Impact evaluation of a programme to improve complementary feeding practices: study design and constraints

Yves MARTIN-PREVEL
Serge TRÈCHE
Francis DELPEUCH

1. INTRODUCTION

The evaluation of a nutrition programme is an essential element for the orientation of future activities. Here we deal with the evaluation of the impact of a nutrition programme or, in other words, with an evaluation that seeks to measure the effects of the programme, with respect to its ultimate objective, which should be the improvement of the nutritional status of the target population. This impact evaluation must be distinguished from the process evaluation of a programme, which consists of seeing whether the programme has in fact been implemented, by measuring the outreach of the strategy among the population, identifying and analysing the main sources of resistance, etc. Naturally an impact evaluation does not, and should not, exclude the need for a process evaluation.

In the first part of this paper we will review the methodological basis for assessing impact in the area of nutrition. Interventions related to protein-energy malnutrition will be considered although the methodological principles developed are valid for impact studies of other forms of malnutrition. However, we have deliberately limited our task to cross-sectional evaluation, which is the most appropriate for assessing the impact of public health programmes. Longitudinal methods, although being more precise and more complex to implement, are directed more towards research-type evaluation.

In the second part, as an illustration, we will describe designs used in Congo to assess two strategies for the improvement of complementary feeding practices.

2. METHODOLOGICAL BASIS

2.1 General principles

The evaluation of impact of a nutrition programme cannot be simply restricted to a description of the nutritional status of the population that benefited from the intervention. It is necessary to compare the nutritional status of one or more control group(s) who may belong to another population, or be the same population before action was taken, or both types simultaneously.
This comparison must ultimately provide an answer to the following question: "*What difference, or what part of the difference between the groups, can be attributed to the*

intervention?" Answering this question is not easy. The degree of plausibility, with which a noted difference can be attributed to the programme, is high if potential confounding factors can be eliminated or at least if their effects can be measured and therefore taken into account in the analysis. The first and essential corollary is that the impact evaluation must be planned and designed at the same time as the programme itself.

Prior consideration is thus necessary, not only to define the study design and the variables to be used to control for confounding factors, but also to choose relevant indicators of results, determine the sample size and method of sampling, establish the budget required, etc. We shall review these aspects from a theoretical point of view, but in practice a feasibility study will, in many instances, be needed to define the conditions under which an impact evaluation can be carried out with a reasonable chance of succeeding.

2.2 Confounding factors

A confounding factor is any element that affects the nutritional status of the target population and which, although not part of the programme, is linked to its implementation. This is the major obstacle to be avoided, or at least kept under control, in impact evaluations.

The following is a typical example: if the programme consists of supplementary feeding of children, but participation in the programme also leads to better access to care, as far as the children's growth is concerned it will be impossible to distinguish between the effects of the supplements and of the care. The modification in access to care is thus a confounding factor when assessing the impact of supplementary feeding.

This example is particularly illustrative but, in practice, confounding factors can be much more difficult to identify. It is therefore necessary to systematically collect a certain number of variables that are known to affect nutritional status, and to examine whether their level is identical either with or without the programme. There are four main sources of confounding:

- non-comparability of the groups
- information bias
- time effects
- regression toward the mean.

Non-comparability of groups

For an impact evaluation, the ideal is for the groups compared to be as similar as possible with respect to all the factors that may affect their nutritional status, the only difference being their participation in the programme. Any difference occurring between the groups can then be attributed to the programme. It is thus necessary to carefully study the comparability of the groups on the basis of potential confounding factors.

Three levels can be distinguished:

- community: health facilities, access to drinking water and electricity, and overall environment;
- family: size, income, educational level, feeding practices and all other socioeconomic variables;
- individual: age, gender, vaccination status and other health information.

