	61·53**
	1000–1499	21·16**
	1500–1999	7·323**
	2000–2499	2·843**
	2500–2999	1·414**
	3000–3499	1·048
	4000–4499	0·835
	4500 or higher	1·321
	Missing	5·450**
Parity (ref.: none)	One	1·072
	Two or three	1·060
	Four or more	1·166
	Missing	0·957

https://doi.org/10.1017/S0021932004006595 Published online by Cambridge University Press

Table 4. *Continued*

Variable	Categories	Adj. odds ratio
Maternal age (ref.: 20–24)	11–14	1·551
	15–19	1·167
	25–29	1·019
	30–34	0·879
	35–39	0·881
	40 or higher	0·579**
	Missing	3·128**
Race/colour (ref.: white)	Non-white	0·987
	Missing	1·027
Maternal education (ref.: elem., incomplete)	No schooling	1·239
	Elementary, complete	1·059
	Secondary	0·924
	College	0·936
	Missing	1·199
Community development (ref.: high)	Low	1·330**
	Medium	1·593**

Note: Data weighted according to the weight given to each matched pair.
*$p<0{\cdot}01$; **$p<0{\cdot}001$.
Source: DATASUS, 2000.

which the lowest mortality was observed was in the range 4000–4499 g, and this slightly decreased the odds of death, but the effect was not significant. Also, infants with missing information on birth weight had a higher odds of dying in the neonatal period (OR = 5·43, $p<0{\cdot}001$). This indicates a higher percentage of low birth weight infants in this category. There was no significant effect of parity on the odds of dying in the neonatal period, but the direction and magnitude of the odds ratio increased from lower to higher parities, which suggests that the higher the number of siblings the higher the risk of neonatal death. Older motherhood significantly decreased the odds of neonatal death (OR = 0·58, $p<0{\cdot}001$). This finding suggests that in the presence of most proximate factors, infants of older mothers do better. Since a very low weighed percentage of infants who died had missing information on maternal age, very high and very low odds ratios for the missing maternal age category may be seen as a result of instability of coefficients due to small numbers.

No significant results were found regarding maternal education and race.

Finally, having a mother who lives in the highest developed community reduced the odds of neonatal death. It was expected that, after controlling for all factors more proximate than community development, the effect of this variable on mortality would disappear. The rationale is that if the idea of proximate determinants is correct, and the factors are exhaustively measured, other factors should not have significant effects once proximate determinants are controlled, since every single distal determinant should work through one or more proximate determinants (Mosley & Chen,

https://doi.org/10.1017/S0021932004006595 Published online by Cambridge University Press

1984). Instead, it was found that living in a community with medium development increased the odds of neonatal death as compared with living in the highest developed community (OR = 1·33, $p<0{\cdot}001$). Also, the effect of living in the lowest developed community as compared with the highest developed one, was to increase the odds of death (OR = 1·59, $p<0{\cdot}001$). The result suggests that other proximate determinants or other indicators of socioeconomic status of the mother remained uncontrolled for, or a net effect, over and above proximate factors, of community development.

Discussion and conclusions

In this article two datasets were probabilistically matched in order to study determinants of infant mortality in the city of São Paulo, Brazil. Even though a one-to-one relationship could not be achieved for every death record, it was possible for most death records (74%). Further studies need to be conducted using probabilistic matching to make use of the data available in the SINASC and SIM Information Systems in a combined way, in order to obtain generalizable results for geographically defined sub-populations in Brazil. Population-based studies are a first necessary step for further in-depth investigations of determinants of neonatal mortality.

