CHAPTER 9

Evaluating the Effects of Associated Factors

To investigate the possibility that our three main groups of children may differ in their susceptibility to the effects of different associated factors, and to compare the magnitude of the effects of individual factors, we have carried out multiple regression analyses to identify the contribution which each of a large number of potential adverse factors makes to the total variance of the same 13 measures of performance as are shown in Tables 8.2 and 8.3, within each of our two extreme abnormal groups and our total random sample. This latter group includes all the cases in the random control group plus 'overlapping' cases which are also in the short-gestation or one of the two light-for-dates groups. The results for the random sample are therefore representative of the whole infant-school population (see Chapter 1, p. 3). No separate multiple regression analysis has been carried out for the cases in the rather light-for-dates group, since all the simpler analyses have shown that their results fall somewhere between those of the very light-for-dates and the random control groups. For reasons discussed in Chapter 2, it is true to say that the short-gestation group represents approximately the same degree of abnormality of the duration of intra-uterine growth as the very light-for-dates group represents of its rate.

The same 15 'confusing' factors as in Table 8.1 have been included in all three of these series of multiple regression analyses, with the addition of two extra maternal characteristics—her expectations of her child's social attainments and her neighbourliness. The first of these was rated on a three-point scale, using a technique developed at the Maudsley Hospital (Rutter, personal communication) to assess the mother's expectations of her child's ability to perform four age-appropriate tasks (*e.g.* dressing, going shopping). The mother's 'neighbourliness' was a measure of her relationships with her neighbours and her sociability, based on the answers she gave to a twelve-item self-rating questionnaire (Wallin 1954).

We have included one or two further factors in each of the series of multiple regression analyses in order to identify the effect of variations in intra-uterine growth. The details differ in the different groups of children, because they were originally defined by different criteria, some of which were specified in terms of intra-uterine growth. In the case of the random sample, we have included each child's gestational age, in days, and birthweight centile grading, in terms of the corrected sex-specific curves for the Newcastle population published by Neligan *et al.* (1974). The centile bands which we have selected as statistically most appropriate are —3, —10, —25, —75, —90, —97 and >97, each of which has been allotted a code (1 to 7). Of course we can only include cases whose gestational age was known (see Chapter 1, p. 3). In the analyses of both the short-gestation and the very light-for-dates groups, however, we have omitted the factor of gestational age (since this was specified in the criteria by which these groups were defined): but we have included birthweight as measured, in grammes rather than in centile form, since the former is a more sensitive indicator of variations in the rate of intra-uterine growth when the groups of children concerned are primarily defined in terms of gestational age-range (255 days or less, and more than 255 days with a weight below the 5th centile, respectively).

The prediction of performance may be considered in two ways: first by studying correlation coefficients and second by multiple regression analysis. The first is the simpler method, and involves calculating correlation coefficients between the measures of performance and each of the associated factors (predictive variables). Three sets of tables are available in the unpublished appendix showing these correlations for the random sample, the short-gestation and the very light-for-dates groups.

On the other hand, multiple regression, which is the procedure we have used, does not simply consider the correlations between the various associated factors (predictive variables) and subsequent performance (dependent variables), but also takes into account the intercorrelations between the associated factors themselves. An associated factor might be highly correlated with subsequent performance but have a relatively low predictive weight in regression analysis because it is duplicated, that is, highly correlated with another associated factor. The relative importance of each of the significant associated factors has been assessed by calculating its *coefficient of individual determination** in relation to each dependent variable.

It follows that conclusions drawn from the use of these two methods of predicting performance will not be identical or even necessarily similar. The simple correlation method considers each associated factor individually; the more complex regression method highlights those associated factors which are most predictive *as members of the group of associated factors.*

The results of this series of analytical procedures are very complex, and any attempt to interpret them without distortion is made very difficult by two facts:—

(1) There are different numbers of children available for analysis in each of the three groups studied (163 in the random sample, 50 in the short-gestation and 61 in the very light-for-dates group). The effects of these differences upon the results of tests of statistical significance are similar to those noted at earlier and simpler stages of the analysis, of course, but rather more difficult to keep in mind and allow for because of the complexity of the analysis at this stage.

(2) The sums of the significant effects of the individual independent variables included in our analyses account for very different proportions of the total variance of the scores for the individual dependent variables, ranging from 0.07 for the total gesture score in the random sample to 0.60 for the verbal IQ in the short-gestation group (see bottom section of Figures 11 to 13, where the magnitude of the multiple R^2 is indicated†).

