

The undesirable consequences of controlling for birth weight in perinatal epidemiological studies

Eve Blair

Abstract

Objective – To compare the effects of controlling for birth weight with those of controlling for gestational age at delivery in perinatal epidemiological studies using two examples.

Setting – Western Australia.

Subjects – Population data: all white births born at 20–46 weeks of gestation in Western Australia during 1985–91 inclusive (n = 147 564). Example 1: All Western Australian births from 1980–89 born either at 33–36 weeks inclusive (n = 13 607), or born with a birth weight of 2050–2900 g (n = 34 107). Example 2: 160 singleton cases of spastic cerebral palsy born to white mothers in Western Australia from 1975–80 and whose gestational age was known, compared with (a) 480 controls individually matched for gender and birth weight and (b) singletons with known gestational age liveborn to white mothers in Western Australia from 1980–81, or 1979–82 if <30 weeks' gestational age at birth (n = 32 031).

Measurements and main results – The risks of cerebral palsy associated with two separate exposures in groups defined by birth weight were compared with those in groups defined by gestational age. The origin of the differences are explained using total population data. The estimates of risk differ when exposure and outcome are both associated with appropriateness of fetal growth. The difference varied with gestational age, being greatest in the moderately preterm (33–36 weeks' gestation).

Conclusion – Epidemiological studies in which appropriateness of fetal growth is an important variable should be based on gestational age at birth rather than birth weight, whatever the neonatal size or maturity.

(*J Epidemiol Community Health* 1996;50:559–563)

This paper concerns three variables of paramount importance in perinatal epidemiology: birth weight, gestational age at birth (GAB), and the appropriateness of fetal growth rate (AFGR). It concludes that since birth weight is biologically dependent on both GAB and AFGR, while the latter two are biologically independent of each other, studies investigating exposures associated with AFGR should be defined in terms of GAB rather than birth weight.

Since no one would agree that cases and controls should be matched for outcome, this

conclusion may seem self evident, but in practice it is frequently overlooked. Eight years ago I had collected data pertaining to controls matched for birth weight in order to investigate the role of AFGR in the aetiology of cerebral palsy. However, children with cerebral palsy tend to be born preterm and hence have a reduced birth weight, whether or not they have had appropriate intrauterine growth rates. Controls were constrained to match a case in terms of birth weight, but since they were selected from the total population in which more than 93% were born at or after term, the matching of birth weight selected a group likely to have the combination of low birth weight and born at term and hence be growth retarded. In birth-weight matched analysis, poor growth appeared protective against cerebral palsy. I was embarrassed at having made such an obvious error, yet it was an error shared by all the experts to whom I had shown my protocol and by many researchers whose papers had come to publication before, and since. Once one has matched in terms of birth weight, it is not possible to correct for GAB instead in analysis, so collection of these data may entail much wasted effort if exposure variables prove to be associated with AFGR. In my study I had to select new GAB matched controls. These two analyses are compared in Example 2.

When control for maturity is considered necessary, the reasons for collecting birthweight defined rather than GAB defined data are compelling.

1. Birth weight is an easily, accurately, and, in developed nations, almost completely available perinatal variable. In contrast, GAB is difficult to measure accurately, is not always available, and when available its accuracy, particularly in population based data collections, is often unknown. If missing GAB is associated with the factor of interest, selection of those with known GAB may give misleading results.¹

2. The biological dependence of birth weight on AFGR is not always appreciated. Birth weight is directly measurable and is the end result of biologically integrating the (varying) rate of weight gain over the total duration of gestation. Optimal rate of weight gain varies between individuals, thus it is the rate of weight gain *relative to this optimal rate* that is considered to be strongly associated with fetal health and therefore of greater interest epidemiologically than the absolute rate of weight gain. This optimal rate is not directly observable, but the *appropriateness* of rate of weight gain (from conception to birth) may be inferred at birth by comparing an infant's birth weight with that of other babies of the same GAB and with the

Institute for Child
Health Research, GPO
Box 855, West Perth
6872, Western
Australia
E Blair

Correspondence to:
E Blair.