On considering the substantive results, the effect of older motherhood was to increase the chances of survival. It is suggested that, after controlling for proximate factors that were associated with disadvantages of infants of older mothers, such as number of prior losses, method of delivery and prenatal care utilization, infants of older mothers have a better chance of neonatal survival. Indeed, though much is assumed about the disadvantages of teenage mothers, motherhood in the early twenties is also likely to be disadvantageous as compared with older motherhood (Hobcraft & Kierman, 2001). Older mothers are more likely to be married, and to have fewer mistimed pregnancies. A pathway can be suggested: women with intended pregnancies tend to be older and more likely to either initiate breast-feeding and also to continue breast-feeding, as compared with mothers with unintended pregnancies, who tend to be younger (Taylor & Cabral, 2002). Older mothers may also be more likely than young mothers to value highly continuity of prenatal care and comprehensive care (O'Malley & Forrest, 2002), and more likely to attend more prenatal care visits, which reduce morbidities throughout the pregnancy period.

There was no effect of Caesarean section on neonatal mortality in these data. The use of forceps, as compared with vaginal delivery, increased the odds of survival. This finding makes sense, since in the hands of a skilled and experienced physician, the use of forceps can hasten the delivery and alleviate possible fetal distress. Its use, then, is ultimately beneficial to neonatal survival.

Birth weight, gestational age, and Apgar scores were the most important predictors of neonatal survival, as expected.

Community development remained significant in the presence of all variables used in this study. The community variable used can be seen as an index that comprises overall quality of life, income level and investment level in the community, as well as education of the household head. It also measures the average household density. More developed communities are more likely to be better served by sanitation connections, which improve infant survival. Community infrastructure may also improve hygienic practices (Sastry, 1996). Interactions between friends and neighbours in communities may also lead to changes in behaviour regarding infant care, and in this sense better-off communities may benefit from an overall level of community education. Also, the community development effect seems to be measuring either unobserved effects related to the pregnancy (for example, maternal smoking or maternal nutrition) and/or unobserved factors related to the conditions experienced by the newborn child within the home. These are all suggested pathways through which community development may affect infant survival that were not captured through the variables used in this study, and may all be reflected in the community variable. It might be possible to move a step further and suggest that where the household head is also the mother of the child, an 'absent father' effect is being observed. Qualitative studies might be the next step in order to understand and minimize health disparities among neonates of different communities in this population.

Acknowledgments

The authors thank the Brazilian Agency for Post-Graduate Education (CAPES), process number 2166/97-6, and two anonymous referees for their very thoughtful comments.

References

Bai, J., Wong, F. W. S., Bauman, A. & Mohsin, M. (2002) Parity and pregnancy outcomes. *American Journal of Obstetrics and Gynecology* **186**, 274–278.

Barros, F. C., Victora, C. G. & Horta, B. L. (2001) Ethnicity and infant health in Southern Brazil. A birth cohort study. *International Journal of Epidemiology* **30**, 1001–1008.

Bianco, A., Stone, J., Lynch, L., Lapinski, R., Berkowitz, G. & Berkowitz, R. L. (1996) Pregnancy outcome at age 40 and older. *Obstetrics and Gynecology* **87**, 917–922.

Clark, S., Blum, J., Blanchard, K., Galvao, L., Fletcher, H. & Winikoff, B. (2002) Misoprostol use in obstetrics and gynecology in Brazil, Jamaica, and the United States. *International Journal of Gynecology and Obstetrics* **76**, 65–74.

Cowden, A. J. & Funkhouser, E. (2001) Adolescent pregnancy, infant mortality, and source of payment for birth: Alabama residential live births, 1991–1994. *Journal of Adolescent Health* **29**, 37–45.

Curtis, S. L., Diamond, I. & McDonald, J. W. (1993) Birth interval and family effects on postneonatal mortality in Brazil. *Demography* **30**, 33–43.

De Hart, M. P. (1994) Racial differences in Apgar scores and mortality and morbidity in low birthweight infants. ScD Dissertation, Johns Hopkins University, Baltimore, MD.

Diamond, I. (2000) Child mortality – the challenge now. *Bulletin of the World Health Organization* **78**, 1174.

Filmer, D. & Pritchett, L. H. (1999) The effect of household wealth on educational attainment: evidence from 35 countries. *Population and Development Review* **25**, 115–132.