We want to compare the magnitude of the effect of each of the associated factors (independent variables) upon each of the measures of performance (dependent

*The coefficient of individual determination is the product of the appropriate correlation coefficient and the standardised partial regression coefficient.

†The multiple R^2 is the sum of the coefficients of individual determination.

variables) and upon the sum of all the dependent variables in each of the three groups, and this is difficult to do with the raw results shown in the Appendix Tables (available from the Editors). We feel that the best way of facilitating the comparison without any distortion is to convert these results straightforwardly into a visual form, as we have done in Figures 11 to 13: the magnitude of the important coefficients of individual determination between all the pairs of independent and dependent variables within each group of children can be appreciated at a glance, and the proportion of the total variance for each dependent variable in each group which is accounted for by all these 18 or 19 independent variables put together can also be compared by looking at the bottom section, where the magnitude of the multiple R^2 is illustrated. It is clear that the sum of the effects of the 18 or 19 independent variables tested in this analysis tends to account for a greater proportion of the variance in the cognitive tests (particularly in the short-gestation group, in which it accounts for more than half the variance of the verbal and non-verbal IQ and the Language Quotient) than in the motor, behavioural and neurological assessments (in several of which it accounts for less than a quarter of the variance in all three groups).

We have also illustrated, in the right-hand portion of each of these three figures, the sum of the important coefficients of individual determination produced by each of the independent variables in relation to all 13 dependent variables. To facilitate comparison between the separate findings in the three groups of children, we have redrawn this portion of the three histograms and combined them in Figure 14, showing only those sums of coefficients which exceed 0.1 (and omitting 'gestational age', which could only be evaluated in the random sample).

If we start by looking at Figure 11, in which the results can be regarded as representative of the whole infant-school population of our City, the most striking feature is the importance of the relationship between the mother's care of her child and the results of a whole range of tests of performance. An important coefficient is illustrated in relation to nine of the 13 tests, and the magnitude of this coefficient is far greater than that of any other independent variable in the case of both verbal and performance IQ and language quotient and behaviour as reported by the teacher. Its magnitude is about equal to that of any other of the independent variables tested in three further instances (Holborn reading test, Bender-Gestalt test of visuo-motor abilities and gesture score). In contrast, and in sharp disagreement with the findings of all other surveys that we know of, the child's social class of origin (as determined by the father's occupation) produces a negligible effect in the shape of a very small coefficient in relation to only two dependent variables—performance IQ and height at seven years. The most obvious explanation for this unique finding is that we have included in our analysis a number of other associated factors, defined in rather more concrete terms, whose effects upon the child's family environment (and so upon his performance) have eclipsed those of the much less specific and indirect label of 'social class'. The effects of these other factors are broadly spread amongst the various dependent variables, so it is legitimate and helpful to look at their sum, as illustrated in the right hand column of Figure 11 and the left hand column of Figure 14. The other variables which can be regarded as making up the 'family factors' are, in order of magnitude of their coefficients, the child's ordinal position in the family, the mother's expectations of his social abilities, whether he was read to regularly, and the mother's extraversion and neighbourliness scores.

The only other factors which are associated with a sum of coefficients sufficiently large to be drawn in Figure 14 are grouped under the heading of 'biological factors'. Of these, the biggest sum of coefficients is associated with mother's height, and Figure 11 makes it clear that the whole of this effect is accounted for by the two dependent variables which are measures of physical growth—height at seven years and weight increment from birth to seven years.

The other three biological factors whose sum of coefficients exceeds 0.1 (Fig. 14) are the mother's age and the child's sex and birthweight. It is apparent from Figure