In practice, the groups are never entirely comparable due to the, all too often, restricted choices of the intervention and control groups and to the ethical and political reasons that motivate this choice. Thus, it is important to know and measure the existing differences between the groups, and follow their development during the course of the programme. It is then possible to take these differences into account when analysing the results by using statistical adjustment techniques. These techniques, however, lead to a decrease in the statistical power of the comparisons, underlining therefore the need to select groups that are as comparable as possible from the outset.

Even more important is to identify and avoid any programme-related systematic cause of differences between the groups, such as for example, any self-selection phenomenon, non-random distribution of the lost to follow-up subjects or an age effect when making before-and-after comparisons. These are real confounding factors which are the source of differences between the groups; thus it will be impossible to take them into account.

Information bias

This is the result of a difference that does not exist in the study groups but in the way information on each group was collected. The simple fact of using a methodology that is not strictly identical for collecting information in the different groups, different teams of researchers, or non-standardized measuring methods, is enough to create an artificial difference between the groups in the results. As a result, there will be some confounding with the effects of the intervention.

Time effects

Like almost every other health phenomenon, the nutritional status of the population is subject to the effects of time (to be distinguished from the effects of age). When measures are taken at different times, it may be difficult to know whether the variations noted are related to the effect of time or to the intervention.

There are mainly two types of time effects:

- secular trend: change in the nutritional status of a population, which is usually moderate unless there are important external events (armed conflicts, for example);
- seasonal trend: cyclical change in the nutritional status, whose effects on the most vulnerable groups can be significant.

Regression toward the mean

According to this phenomenon, when subjects are selected from the ends of a distribution, the value of the variable measured among these subjects tends to move spontaneously towards the central value when a second measurement is taken. The explanation of this phenomenon is that some of the individuals selected have been chosen at random, following fluctuations in the variable or mistakes in measurement. This, however, has little chance of reoccurring for the same individual in a second measurement.

It should be noted that this phenomenon is only a handicap for an impact evaluation if the programme is directed at an individual and not a population level and concerns individuals selected for extreme values of their nutritional indices.

Note

In describing the various confounding factors, we have referred to cases in which these factors lead to incorrectly attributing an observed difference to the intervention. The reverse is equally possible and the confounding factor then makes it impossible to observe a difference that nevertheless exists. The risk of drawing an incorrect conclusion is just as serious in both cases.

2.3 Study design

The best design is a randomized double blind experiment. If well conducted, this yields the most conclusive results regarding the effects of an intervention on the nutritional status of the target population; it is then possible to exert maximum control over confounding factors. But in practice it cannot be carried out at a public health programme level. In addition, a random experiment is practically impossible for many reasons, especially ethical and political reasons, and so it is only possible to use so-called "quasi-experimental" designs.

Here we will discuss the main feasible types of design, and will briefly describe their advantages and limitations. Variations exist or could be invented for each type, but we will refer to the general principle. As we have already suggested, impact evaluation compares an intervention group with a control group. The comparison may be in space ("with vs without programme"), in time ("before-and-after"), or both at once (combined designs, which are the most effective but also the most difficult to carry out) as shown in Figure 1.

"With vs without programme"

Intervention is among the "A" population. After a certain time, depending on the type of programme, the nutritional status of the target group of the "A" population is compared to that of the "B" population, which is not submitted to the intervention.

As described above, it is important that groups A and B are as comparable as possible. However, even if this level of comparability is respected for all the factors affecting nutritional status, a difference noted between the groups cannot with certainty be attributed to the programme because there is no proof that this difference did not exist prior to the study.

Another problem of this type of study design is the choice of the population sector as subject of the intervention. At this level, ethical or political constraints are often an obstacle to valid comparisons. In addition, in order to ensure comparability, it is desirable that population sectors are geographically close. The proximity of the intervention and control groups is, however, the source of the so-called "contamination" effect, i.e. the subjects in the control zone can be contaminated by the programme. This makes it difficult to identify clearly participants and non-participants in a programme. Moreover, this type of study design is the most susceptible to self-selection bias.

