Practical Problems in Conducting Comparative Trials of Feeding Regimens in Very-Low-Birth-Weight Infants

A. F. Williams

The University Department of Paediatrics, John Radcliffe Hospital, Oxford OX3 9DU, England

It has been suggested that the growth of very-low-birth-weight infants fed human milk may be restricted by protein (1) and energy deficiency. If this is the case, it is logical to suppose that enrichment of human milk with human milk proteins and fats will accelerate their growth. It is tempting to suppose that this hypothesis could be proven simply and objectively by a clinical trial. An account of the practical problems encountered in testing this hypothesis in clinical practice is given here. We conclude that there are many obstacles to relating gross nutrient intake and growth differences between very-low-birth-weight infants by clinical trial methods. These factors include difficulty in measuring growth and nutrient intake with accuracy and large variations in morbidity between and within infants.

POOLED HUMAN MILK AND THE VERY-LOW-BIRTH-WEIGHT INFANT—THE PROBLEM

Most of the controversy over the suitability of human milk for feeding the very-low-birth-weight infant relates to pooled human milk collected by human milk banks. Human milk is considered by some to be too low in nutrient concentration (1) to provide sufficient substrate for the very-low-birth-weight infant to achieve fetal growth rates. However, cross-sectional standards of birth weight at various
gestational ages may not represent patterns of longitudinal growth for individual infants born prematurely, and the clinical sequelae of growth rates below those seen in the fetus have not been characterized. There is therefore room for caution in stating that human milk is quantitatively deficient in nutrients, and when nutrient quality is considered, there is strong in vitro evidence to suggest that the constituents of human milk are more appropriate than those of artificial formulas. Many of the proteins can be shown to possess biological functions, potentially conferring specific and nonspecific immunity and digestive functions (e.g., bile-salt-stimulated human milk lipase) on the infant. These systems are elaborate, and it is attractive to propose that the very-low-birth-weight infant fed pooled human milk will be clinically advantaged, but clear direct proof as shown from clinical trials is, in most cases, lacking.

Most clinical studies have compared groups of infants fed human milk with groups fed artificial formulas. These are beset with methodological problems. First, it is difficult to achieve true random allocation of infants—in many studies, the group of infants fed artificial milk has been the group whose mothers did not wish to breast feed. Second, numbers have usually been small. Third, clinical staff know, as do the experimenters, to which group the patient belongs; the effects of staff attitudes on perinatal clinical trials have been remarked on previously (2), and these may be equally important in feeding studies.

Thus, the problem is not simple. Many questions require experimental testing. Clinicians must decide the relative importance of each question according to their practice, as this factor affects the outcome measure that is chosen. One might ask "Does pooled human milk protect the infant from infection?" The outcome measure would be the incidence of proven infection; however, measurement of a significant protective effect would require a large study in the special-care nurseries of industrialized societies, where antibiotics are used in abundance. By contrast, Naryanan (3) has demonstrated such effects in a nursery in India.

In developed countries, the question of most concern is "Is the growth of very-low-birth-weight infants fed pooled human milk
restricted by nutritional deficiency?” This is the question that is discussed here. How easily can it be objectively answered by a clinical trial?

**The Hypothesis**

If the growth of very-low-birth-weight infants is restricted by nutritional deficiencies in pooled human milk, it is logical to suppose that enrichment of pooled human milk will accelerate growth. Consideration of the protein content of human milk and the maturity of biochemical pathways in the very-low-birth-weight infant (4) suggests that supplementary human milk proteins will be well tolerated. The hypothesis to be tested, therefore, is that enrichment of pooled human milk with human milk proteins and human milk fats will accelerate the growth of preterm infants without a metabolic cost.

**The Contribution of “Lactoengineering” to the Study**

The term “lactoengineering” has been employed by Baum generally to describe methods of altering human milk composition. This is methodologically a valuable tool because it allows controlled preparation of human milks of widely differing composition, allowing blind comparisons of enriched human milk and pooled human milk to be made for the first time.

**The Method of “Lactoengineering”**

For the study considered here, the purpose of lactoengineering was simply to provide a human milk substantially enriched in composition in comparison with pooled human milk supplied by the human milk bank. In the case of the John Radcliffe Hospital Milk Bank, “drip” breast milk (DBM) is collected. The composition, collection, and processing of this milk are described elsewhere (5). Composition was altered by first skimming fat from a 4-liter batch of milk (centrifugation at 1,500 g, 4°C, 20 min). Skimmed milk was then concentrated fourfold by ultrafiltration using an Asahi Medical AM10 “hollow-fiber” dialysis cartridge (95% mo-
The skimmed milk concentrate prepared in this way was lyophilized to a powder. Powder and fat from each 4-liter batch was then resuspended in another 4-liter batch of human milk. This preparation was described as a "human milk formula" (HMF). It was pasteurized and microbiologically screened before use.