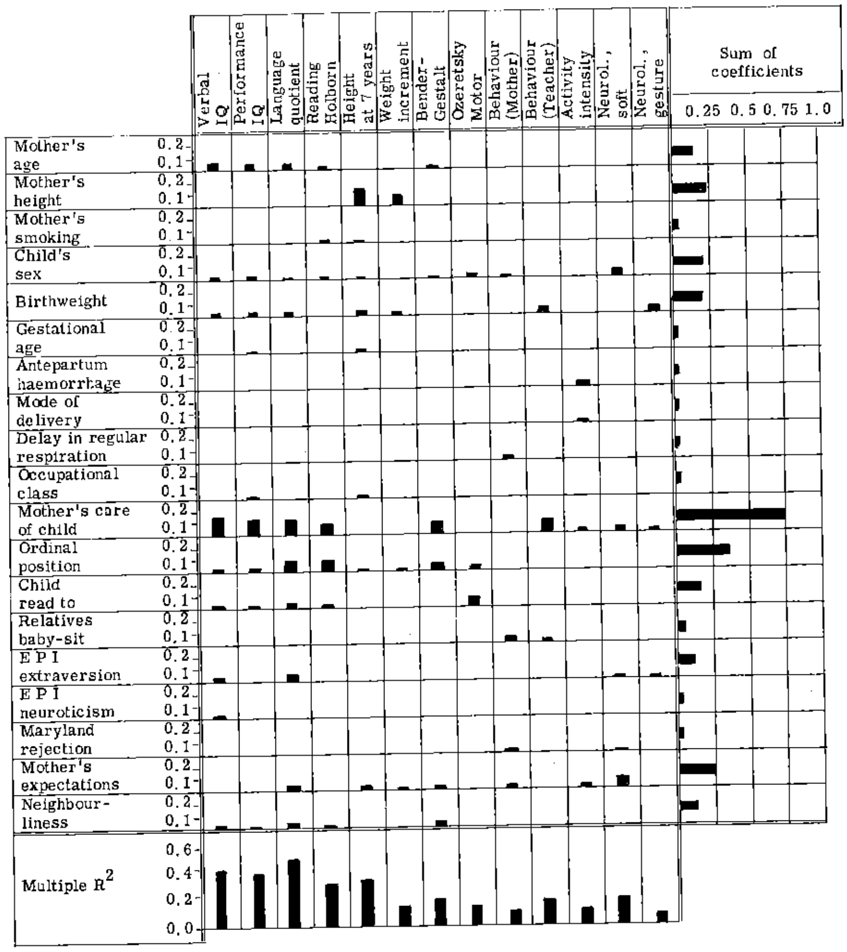

Fig. 11. Results of multiple regression analysis by step-wise procedure in random sample. Coefficients of individual determinations of all significant independent variables are shown.

11 that their effects are distributed much more widely, over five, nine and seven dependent variables, respectively. Gestational age and smoking during pregnancy are both associated with negligible effects (Fig. 11). The same applies to the three 'clinical factors' of antepartum haemorrhage, mode of delivery and delay in onset of regular respiration after birth.

When we look at Figure 12, it is obvious that the pattern of coefficients in the short-gestation group resembles in certain respects the one which we have just been studying. The biggest single sum of coefficients is associated with the mother's care of her child and is made up of important effects upon seven dependent variables: her expectations of the child's performance and whether he was read to are still important