Filmer, D. & Pritchett, L. H. (2001) Estimation wealth effects without expenditure data – or tears: an application of educational enrollment in the State of India. *Demography* **38**, 115–132.

https://doi.org/10.1017/S0021932004006595 Published online by Cambridge University Press

Giffin, K. (1994) Women's health and the privatization of fertility control in Brazil. *Social Science and Medicine* **39**, 355–360.

Guo, G. (1993) Use of sibling data to estimate mortality effect in Guatemala. *Demography* **30**, 15–32.

Haaga, J. G. (1989) Mechanisms for the association of maternal age, parity, and birth spacing with infant health. In *Contraceptive Use and Controlled Fertility.* National Academy Press, Washington, DC, pp. 96–139.

Hobcraft, J. (1993) Women's education, child welfare and child survival: a review of the evidence. *Health Transition Review* **3**, 159–175.

Hobcraft, J. & Kiernan, K. (2001) Childhood poverty, early motherhood and adult social exclusion. *British Journal of Sociology* **52**, 495–517.

Hosmer, D. W. & Lemeshow, S. (1989) *Applied Logistic Regression*. Wiley and Sons, New York.

Liang, K. Y. & Zeger, S. L. (1993) Regression analysis for correlated data. *Annual Review of Public Health* **14**, 43–68.

Machado, C. J. (2002) Early infant morbidity and infant mortality in the city of São Paulo, Brazil: a probabilistic record linkage approach. PhD Dissertation, Johns Hopkins University, Baltimore, MD.

Mahy, M. (1999) Perceptions of child mortality and their effects on fertility intentions in Zimbabwe. ScD Dissertation, Johns Hopkins University, Baltimore, MD.

Mosley, W. H. & Chen, L. C. (1984) An analytical framework for the study of child survival in developing countries. *Population and Development Review* **10** (supplement), 25–45.

Moss, W., Darmstadt, G. L., Marsh, D. R., Black, R. E. & Santosham, M. (2002) Research priorities for the reduction of perinatal and neonatal morbidity and mortality in developing countries. *Journal of Perinatology* **22**, 484–495.

O'Malley, A. S. & Forrest, C. B. (2002) Beyond the examination room: primary care performance and patient–physician relationship for low income women. *Journal of General Internal Medicine* **17**, 66–74.

Potter, J. E., Schmertmann, C. P. & Cavenaghi, S. M. (2002) Fertility and development: evidence from Brazil. *Demography* **39**(4), 739–761.

Rutstein, S. O. (2000) Factors associated with trends in infant and child mortality in developing countries during the 1990s. *Bulletin of the World Health Organization* **78**, 1256–1270.

Sastry, N. (1996) Community characteristics, individual and household attributes, and child survival in Brazil. *Demography* **33**, 211–229.

Scheuren, F. & Winkler, W. E. (1993) Regression analysis of data that are computer matched – Part 1. *Survey Methodology* **19**, 39–58.

Sposati, A. (ed) (1996) *Mapa da exclusão/inclusão social da cidade de São Paulo*. Educ, São Paulo.

Sposati, A. (ed) (2000) *Mapa da exclusão/inclusão social da cidade de São Paulo – Dinâmica social dos anos 90.* Educ, São Paulo.

Taylor, J. S. & Cabral, H. J. (2002) Are women with an unintended pregnancy less likely to breastfeed? *Journal of Family Practice* **51**, 431–436.

Villar, J. & Bergsjo, P. (1997) Scientific basis for the content of routine antenatal care. I. Philosophy, recent studies, and power to eliminate or alleviate adverse maternal outcomes. *Acta Obstetricia et Gynecologica Scandinavica* **76**, 1–14.

Wells, J. C. K. (2000) Natural selection and sex differences in morbidity and mortality in early life. *Journal of Theoretical Biology* **202**, 65–76.