The protein and energy composition of the experimental "human milk formula" and the control "drip" breast milk are shown in Table 1. Energy content was determined by bomb calorimetry (16) of triplicate lyophilized samples. Protein content was determined by the Biuret technique (7).

The Design of the Controlled Clinical Trial of "Human Milk Formula"

The study was a randomized blind comparison of enriched human milk with unaltered pooled human milk. Random allocation of very-low-birth-weight infants to feeding groups is, in practice, difficult because many mothers elect to begin expressing milk to feed their own infants. Exclusion of these infants would be selective, and in any case, experience shows that these infants, in later weeks, often require an additional source of milk as part or the whole of their dietary requirement.

In this study, infants receiving any enteral feeds at 7 days of age were randomly allocated, using random-number tables, to a control

<table>
<thead>
<tr>
<th>TABLE 1. Composition of the feeds</th>
</tr>
</thead>
<tbody>
<tr>
<td>Drip breast milk</td>
</tr>
<tr>
<td>------------------</td>
</tr>
<tr>
<td>Protein (g/100 ml)</td>
</tr>
<tr>
<td>Gross energy (kcal/100 ml)</td>
</tr>
</tbody>
</table>

1Manufacturer's information. Supplied by Kimal Scientific Products Limited, Uxbridge, U.K.
group or an experimental group without regard to the mother’s own milk supply. The control group received pooled human milk from the milk bank when mother’s own milk was insufficient to meet the fluid demands of the infant or was unavailable; the experimental group received “human milk formula” according to the same clinical indications. Mothers were not told which feed their infants would receive, and consent was obtained on this understanding. It was felt that this might influence their decision to continue breast feeding. The “drip” breast milk and the “human milk formula” were supplied in containers of identical appearance so that the clinical staff, parents, and anthropometrists were unaware which feed a baby was receiving. Twenty consecutive inborn infants with birth weight <1,800 g (excluding small-for-date infants) were identified for inclusion in the study. Two sets of parents refused consent, so that 18 infants were studied. All feeds were administered by orogastric tube, the volume of feed given was measured by syringe and recorded by the nurses. The recommended intake for each baby was 180 ml/kg/day, although in practice this was varied at the discretion of the staff. Babies were withdrawn from the trial at discharge from hospital or if clinical staff felt they were “failing to thrive.”

Weight was recorded daily, or every other day, by nursing staff who used an Avery balance with resolution of ± 10 g. Length was measured weekly using a Holtain neonatal stadiometer (8), but this measurement was sometimes omitted in the sickest babies who failed to tolerate the necessary handling. Head circumference was measured using paper tapes. Measurements of length and head circumference were made in all cases by the same pair of trained observers.

In order to detect dietary protein overload in babies fed the “human milk formula,” venous blood was taken once weekly from all babies, and plasma tyrosine estimated by a fluorometric method (9).

THE OUTCOME OF THE STUDY

Objective comparison of the treatment regimens was frustrated by three major considerations. First, dropout rates were consider-
ably higher than expected; second, accurate measurements of growth over short periods were difficult; and third, a problem was created by variation in the weight trajectories of individual babies. This tended to frustrate between-group comparisons of weight gain at set postnatal ages. These aspects of the study bear separate consideration as problems for the clinical investigator.

**Dropout Rates**

The nursery in which this trial was carried out is a regional neonatal intensive-care referral center. None of the babies studied was outborn, but eight babies out of 18 recruited to the study were transferred to other hospitals in the region after an average stay of only 3.4 weeks (2.3–4.5 weeks). If babies transferred out were a random sample of those included, an objective comparison between the two groups would of course still have been possible. However, babies are transferred out when they are no longer sick, and at this stage, they have usually just begun to gain weight. The babies who remain are, therefore, a selected sick population, and it is likely that nutrition accounts for a much smaller proportion of the variance in growth in this group than in a group of well babies. Simple comparisons between groups receiving different gross nutritional intakes are thus reduced in their discriminatory power.

How might this problem be overcome in a trial design? Multi-center studies might be an answer in the future, provided the number of infants studied was sufficiently large to control for variations in nursery practice by randomly allocating to control and experimental groups within nurseries and not between nurseries.