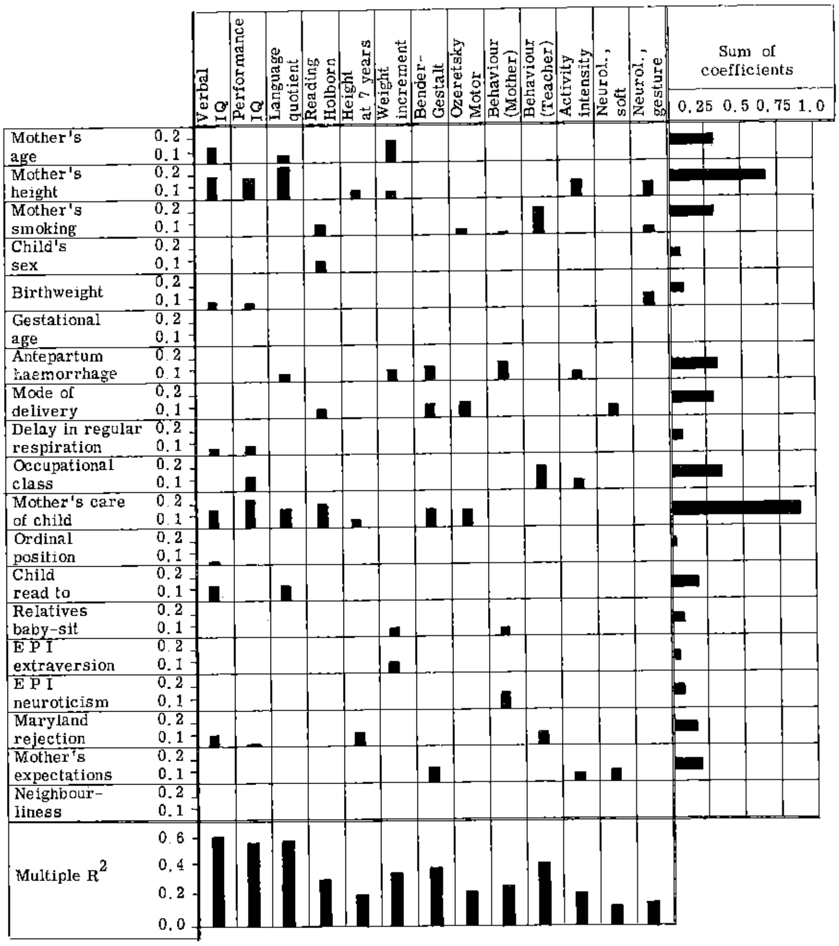

Fig. 12. Results of multiple regression analysis by step-wise procedure in short-gestation group. Coefficients of individual determinations of all significant independent variables are shown.

among the 'family factors'. However, this time the father's occupational social class makes an important contribution, and the mother's rejection score in the Maryland Parent Attitude Survey has displaced her neuroticism as the relevant component of her personality.

When we turn to the biological factors in the short-gestation group, it is obvious from the middle column of Figure 14 that they are making a far greater contribution, over-all, than they did in the random sample. And the very large sum of coefficients associated with mother's height is differently distributed in this case (see Fig. 12). Although the effect upon the two growth measures (height and weight increment) is still visible, it is entirely overshadowed by the effects upon the verbal and performance

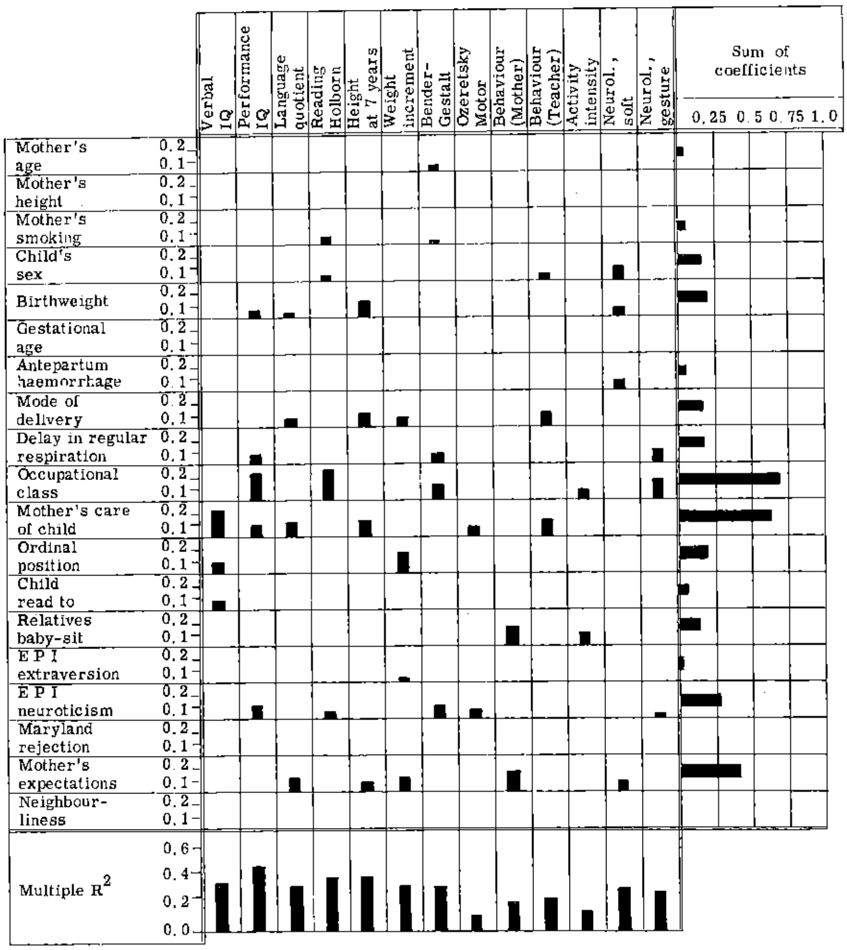

Fig. 13. Results of multiple regression analysis by step-wise procedure in very light-for-dates group. Coefficient of individual determinations of all significant independent variables are shown.