Accepted for publication
April 1996

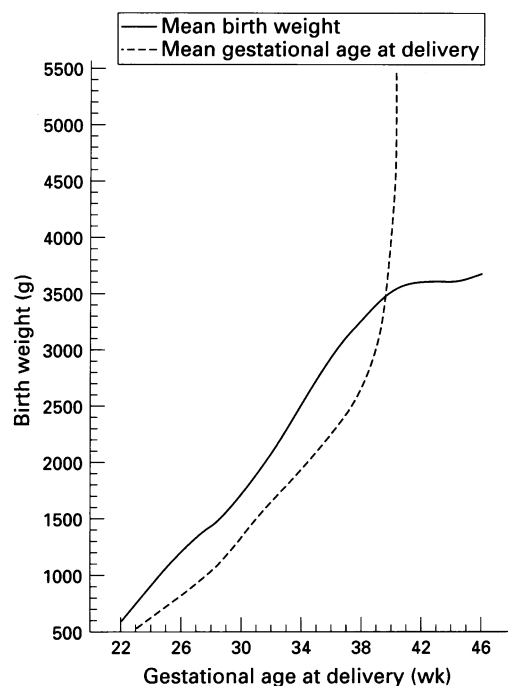


Figure 1 Mean birth weight in relation to gestational age week at delivery and mean gestational age at delivery in relation to 100 g birthweight band. (Western Australia midwives' data 1985–91.)

same non-pathological characteristics determining growth, such as gender and maternal parity and height. Common measures of AFGR are therefore (i) the centile position on the birth weight distribution of babies of the same GAB and characteristics and (ii) the ratio of the birth weight to the median birth weight of babies of the same GAB and characteristics. Thus, the value assigned to AFGR in epidemiological studies is obtained from birth weight and GAB, masking the biological dependence of birth weight on AFGR and the independence of AFGR and GAB.

3. Even if it is accepted that GAB is the defining factor of choice, figure 1 shows that mean birth weight is highly correlated with GAB over much of the GAB range. This has led many researchers, faced with the difficulties of obtaining good GAB data, to consider birth weight an acceptable surrogate measure of maturity and to control for birth weight by individual birthweight matching, by stratifying or by defining the sample in terms of birth weight. It has already been suggested² that this is unsatisfactory in studies of very small or immature

infants whenever fetal growth is an important variable. This paper examines the effects of substituting control of birth weight for control of GAB in perinatal epidemiological studies investigating factors associated with AFGR regardless of newborn weight or age by using population based data.

Methods

EXAMPLES IN WHICH THE ASSOCIATION BETWEEN TWO VARIABLES DIFFERS BETWEEN BIRTH WEIGHT AND GAB DEFINED SAMPLES

Example 1

The risk of cerebral palsy associated with hypertensive disease of pregnancy (HDP) was investigated in two samples drawn from the Western Australian maternal and child health database for livebirth cohorts 1980–89.³ The first sample was defined by birth weight (2050–2900 g); the second was defined by GAB (33–36 weeks). This example was chosen because both HDP and cerebral palsy are associated with poor fetal growth,⁴ and the cut off points were chosen because the median birth weights are 2050 g at 33 weeks and 2900 g at 36 weeks in this population.

Table 1A shows that there was no association between HDP and cerebral palsy in the cohort defined by GAB, the odds ratio (95% confidence interval) being 1.03 (0.5,2.2). Table 1B shows a significant protective effect of HDP in the birth weight defined cohort, with an odds ratio of 0.52 (0.32,0.85).

Example 2

The difference between associations for any given population depends on the correlation between the exposure variable and AFGR. For any outcome, the maximum difference will therefore be obtained when AFGR is the exposure variable, the example mentioned in the introduction.

The distribution of centile positions on the gestation-specific birthweight distribution (*the birthweight centiles*) of 160 white, singleton cases of spastic cerebral palsy with known gestational age was compared with that of (a) 480 controls with known gestational age, 3 controls being matched to each case on gender and birth weight and (b) the total Western Australian singleton population with known gestational age, liveborn to white mothers in Western Australia in 1980–81 or 1979–82 if <30 weeks GAB.⁵ Since the birthweight centiles are specific to the week of GAB, comparison with an unselected population effectively matches on week of GAB. The birthweight centile distributions are compared in table 2 and show that for a single group of cases, a birth weight centile below the third seems to reduce the risk of cerebral palsy when birth weight matched controls are considered, but greatly increases the risk when compared within week of GAB.