The group of three babies removed by clinical staff (because of failure to thrive) deserves separate consideration. It included both of the only two babies who were totally dependent on pooled banked human milk (DBM) for their nutrition and one of the three babies totally dependent on enriched human milk (HMF) for their nutrition. Thus, although two out of the three babies totally dependent on "human milk formula" were considered to be growing adequately by clinical staff, none of the babies dependent on unenriched pooled
human milk was. Because this was a blind trial, this observation may be of some interest, but the numbers are of course insufficient for any statistical comparison.

**Short-Term Measurements of Growth**

Weight gain is the most commonly used measure of growth in the nursery. The relative insensitivity of mechanical balances means that the most accurate measurements of weight gain are made by plotting a regression of weight against time, taking the slope of the line as weight gain in grams per day. There are two important qualifications that relate to the accuracy of this technique. First, change in weight with time in infants born preterm (unlike change in length or head circumference with time) appears to obey a power curve equation more closely than a linear regression equation (see Fig. 1). The corollary of this in making comparisons of growth rate between individuals is discussed later, but for the present, let it be assumed that the error involved in calculating weight gain over a short period of enteral feeding (e.g., 1 week) using a linear regression method is very small.

Second, the proportion of the variance in weight that is attributable to time is expressed by the square of the correlation coefficient, $r^2$. In a sick baby, this may be very low (see Fig. 2). In 6 out of 18 babies in the second week of life in this study, $r$ was $<0.7$, but in another 6, it was $>0.95$. In the first case, 50% of the variance in weight can be accounted for by time, and in the second, $>90\%$ is accounted for by time. The reason for this discrepancy cannot be assumed to be attributable to differences in milk intake; it might be equally attributable to greater error in the measurement of weight in the first group. It is difficult to separate these two effects, and, hence, it is difficult to be sure that weight gain is a comparable measure of the effect of milk intake in the two groups.

When the effect of nutrition on growth is investigated in very-low-birth-weight infants, accurate short-term measurements of growth are vital because periods of enteral feeding when the baby is in a
FIG. 1. Growth chart of one of the 26-week-gestation infants studied. Weight does not show a linear relationship with time, showing that weight gain interacts with postnatal age.
FIG. 2. Relationship of weight and time in one of the study infants during the second week of life. Large day-to-day variations in weight in sick infants can make linear regression methods unsuitable for calculation of weekly weight gain.

steady clinical state are few. They are punctuated by periods of intercurrent illness, which may involve curtailment or reduction of nutritional intake. The problem of measuring linear growth over short periods is a pragmatic one. Although measurement of crown-heel length has been shown to be a technique capable of accuracy (8), many sick babies do not tolerate the manipulation necessary to make reproducible readings. Thus, accurate measurements of crown-heel length can be difficult in the early weeks. Measurement of shorter distances (e.g., foot length) has been suggested (6) as an alternative, but there are no longitudinal studies that indicate that growth of the foot parallels growth in crown-heel length in the preterm infant. In addition, the coefficient of variation in the measurement is approximately 1.5% in our experience, and this distance
is close to 33% of the weekly increment expected from the available intrauterine growth standards (10).

Comparing Rates of Weight Gain Between Individuals—The Effect of Weight Trajectories

Although we have referred to intrauterine growth as a measure of optimal postnatal growth (11) for very-low-birth-weight infants, it should be noted that so-called “intrauterine growth standards” are, more strictly, cross-sectional plots of birth weight and gestational age. It cannot be inferred that the growth trajectory of any individual infant should follow such a course. Indeed, as Brandt (12) has shown in longitudinal studies, the weight trajectory of growing preterm infants is quite different from that described by a cross-sectional plot of birth weight and gestational age.

Weight gain in growing preterm infants approximates more closely to a power curve relationship, and this is clearly seen in Fig. 1, showing the growth of an individual infant included in the study. Much variation between infants can be accounted for by the very different patterns of early weight loss over the few weeks after birth. This feature is not purely an effect of variation in nutritional intake; it can be heavily influenced by features such as severity of illness, drug therapy (e.g., diuretics), and water and electrolyte balance.

In the case of the study infants, early weight loss was very variable, even within each of the two feeding groups (Table 2). This meant that the phase of weight gain—starting at the time when infants regain birth weight—begins at very different postnatal ages.

