CHAPTER 1

The Study Population: Definition and Selection

To learn about any later effects of the variations in the duration and the rate of intra-uterine growth — which were our primary interest — we needed first to select three groups of children, appropriately defined in terms of birthweight and gestational age, and then to examine them at appropriate ages and by appropriate methods to detect any significant differences in their development. We also needed to have information about the other factors, discussed in the Introduction, which might modify the effects of our factors of primary interest. These potential modifiers include the social background, clinical complications of the pregnancy and the neonatal period, and infections or injuries of the central nervous system in later childhood. If the findings of our local investigation were to have more general application, it was important to be able to define the relationship between our study population and a geographically defined population of a kind that would be suitable for comparisons to be made, and relevant conclusions drawn, by workers elsewhere.

The opportunity to fulfil all these requirements (though with varying degrees of completeness and excellence) was provided by the Newcastle Survey of Child Development (NSCD). This survey enrolled all the survivors of the first month of life who were born to Newcastle mothers during the three years 1960 to 1962. Information was obtained through the Newcastle Maternity Survey concerning the mother's previous obstetric and social history, the pregnancy and delivery, and the baby's condition at the time of birth. The NSCD attempted to keep in touch with as many as possible of these children, obtaining information about their health, development and attainments up to the age of 10 years. All the information was simple and, for the children who remained accessible, it was provided by the midwives, health visitors, doctors and teachers who were routinely responsible for the children concerned. The information was standardised and relied on specified definitions whenever possible. Postal enquiries were made for children who had moved away from the area. We were able to categorise 97.8 per cent of our originally enrolled population as being at least presumably normal, definitely abnormal, or doubtful in developmental status (Neligan *et al.* 1974).

Because of practical considerations, we had to limit the number of children from this geographically defined population for inclusion in our study of the later effects of variations in intra-uterine growth. We planned to carry out a very detailed assessment of each selected child over a wide range of aspects of development, in the hope of detecting the most minor or subtle effects; and all observations were to be made by specialists in the appropriate fields, working either full-time or part-time in the study. Our programme for each child selected started with a home visit by the full-time research health visitor, to invite the mother's co-operation and make a detailed social assessment. The child and mother were then brought up to our research unit for the

first battery of tests when the child was five years old, and were brought up again when he was six and again when he was seven years old. The reasons for retesting in this way were to make it possible to apply a much wider range of tests than could be applied at a single session, and to apply tests for particular functions at an optimal age. The details of the tests applied by the full-time psychologist and by the child psychiatrist, the speech therapist and two research child psychiatrists working part-time in the study, are discussed in the appropriate later chapters of this book. We estimated that the maximum number of children who could be tested during the two years for which the full-time research workers were funded would be approximately 400. We decided to exclude all twins and other multiple births because of the distorting effect upon subsequent performance demonstrated by McKeown and Record (1971), and all children whose mothers were unmarried when they were born because of the special difficulty of following them up, and the distorting effects of their varied environmental experience. We also excluded children known to have died before the age of five years.

The Three Groups

Within this practical limit of about 400 children we aimed to select three groups, one of children who were born too soon, one of children who were born too small, and a comparison group.

Comparison Group (Random Controls)

Whatever practical difficulties we might encounter, the principles which should guide the selection of our two abnormal groups were relatively easy to determine once we had accepted the ideas about intra-uterine growth which we have described in the Introduction: we simply needed to define each group from the data on gestational age and birthweight which were available to us through the Maternity Survey. It was more difficult to decide the principles upon which we should base our selection of the comparison group. We could use the data available to us as the basis for some matching procedure or for some form of random sampling. There was a precedent for the former procedure in the study of low-birthweight children in a national sample reported by Douglas (1960). He selected his controls by a process of pair matching (controlling for sex, ordinal position in the family, mother's age, social group and degree of overcrowding in the house). He had to conclude, however, not only that it had been impossible to achieve perfect matching for all the variables in all pairs, but also that "a blind reliance on the method of paired controls would have led to a misleading interpretation of the differences observed". Billewicz (1964), after assessing the relative merits of the matching and random sample methods, had decided that "random sampling is always easier to administer and may also be more efficient from a statistical point of view". We therefore preferred to use a random sampling procedure, which would have the additional advantage of enabling us to obtain normative data concerning the items in our test battery in our geographically defined population, and descriptive data concerning our population in terms of the tests administered.