IQ and the language quotient: and there is a visible effect upon activity intensity and the gesture score.

Another new effect not seen in the random sample is that of smoking in pregnancy (Fig. 14), which is made up of a major effect upon the child's behaviour as reported by the teacher and detectable effects upon reading, the Ozeretsky Test of Motor Abilities, the child's behaviour as reported by the mother and the gesture score (Fig. 12). Furthermore, in this group of children, two of the 'clinical factors' (antepartum haemorrhage and mode of delivery) are associated with effects of considerable magnitude (Fig. 14), spread over five and four dependent variables respectively (Fig. 12).

Before we discuss and try to evaluate the differences between these results in the short-gestation group and those we noted in the random sample, it seems best to look at the very light-for-dates group in the same way. In Figure 14 it is obvious that the 'family factors' are playing an even more dominant rôle than in the other two groups, and that among them the trend towards an increase in the magnitude of the sum of coefficients associated with the father's occupation (noted in the short-gestation group) has raised this factor to a position of even greater importance than the mother's care of her child. Her expectations of her child's abilities have also assumed

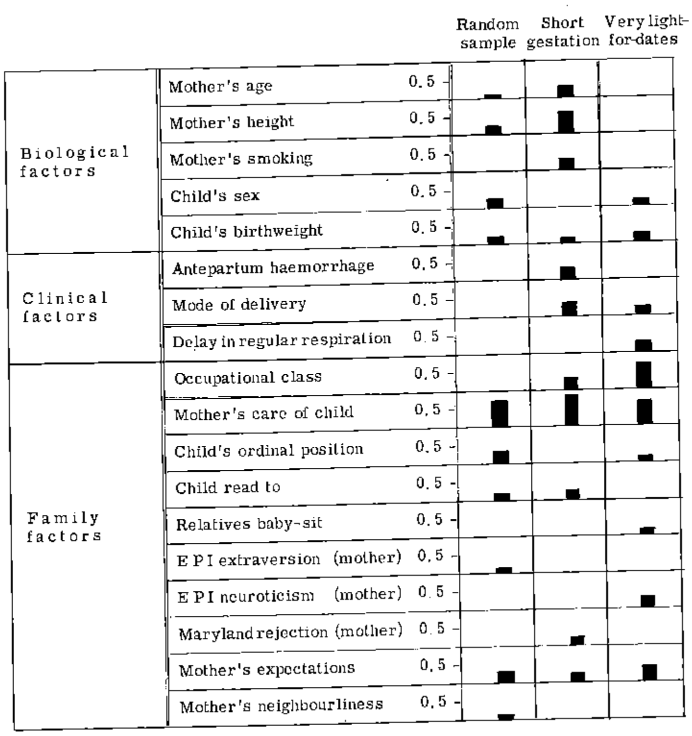

Fig. 14. Sum of coefficients greater than 0.1 from Figures 11, 12 and 13.

greater absolute importance than in the other two groups of children, and her neuroticism score and the extent to which relatives baby-sit have both assumed visible importance for the first time.

Among the biological factors, it is interesting to note that differences in birthweight are still of over-all importance, even in a group of children selected by a criterion which restricts their birthweights within a narrow range, and it is interesting to see that the biggest effect is upon the child's height at the age of seven years (Fig. 13). None of the maternal factors (age, height, smoking) has any observable effect in this group. As in the short-gestation group, two clinical factors are important, but delay in regular respiration has displaced antepartum haemorrhage as the effective partner of mode of delivery.

Discussion

Before attempting to draw any firm conclusions of practical significance from the results of our multiple regression analysis as summarised in Figures 11 to 14 we would like to be sure that they make sense when examined in the light of previous knowledge and of plausible mechanisms.

We have no doubt that on the whole they do make sense, particularly if we start by confining our attention to the random sample (Fig. 11), the group for which we have the greatest amount of relevant previous knowledge in the shape of results from geographically defined populations. In the group of biological factors, our children's performance in all cognitive tests was related to maternal age, as has been reported by Illsley (1967), who considered this effect to be a reflection of associated social factors (so that perhaps we should have included this variable in our 'family factors' group). The fact that maternal height shows an important association with both measurements of the child's growth is consistent with Hewitt's (1957-8) report of a highly significant correlation between maternal height and child's height at the age of five years, presumably on a genetic basis. The widespread effects of the child's sex, covering cognitive, growth, behavioural and neurological variables, are consistent with many published observations, including those concerning the parent population (Neligan *et al.* 1974). The widely reported effects of low birthweight formed the starting point of this whole investigation, and the relatively trivial effects of variations in gestational age were reported by Neligan *et al.* (1974), as was the complete absence of significant over-all effect from the three clinical factors listed.