Table 1 Distribution of hypertensive disease of pregnancy and cerebral palsy. Western Australian livebirths 1980–89

(A) born between 33 and 36 weeks' gestational age at delivery inclusive		
	Cerebral palsy	
Hypertensive disease of pregnancy	+	–
+	8	1887
–	48	11664
Odds ratio = 1.03; 95% confidence interval 0.50, 2.2		
(B) born between 2050 g and 2900 g birth weight inclusive		
	Cerebral palsy	
Hypertensive disease of pregnancy	+	–
+	7	3571
–	114	30415
Odds ratio = 0.52; 95% confidence interval 0.32, 0.85		

Table 2 Distribution of appropriateness of fetal growth rate, estimated by centile position of the gestation week-specific birthweight distribution of white Western Australian singleton livebirths⁵

	Birthweight centile				Total no
	<3rd	3rd-10th	10th-90th	>90th	
Population:					
Cerebral palsy cases % (no)	15.0 (24)	13.4 (23)	65.0 (104)	5.6 (9)	160
Birthweight matched controls % (no)	22.1 (106)	31.8 (66)	59.2 (284)	5.0 (24)	480
Total population %	3	7	80	10	32 031
Odds ratio (birthweight matching)	0.62	0.95	1	1.02	
(95% confidence interval)	(0.23,1.64)	(0.34,2.66)	Reference	0.21,4.90	
Odds ratio (total population)	6.2	2.5	1	0.69	
(95% confidence interval)	(2.55,14.8)	(1.04,6.16)	Reference	0.18,2.64	

THE ORIGINS OF DIFFERENCE IN ODDS RATIOS

Figure 1 shows that while there is a high correlation between birth weight and GAB for much of the range of birth weight and GAB, the approximately linear relationship breaks down above 3000 g birth weight or 37 weeks' GAB, and thus is not applicable to the majority of births. Figure 1 also shows that the mean distribution of birth weight for GAB is not the same as the mean GAB for birth weight. Furthermore, while there is a near Gaussian distribution of birth weight around the mean birth weight of each GAB week, this is seldom true for the GAB distribution around the mean GAB in each birthweight band because, as shown in figure 2, there is a very peaked distribution of GAB for the total population. The vast majority of births occur close to 40 weeks of gestation (at term). Therefore, in any birthweight stratum there is going to be a disproportionate number of infants born at term and since term infants in birthweight strata below the term median birth weight are relatively poorly grown, random selection of infants from low birthweight strata will result in a group biased towards poor intrauterine growth. The reverse is true of high birthweight strata, where randomly selected birthweight matched controls will be biased towards above average intrauterine growth.

MEASURING THE BIAS IN AFGR ENGENDERED BY CONTROLLING FOR BIRTH WEIGHT

The degree of bias varies with both GAB and AFGR. The method of measuring bias is illustrated by the following simplified example

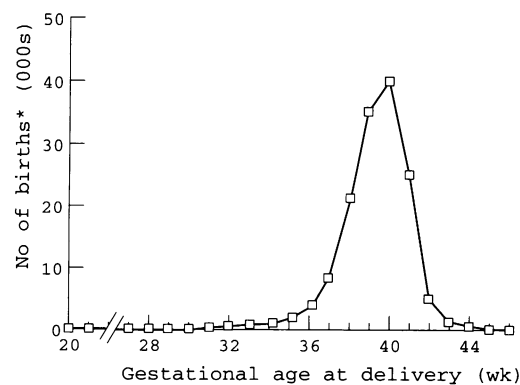


Figure 2 Distribution of gestational age at delivery for white Western Australian births, 1985-91.

of the estimation for GAB = 36 weeks. Figure 3A shows that in the white Western Australian population born in 1985-91, the median birth weight was 2899 g at GAB = 36 weeks. If non-pathological characteristics determining growth are ignored and it is assumed that no infants have precisely the same AFGR (simplifications not used in the actual calculations), then 50% of births will be more poorly grown and 50% better grown than an infant of birth weight 2899 g born at 36 weeks' gestation. In order to assess the probability distribution of GAB at any point birth weight from our sample, the width of a band of birthweight values centred on the birth weight under consideration was chosen so as to include at least 100 births. At 2899 g the necessary birthweight band width in our sample was 40 g (2879-2918 g). Figure 3B group (i) shows the observed GAB distribution of infants of birth weight 2879-2918 g. Most births (92.5%) within this birthweight band have a GAB >36 weeks and hence (if we consider 2879-2918 g as an indivisible unit of birth weight for the purposes of this example) a lower AFGR: ie, their centile is below the median. I express the degree of bias introduced by birthweight matching as: the birthweight/GAB matching ratio. This is the:

Probability of choosing a more poorly grown control with birthweight matching/Probability of choosing a more poorly grown control with GAB matching

$$= 0.925/0.5 = 1.85,$$

whereas if birthweight matching were equivalent to GAB matching with respect to AFGR the ratio would be unity.