We thought that the number which would be appropriate for our purposes would

be about 200, leaving a similar number to be distributed between our two abnormal groups. In case of any seasonal variations, either in antecedent factors or in later effects, we decided to select children born during a whole year. Those born between 1st June 1961 and 31st May 1962 were then reaching the age of five years, appropriate for our first round of testing. We therefore selected the children in our comparison group as about a one in seven sample from among the 3329 Survey children born during that year whom we knew to be in Newcastle schools (either private or local authority), but although the selection of the individual children was random, the numbers in individual schools were in the same proportion as those of the Survey children who had been born during the previous year. This procedure enabled the sample to be structured in such a way that its social-class distribution was likely to resemble that of our whole NSCD population of the same age — and this turned out to be the case (Chapter 2). Also, the school experience of our children during the two years over which our testing programme would be spread should resemble that of the rest of the population. Yet the selection of the actual children to be included in the comparison group, from among the Survey children of the appropriate age known to be in each school, was entirely random (and of course by chance included a number of children who were also in one of our two abnormal groups).

The number of children selected by this procedure was 229 (of whom 22 also fell into one or other of the abnormal groups). The mothers of 20 (8.7 per cent) refused to co-operate. Their social-class distribution was of interest : one was in social class II, nine in class III and 10 in class V, which must have produced a slight distortion of the social-class distributions of those who did co-operate (see Chapter 2). Our procedure with regard to the 22 'overlapping cases' throughout the analyses reported in this book has been to exclude them from the 'random control' but to include them in the 'random sample' group. This means that the five-year-old random control group consists of 187 children (a few losses occurred during the later stages of our testing programme, at six and seven years).

Birthweight and Gestational Age

We would need to use both birthweight and gestational age to define our two abnormal groups of children in terms of their intra-uterine growth characteristics. Of course we could only include in these groups babies whose mothers' last menstrual dates were known within an acceptable degree of certainty, and in general we adopted the same criterion as the Newcastle Maternity Survey, namely to within five days (Neligan 1965). However, as we have already mentioned in the Introduction, the raw data of the Newcastle Maternity Survey (Neligan 1965) resembled that of some other large populations (Lubchenco *et al.* 1963, Gruenwald 1966) in that it contained an obvious excess of very heavy babies born at a reported gestational age of less than 37 weeks. This had been shown to produce a bimodal distribution of the birthweights recorded at 36 weeks or less, a finding which was interpreted as identifying a misplaced group of babies with birthweights distributed around the heavier mode, whose true gestational age was one or two months greater than that calculated from the mother's reported dates of her last menstrual period (Neligan 1965, Gruenwald 1966). We were happy to use the procedures suggested by these observations to

'correct' the centile curves relating birthweight to gestational age in the Newcastle raw data for 1960-62 births, with the results illustrated in Figure 1. Only the 10th and 90th centiles are shown, since these were the ones used for selecting the cases in our two abnormal groups.

To check whether the effects of our correction procedure were reasonable, we have compared the position of the significant parts of our curves with those derived from three other large, geographically defined populations which have been published subsequently (Table 1.1). It can be seen that there is remarkably good agreement with the Aberdeen data (Thomson *et al.* 1968), the difference being 90g or less for five of the six pairs of values; and we have given reasons (in the Introduction) for regarding the Aberdeen norms as the best available. Unfortunately they were not available to us at the time in question; and our own data were based upon too few births for sex-specific curves to be derived from them. As a result, and because of the known higher mean birthweight of boys, we knew that girls were likely to be over-represented in our total light-for-dates group defined in terms of our weight-for-gestation curves.

TABLE 1.1

Comparison of certain key percentile values of birthweight (g) of children born in geographically defined European populations

Percentile Values	*Newcastle (1960-62)*	*Aberdeen (1948-64)*[1]	*Britain (1958)*[2]	*Oulu (1966)*[3]
10% at 32 weeks	1380	1310	1100	1150
10% at 36 weeks	2045	2330	1600	2000
10% at 40 weeks	2850	2840	2800	2800
10% at 42 weeks	2910	2920	2850	3025
90% at 32 weeks	2720	2630	2700	2900
90% at 36 weeks	3470	3530	3550	3625

[1]Thomson *et al.* (1968)
[2]Butler and Alberman (1969)
[3]Rantakallio (1969)

Born Too Small (Light for Dates)

We were able to use the corrected curves illustrated in Figure 1 to define our two abnormal groups. We chose the 10th centile as our cut-off point for defining poor intra-uterine growth so as to include a sufficiently large number of children to allow us to subdivide this group for purposes of statistical analysis, by using the 5th centile curve to identify any progressively more severe adverse effect of progressively more severe impairment of intra-uterine growth. As can be seen from the results of our analyses in later chapters, this proved to be a most helpful procedure, so that in nearly every case we have separated this group into two subgroups, one called *rather* light-for-dates (birthweight between 5th and 10th centiles), the other *very* light-for-dates (birthweight on or below 5th centile). There were three children whose birthweights fell below the 10th centile and whose gestational ages were less than 255 days (Fig. 1). These have been included in the short-gestation group for purposes of analysis.