<table>
<thead>
<tr>
<th>TABLE 2. Time taken to regain birth weight</th>
</tr>
</thead>
<tbody>
<tr>
<td>Group</td>
</tr>
<tr>
<td>-------</td>
</tr>
<tr>
<td>Control</td>
</tr>
<tr>
<td>Experimental</td>
</tr>
</tbody>
</table>

*Before regaining birth weight.
Cross-sectional comparisons of weight gain between feeding groups at set postnatal ages cannot, therefore, be made without taking into account the point that each infant has reached on his individual weight trajectory. Infants may start to gain weight at quite different postnatal ages but finish at similar points attained through different routes. Weight velocities at set postnatal ages can, therefore, be spurious. For this reason, weight gain, although a reasonable measure of nutritional state in an individual infant, has to be looked at very critically when one is making comparisons between individuals. For interindividual comparisons, measurements of linear growth velocity would not appear to be subject to these errors, since their relationship with postnatal age approximates very closely a linear regression with a constant slope at all postnatal ages (Fig. 1).

The Relationship Between Weight Gain and Gross Nutrient Intake in Individual Infants

Weight gains (g/kg/day) of the seven babies who were studied from birth to discharge are shown for each individual in Fig. 3. Each data point represents the slope of the regression of weight on

<table>
<thead>
<tr>
<th>Weight</th>
<th>HMF</th>
<th>MBM</th>
<th>Mixed Maternal and Donor Milk</th>
</tr>
</thead>
<tbody>
<tr>
<td>Gain</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Gm/Kg/d</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

FIG. 3. Weight gains of the seven infants completing the study. Each point represents the weight gain shown during a week of pure enteral feeding. Note the large within-individual variations.
time for a 1-week period when enteral feeding was uninterrupted. This figure has been divided by the baby's average weight during the 1-week period to express weight velocity in the units shown. In the case of most babies, few weeks of uninterrupted enteral feeding were identified, explaining the paucity of data points for some individuals. Only in the case of two babies fed entirely on the "human milk formula" (HMF) could gross energy intake be stated accurately, because the energy content of the feeds had been determined by bomb calorimetry and the volume given had been measured by syringe. Although individual babies can show a wide range of weight velocities at different postnatal ages, the correlation of gross energy intake, where this was known, with weight gain was very low \( r = 0.371 \), 8 df).

A study (13) examining the relationship between gross energy intake of babies fed raw expressed maternal breast milk with weight gain made similar conclusions. Several explanations may account for this. First, wide interindividual differences in nutrient absorption exist; in support of this as an explanation is the improved correlation of metabolizable energy intake with weight gain observed by others (14). Equally, energy utilization is likely to vary according to postnatal age and morbidity. Finally, in this study and in the raw expressed breast milk study referred to above (13), the range of gross energy intake provided was similar and quite high (here between 100 and 210 cal/kg/day); it is possible that energy absorption and tissue deposition are maximal at a given energy intake, making the relationship between gross energy intake and weight gain nonlinear.

Biochemistry

Two babies were found to have plasma tyrosine concentrations well outside the normal range (400 and 600 μmol/liter). These two babies belonged to a group of four infants who were dependent on human milk formula for their nutrition (one of these babies was subsequently transferred to another hospital). In one case, however, the baby had received human milk formula for only 24 hr (day 8)
and had therefore not been receiving a high protein level. This
prompted us to examine vitamin C intake. Normal clinical practice
is to provide a multivitamin suspension (0.3 ml) from the seventh
day of life. This contains 25 mg of vitamin C. Although the vitamin
C content of mature expressed breast milk is quoted as 3.8 mg/100
ml (15), the vitamin C content of pooled drip breast milk was found
to be lower than this in all of six pools studied (mean 1.26 mg/100
ml, range 0.8–1.9 mg/100 ml). The concentration of vitamin C in
these samples fell to only 0.48 mg/100 ml (range 0.07–0.8 mg/100
ml) after pasteurization. Babies receiving pasteurized human milk,
therefore, have very low intakes of vitamin C. Following this ob-
servation, subsequent babies received supplementary vitamin C (60
mg daily). No further hypertyrosinemia was seen.

CONCLUSIONS

Care was taken to design a blind randomized control trial of
feeding on enriched human milk, but high dropout rates prevented
statistical comparisons between groups. Other pitfalls in between-
group comparisons were highlighted, which suggests that weight
gain is an outcome measure subject to interaction with postnatal
age and early weight loss and that where gross energy intake could
be stated with confidence, it showed poor correlation with weight
gain.

These findings suggest that clinical trials of feeding regimens for
very-low-birth-weight infants will require large numbers of infants
for study and careful choice of outcome measures. There is much
current pressure to change feeding regimens, often in favor of
artificial formulas, but without objective demonstration of the su-
periority of one regimen over another in clinical trials, these efforts
should clearly be resisted.

ACKNOWLEDGMENT

This research was supported by a grant from the International
Council of Infant Food Industries.
REFERENCES