This ratio may be calculated for the matching of a subject of any birthweight centile: eg, figure 3A also shows that the 10th birthweight centile at 36 weeks is 2299 g. Figure 3B, group (ii) shows the GAB distribution of the 50 g birthweight band around 2299 g and it can be seen that, making the same simplifying assumptions as above, (100-48.5) 51.5% have a greater GAB and hence have a lower birthweight centile, giving a birthweight/GAB matching ratio of (0.515/0.1) 5.15.

The degree of variation in bias is illustrated in figure 4 by presenting the ratio in each GAB week for a subject at the 50th birthweight centile and at the 3rd birthweight centile. The ratio varies (i) with GAB in a complicated manner, being most extreme in moderately

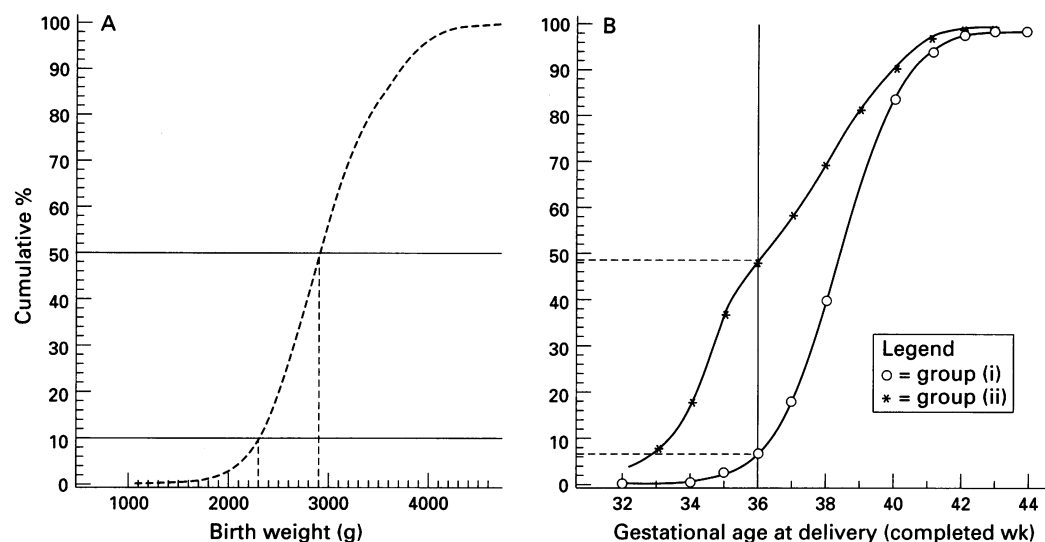


Figure 3 Estimation of the birthweight/GA matching ratio at 36 weeks of gestation. A. Birthweight distribution of infants born at 36 weeks. B. Gestational age distribution of infants of (i) birth weight 2879–2918 g (50th birthweight centile at 36 weeks) and (ii) birth weight 2274–2323 g (10th birthweight centile at 36 weeks).

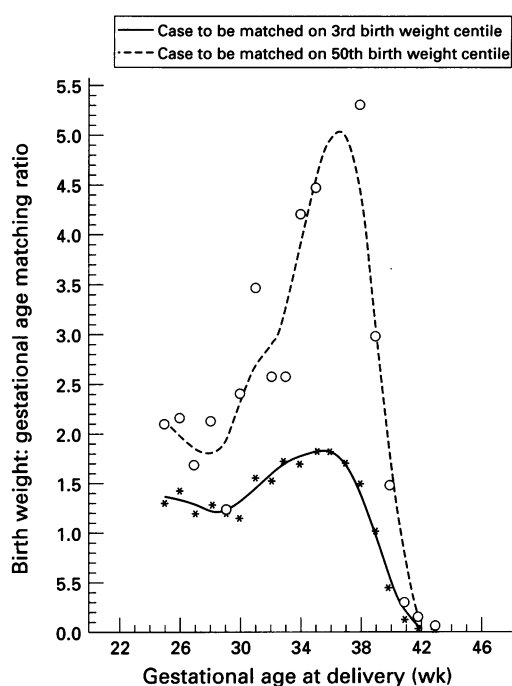


Figure 4 Birthweight/gestational age matching ratio in relation to gestational age at delivery for cases to be matched at the 50th or 3rd birthweight centile. (Western Australia midwives' data 1985–91.)

preterm infants, (ii) with the AFGR of the subject to be matched and (iii) with the range of GAB or birth weight considered to constitute a match (not illustrated), making post hoc correction inadvisable.