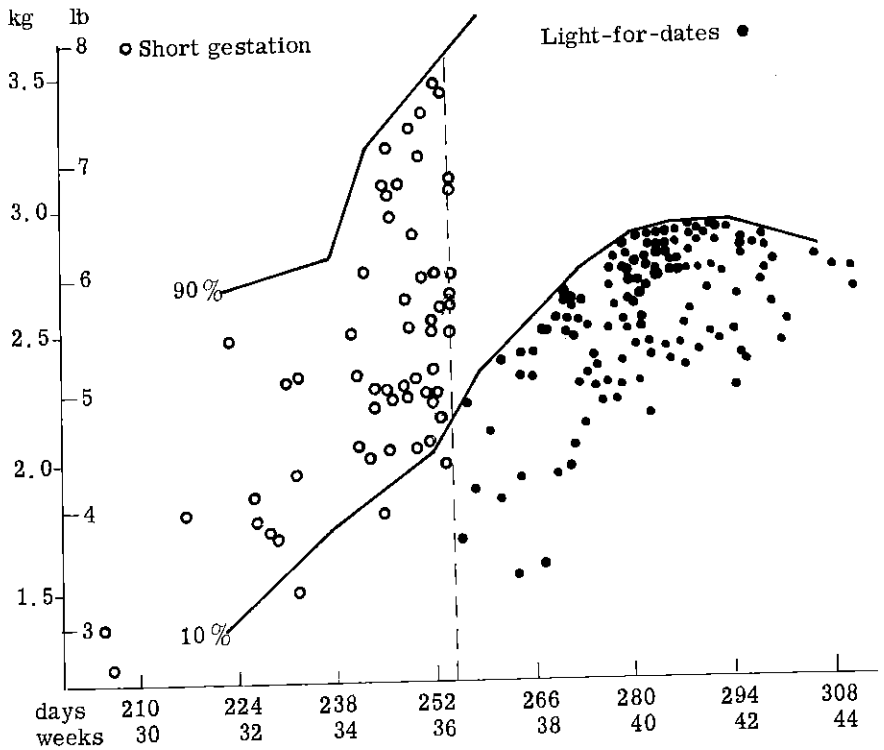

Fig. 1. Birthweight and gestational age of abnormal groups, with relevant portions of 10th and 90th centile curves.

Our final computer programme identified 337 survivors of the first month who had been born during the same year as our random control group and who fell within the light-for-dates group. Unfortunately the programme originally available to us through the Maternity Survey had a limited capacity, and failed to identify all these cases at the time when they were being recruited for our study. However, we were able to satisfy ourselves that this failure was entirely random and these cases, together with those who had moved away from the City by the age of five years, account for the large number who are neither grossly handicapped nor examined by us at the age of five years in Table 1.2 (no mothers in this group refused to co-operate). Those not seen were very similar to those who were examined with respect to the characteristic which placed them in this group, namely their birthweights: those examined had a mean birthweight of 2537g (see Table 2.1, p. 10), and for those not seen by us the corresponding figure was 2559g.

The other, equally important, point about those who were not seen by us and assessed in great detail is that the information available to us through the Newcastle Survey of Child Development made it possible to rule out any gross handicap in the great majority of cases, as shown in Table 1.2. Using the data from either the school entry proforma or the health visitor proforma and postal follow-up, and applying the

criteria described by Neligan *et al.* (1974), we were able to classify these children as either 'normal' or 'probably normal'. This left only 1.2 per cent of the whole group concerning whom we had inadequate information to make this important assessment, and even in these cases the data from the Maternity Survey made it possible to rule out the presence of any gross congenital malformation. The results of our very detailed assessments of the 141 children whom we did examine on one or more occasion can therefore be accepted as truly representative of this group, with a negligible risk that unidentified cases of gross handicap could alter the picture.

Born Too Soon (Short Gestation)

The criteria to be used for identifying the children in this group were more difficult to define. A minor difficulty was the fact that the computer programme originally available to us was too inflexible to allow us to put the cut-off point for gestational age at 259 days (37 weeks), which was becoming accepted as the lower limit of full-term gestation. Faced with the choice between 255 and 262 days, we decided upon the former since it more clearly identified an abnormal group.