"Before-and-after"

The nutritional status of the target group in the A population is first measured at baseline, just before the programme starts, and again after a certain time. This type of design aims at eliminating comparison problems by using the population in the programme as a control group.

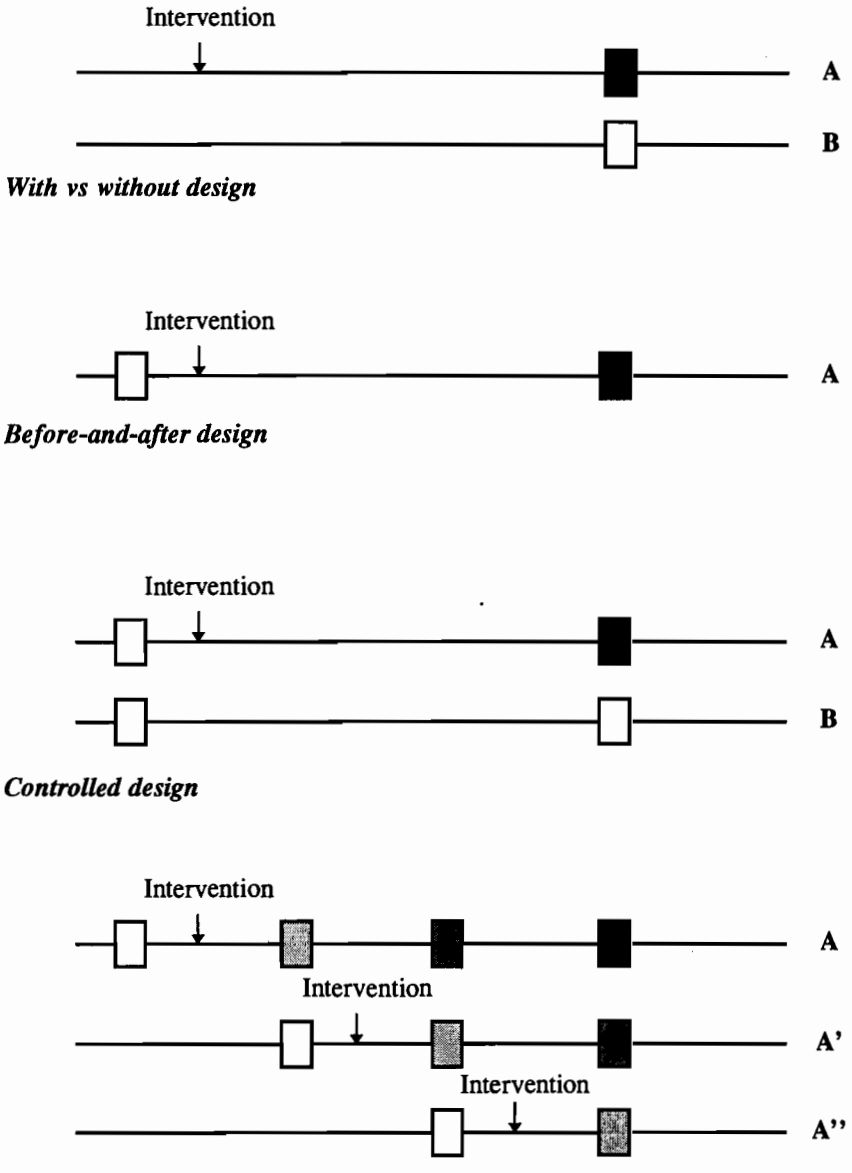

Figure 1
Types of study design for impact evaluation

This design is, however, subject to the effects of time, which can bias the study in a number of ways:

- When the trend in nutritional status is assessed for the same children before and after the programme, the age effect comes into play: for all anthropometric indicators the low values in comparison with the standard reference are more or less frequent depending on age, irrespective of any programme.
- If the programme only concerns a part of the population, generally chosen by extreme anthropometric index values, regression toward the mean occurs.
- Finally, even when evaluation concerns all age groups in the population before and after the programme, there is at least the effect of secular trend. The effect may be mitigated if the duration of the study is short. It can also be estimated indirectly (data on mortality or morbidity, or growth monitoring, vaccination rates, etc.), but such an estimate will always be partial in the absence of a control group.