Discussion

Perinatally, AFGR has far reaching consequences, second only to GAB, so it is important to be aware that controls selected by matching of birth weight are biased with respect to AFGR. For unbiased control selection, the population of potential controls must have the same distribution of AFGR as that of the population from which the cases arose. This is true

for GAB matched controls by our definition of AFGR but is seldom true of birthweight matched controls.

A second reason to control for GAB rather than birth weight is that GAB matched results have a simple and valid interpretation: in Example 2 above, if two infants are born at the same GAB, the one with lower AFGR is at higher risk of spastic cerebral palsy. However, the same low birth weight may be seen in a preterm infant with appropriate growth or in a poorly grown infant born at term. Both infants are disadvantaged, but have distinct diseases with different aetiologies and are subject to different continuing risks. As previously recognised in very low birthweight infants,² matching on birth weight controls for neither GAB nor AFGR but partially controls for both so that the independent contributions of GAB and AFGR cannot be untangled.

WHEN IS CONTROL FOR MATURITY APPROPRIATE?

The rapid development of fetal life makes independence between succeeding events less likely in perinatal than postnatal life. Since preterm birth is an event of such significance, GAB is very likely to be a step in the aetiological pathway toward any outcome,⁶ rendering control for maturity inappropriate when determining the contribution of exposure to outcome. In such cases the time order of events must be considered as well as the nature of the association between each pair of associated variables.⁷ However, the independent contribution of exposure to outcome is not the only goal of research. Researchers in this field are already well aware that prognosis is strongly associated with GAB but efforts directed towards reducing the incidence of preterm birth have had limited success. It is then reasonable to seek to develop strategies to minimise the detrimental effects of preterm birth or to enhance the accuracy of prognosis for a particular

infant after birth (once GAB is determined). In both these cases GAB stratified analysis is appropriate.

SUMMARY

When control for maturity at birth is desirable in studies of newborns of any weight or maturity, it is desirable to control for GAB rather than birth weight, firstly because results obtained by GAB matching are easily interpreted and secondly because birthweight matched controls are biased in respect of AFGR – a variable of far reaching importance. It is hoped that awareness of these consequences of birthweight matching will lead to greater efforts to improve the accuracy and availability of GAB data. As a first step, GAB should always be included as a variable in perinatal data sets and, since accuracy of the GAB estimate is largely determined by its source and timing,⁸ these data should also be included in the data

sets so the magnitude of the problem with GAB data quality can be evaluated.

I would like to thank my colleagues at the Institute for Child Health Research for their constructive criticisms.

- 1 Hall MH, Carr-Hill RA, Fraser C, Campbell D, Sampher ML. The extent and antecedents of uncertain gestation. *Br J Obstet Gynaecol* 1985;**92**:445–51.
- 2 Arnold CC, Kramer MS, Hobbs CA, McLean FH, Usher RH. Very low birth weight: a problematic cohort for epidemiological studies of very small or immature neonates. *Am J Epidemiol* 1991;**134**:604–13.
- 3 Stanley FJ, Croft ML, Gibbins J, Read AW. A population database for maternal and child health research in Western Australia using record linkage. *Paediatr & Perinatal Epidemiol* 1994;**8**:433–47.
- 4 Blair E, Stanley F. Intrauterine growth and spastic cerebral palsy. I. Association with birthweight for gestational age. *Am J Obstet Gynecol* 1990;**162**:229–37.
- 5 Blair E, Stanley F. *Intra-uterine growth charts*. Canberra: Commonwealth Dept of Health, Australian Government Publishing Service, 1985. Cat no 85 0184 3.
- 6 Kiely JL. Some conceptual problems in multivariable analyses of perinatal mortality. *Paediatr & Perinatal Epidemiol* 1991;**5**:243–57.
- 7 Blair E, Stanley F. When can cerebral palsy be prevented? The generation of causal hypotheses by multivariate analysis of a case-control study. *Paediatr & Perinatal Epidemiol* 1993;**7**:272–301.
- 8 DiPietro JA, Allen MC. Estimation of gestational age: implications for developmental research. *Child Development* 1991;**62**:1184–99.