A much greater difficulty was the fact (discussed on pages xi and xii) that in our raw Maternity Survey data there was a clear excess of heavy babies in this gestational age-group, so that a correction factor had to be applied in producing the curves shown in Figure 1. Therefore in selecting the children to be examined in our study we needed to find some procedure, as clearly defined and objective as possible, for eliminating these presumably misplaced full-term babies, whose calculated gestational ages were one or more months less than their true ages. Since we would expect the great majority of their birthweights to fall above the 90th centile curve when plotted against their calculated gestational ages, we decided to eliminate from this short-gestation group all children whose birthweights fell above the 90th centile curve. By doing this we must have excluded the heaviest 10 per cent of the children who really belonged there: but this seemed (and still seems) a fair price to pay for a procedure which was objective and presumably excluded the great majority of those who were misplaced.

Again, our initial computer programme failed to identify all the survivors of the first month who should have been included in this group, but again we were able to satisfy ourselves that the failure was random, and the mean gestational age of those who were not examined by us, either because of this failure or because they had moved away from the City, was almost identical with the 244 days of those who were (see Table 2.1, p. 10). However, the usual tendency for greater mobility among members of the upper social classes was exaggerated in this group, presumably by chance, and this selective loss largely accounted for the distorted social-class distribution of those whom we did see (see Table 2.3, p. 12). The effects of this distortion have been allowed for by appropriate methods of statistical analysis in assessing the results reported in later chapters, as have the effects of many other potentially 'confusing' factors.

Of the 151 children correctly identified as members of the short-gestation group, we again had information concerning the great majority of those who were not examined by us from the Newcastle Survey of Child Development which enabled us to

grade them as 'normal' or 'probably normal' (Neligan *et al.* 1974). But there was a larger proportion of this group, 6.0 per cent, concerning which even this basic information was not available (Table 1.3).

Discussion

Our three main groups of children appear to be well suited to our two main purposes of (*a*) comparing the subsequent performance, at school age, of a group who were born too soon with a group who were born too small, and with a representative sample of the population from which they were derived; and (*b*) ensuring that our results are relevant to other similar geographically defined populations. Prospectively and independently recorded information was also available concerning complications of pregnancy and the perinatal period, and possible 'confusing' factors in our children's postnatal environment, in addition to the data collected specifically for the purposes of our study.

Our random sample of 229 children was fully representative of the NSCD children known to be in Newcastle schools at the age of five years, and the results of the large battery of tests applied to them could be regarded as providing normative data concerning our own and other similar populations (even though 20 mothers refused to co-operate). This group also obviously provided the ideal standards of comparison with the short-gestation and light-for-dates groups, whose members were derived by objective criteria from the same geographically and chronologically defined cohort of births.

TABLE 1.2

Status at age five years of all 337 cases in light-for-dates group

	No.	%
Grossly handicapped (1 mentally subnormal, 1 brain damaged)	2	0.6
Considered normal	172	51.0
Considered probably normal	18	5.3
No usable information available	4	1.2
Examined by us at five years	141	41.9
Total	337	

TABLE 1.3

Status at age five years of all 151 cases in short-gestation group

	No.	%
Grossly handicapped (1 deaf and spastic, 1 mentally defective and spastic)	2	1.3
Considered normal	71	47.0
Considered probably normal	10	6.6
No information available	9	6.0
Examined by us at five years	59	39.1
Total	151	

The children examined by us in the two abnormal groups were representative of those who were born too soon and those who were born too small, in the same cohort of births, by our clearly defined criteria. Those whose birthweights fell below the 5th centile were, by definition, representative of the 5 per cent with the slowest nett rate of intra-uterine growth in the cohort, and numbered 74 (see Table 2.1, p. 10). Those whose gestational ages were 255 days or less numbered 59, and were representative of the (approximately) 5 per cent of the cohort who had the shortest duration of intra-uterine growth (compare the total of 151 children in Table 1.3 with the 337 in Table 1.2, who by definition constitute 10 per cent of the cohort). In other words, the two extreme abnormal groups (short gestation and very light-for-dates) represent the same degree of the biological abnormalities in terms of which they were defined, and in comparing the performances of the children within them we are truly comparing the effects of these two abnormalities of intra-uterine growth (provided we can make adequate allowances for the possible effects of 'confusing' associated factors).

The fact that we only examined a minority of the children correctly identified as members of our two main abnormal groups (Tables 1.2 and 1.3) is compensated for by two important pieces of information available to us from other surveys. Firstly, we know that those not seen were almost identical to those whom we examined in terms of their mean gestational ages and birthweights respectively (and the effects of any differences in terms of confusing factors such as social class could be allowed for in our statistical analyses). Secondly, we had information concerning all but 2.7 per cent of the children in the two abnormal groups combined, which could exclude the possibility of any gross handicaps of a degree and frequency sufficient to make nonsense of the implications of our sensitive and sophisticated test programme.