Combination of designs

Its objective is to make a comparison, both in terms of space and time, to avoid the disadvantages inherent in the two types.

The "controlled design" consists of measuring the nutritional status of the target population before and after the programme, both within the intervention group and the control group. The before-and-after comparison in the latter group permits an estimate of the secular trend and consequently a correction of the before-and-after comparison in the intervention group. In such a design, if the comparability of the two groups is verified before the programme begins and then checked throughout its implementation, it may be assumed that the programme is responsible for the improvement of the nutritional status. In addition, a number of confounding factors can easily be identified and their effects measured. This can sometimes be taken into account in the analysis, but it reduces the number of degrees of freedom, and so decreases the statistical power of the comparisons. This "controlled design" is methodologically the most effective. On the other hand, the choice of the population control group is difficult. In addition to the problems mentioned in connection with the simple "with vs without programme" design concerning geographical selection, it has to be recognized, at the ethical and political level, that a particular sector of the population will remain outside the programme for a certain period of time to allow the changes in the nutritional status of the intervention group to be assessed.

The "staggered implementation" design offers a compromise to the acceptability problems mentioned above. Instead of leaving a population sector without any programme for a long period of time, the programme is implemented gradually, community by community, within a more acceptable delay. This allows on the one hand a "with vs without programme" comparison and an estimate of the secular trend, on the other hand, by successively measuring those entering the programme (who are the control group) and even an estimate of the "dose-response relationship". However, the analysis is more complex and less powerful than the "controlled" design.

In addition, comparability is more difficult to achieve and to control in the successive communities. Finally, and above all, the order in which the communities are included in the

study is often linked to problems of accessibility (geographic, cultural, etc.) which are probably not independent of nutritional risks. This can therefore introduce a confounding factor.

Note

Until now we have considered that study designs are directed at communities. However, they can be designed for individuals, which raises additional obstacles. For example, the self-selection bias becomes much more important, and often cannot even be assessed. The problem of non-random distribution of the lost to follow-up cases also becomes difficult to control. Finally, there is an additional non-negligible source of bias related to missclassification of individuals in the study groups.

2.4 Outcome indicators

In the vast majority of cases, these are anthropometric indices and we will simply explain certain principles for their use in impact evaluations. Among the available indices (weight-for-age, weight-for-height, height-for-age, mid-upper-arm-circumference, etc.), the choice of an outcome indicator will depend on:

- the objectives of the programme and of the evaluation;
- the basic situation (which index is already at a low level);
- the duration of the study (with respect to the sensitivity of the indices);
- the technical and financial resources for the collection of data.

As far as the expression of the indices is concerned, one should bear the following in mind:

- growth velocities are usually more responsive to intervention than gross anthropometric indices;
- expression of indices in centiles or Z-scores is standardized on a reference population but not on age;
- the mean indices are less responsive to errors (in measurement, estimates of age, etc.), and more powerful for comparisons than the percentage of subjects below a threshold value;
- a percentage, on the other hand, is more relevant when making health policy decisions.

2.5 Sampling

The most important aspect is choosing the age groups for the impact evaluation. Age groups with the greatest chance of responding to the intervention should be naturally chosen, and not necessarily the whole group subject to the programme. Extending the study group beyond the sensitive age groups will weaken the evaluation and could attenuate or conceal the real effect; on the other hand, a smaller age group would be less powerful.

The sample size is another important factor. Although there are methods of calculation which can determine the number of subjects required, they will not be covered here. We would simply like to stress that such a calculation implies knowing or estimating a certain number of parameters, and making some hypotheses about the expected results. This is often an additional reason for carrying out a prior feasibility study.

The role of sample selection in the majority of confounding factors mentioned above (comparability of groups, age effects, regression toward the mean, classification errors) has to be taken into account.