When we come to the 'family factors' we encounter a major surprise: the child's social class of origin, as judged by the father's occupation, is associated with only two very small effects, a finding which is in conflict with all previous evidence on this subject of which we are aware. We suggested earlier that the effects which are usually attributed to the label of social class may in our study have been more correctly (and with greater practical implication) assigned to other more concrete factors which are strongly associated with social class. As in the results of the analysis of covariance, so in our multiple regression analysis; the factor of overriding importance is the mother's care of her child as assessed by the City's health visitors when the child was three years old. Neligan *et al.* (1974) drew attention to the importance of this factor, and to its association with social class, but they did not assess the extent to which it accounted for the effects attributed to social class in the parent population. They did report the effects of ordinal position in the family, confirming the findings of Illsley (1967) and Davie *et al.* (1972). We are not aware of any previous publications concerning the effects of reading to the child regularly, or of the mother's extraversion, her expectations of his abilities at the age of five years, and her neighbourliness, but it seems reasonable to find such associations, particularly with the results of cognitive tests and assessment of behaviour. It would also be no surprise to find relationships between these factors and social class (although we have not attempted to define them in our data).

In our random sample, therefore, the results of the multiple regression analysis make sense in terms of previously published information or of reasonable expectation, or both.

When we look at the comparisons between the random sample and the two extreme abnormal groups (Fig. 14, supplemented by reference to Figs. 12 and 13 when indicated), most of these also make obvious sense. The fact that clinical factors now play a recognisable part is consistent with the higher incidence of breech delivery (in both groups) and antepartum haemorrhage (in the short-gestation group), and the same applies to delay in establishing regular respirations in the very light-for-dates group, where the higher incidence of fetal distress (Table 2.2, p. 12) may reasonably be interpreted as an indication of the same placental inadequacy as may have contributed to the impaired intra-uterine growth.

The differences in the biological factors in the short-gestation group are more difficult to explain. The relatively large effect of mother's age is confined almost entirely to the child's weight increment; and although there is again some association between her height and the child's growth variables (Fig. 12), a major part of the very large effect of her height is due to associations with three cognitive variables. Also, the effect of her smoking during pregnancy, seen only in this particular group, is almost confined to the child's behaviour as reported by the teacher. It is difficult to see any direct causal connection between these particular associated factors and most of the outcome variables mentioned, and the likely explanation would appear to be that they are also associated with other factors in the family environment which we have not measured, and which are the actual cause of the effects which we have recorded.

Turning now to the factors which we have recorded under this heading, any differences in detail in other parts of Figure 14 are overshadowed by the progressive magnitude of the effect of the father's occupational class—from nil in the random sample through noticeable in the short-gestation group to greater even than the mother's care of her child in the very light-for-dates group. We have no satisfactory explanation of these differences: the very light-for-dates group does not differ significantly from the random sample in terms of social-class distribution or mother's care of child, but both these associated factors have a significantly less favourable distribution in the short-gestation group (Tables 2.3, p. 12 and 2.4, p. 13). Figures 12 and 13 show that the individual coefficients associated with social class differ in the two groups of children, performance IQ being the only dependent variable upon which important effects appear in both figures. We do not believe we can resolve our uncertainties concerning the explanation for these differences between the distribution of the effects of the family environment in our three groups of children either by further and more complex analysis of the same data, or by including other data available to us.

However, a number of general conclusions can be drawn from our multiple regression analysis. Among the large range of associated factors which we have studied as independent variables, those we have grouped under the heading of 'family factors' account for much the largest proportion of the effects which we have observed, particularly in the random sample and the very light-for-dates groups. In fact in the latter group more than two-thirds of the recorded effects are accounted for by the 'family factors', half of them under the headings of social class and mother's care of child. Drillien (1970) has suggested that this group of children may be particularly vulnerable to the effects of an adverse environment. Both abnormal groups of children appear to be more at risk of harmful effects from clinical factors (presumably through the mechanism of minimal brain-damage). The apparent important effects of biological factors only in the short-gestation group are less easy to understand but merit further investigation.