2.6 Conclusion: choice, validity, interpretation

Any evaluation of the impact of a nutritional programme therefore requires a well prepared study design. Choices and decisions have to be made and these must be considered carefully according to the objectives of the programme, the resources available, field considerations, basic data, etc. One might decide not to go forward with the impact evaluation when it is impossible to elaborate an evaluation system that has a reasonable chance of success.

In any event, if one decides to go forward, implementation of a design that is effective *a priori* does not guarantee any conclusions that may be drawn. It will always be necessary to verify that, despite all the precautions taken, confounding factors have not slipped into the study. This is essential for interpreting the results and for the internal validity of the evaluation, i.e. for ensuring that what has been measured does actually represent the situation in the groups concerned.

The problem of external validity has to be considered, namely to what extent the conclusions of the impact evaluation may apply to other population groups. This is primarily a problem of representativeness, and it is also necessary to know whether the programme evaluated can be replicated easily for other population groups. In this connection, it should be emphasized that special management of a programme at the level of a pilot zone will limit the external validity of the evaluation.

3. IMPACT EVALUATION OF STRATEGIES TO IMPROVE COMPLEMENTARY FEEDING IN CONGO

In Congo, two strategies to improve complementary feeding of young children were set up in test zones: one concerned urban areas and the other rural areas. The work was initiated by the Laboratoire d'Etudes sur la Nutrition et l'Alimentation[1] of the DGRST-Orstom Centre[2] in Brazzaville, in close cooperation with the family health department, and as part of a project to support nutrition activities, financed by the French cooperation fund and administered by UNICEF.

We will not provide figures for the results of the impact evaluation (in any case it has not yet been concluded), but we will try to illustrate the various methodologic problems described in the first part.

3.1 Brief description of the two strategies

The urban strategy

The urban strategy is directed at families who nearly always give their infants a complementary food consisting of a commercial flour for infants. The pilot programme is based on the promotion and marketing of "Vitafort", a flour for infants and young children

[1] Laboratory of nutrition and food studies.

[2] DGRST: Department for scientific and technological research. Orstom was renamed Institut de Recherche pour le Développement in 1999.

(Tchibindat and Trèche, 1999) in Poto-poto, one of the oldest urban areas of Brazzaville. Its purchase price, for comparable nutritional quality, is three to five times less than that of imported flour for infants. Moreover, the price per unit of energy is the same as that of the fermented maize dough used locally to prepare traditional porridge. The intervention zone was defined on the basis of the limits of the health districts corresponding to two Centres de Santé Intégrés (CSI)[3] in the new Programme National de Développement Sanitaire (PNDS)[4]. The intervention strategy includes action to promote awareness of complementary feeding problems among the staff of health centres, especially those in charge of monitoring growth. This action has been duplicated with the mothers and includes the promotion of breastfeeding, information on recommended age of introduction of complementary food and on Vitafort flour.

The rural strategy

In the rural environment, the intervention zone comprises the Kukuya plateau, 400 km north of Brazzaville. Details of this programme are described elsewhere (Moukolo *et al.*, 1999). It is based on the training of *"nutritional education and food technology extension workers"* recruited locally, whose essential task is to act as a relay between health structures and the community. Their role includes nutrition education, mainly aimed at promoting breastfeeding and correct age at introduction of complementary food, as well as teaching food technology that allows protein-rich porridges of high energy density to be prepared from local foods.

3.2 Evaluation methods: study design

An identical design was chosen to evaluate these two intervention strategies in urban and rural areas and it included the following:

- A "process" evaluation, involving observation and analysis of strategy outreach among the target population. The indicators adopted related to the extent of information on the improved complementary foods proposed, their reputation, and rate of utilisation, changes in complementary feeding of young children, etc. The method used was to carry out successive cross-sectional surveys on representative samples of the population subject to the programme;
- An impact evaluation of the nutritional status of children subject to the programme. This is a "controlled" type of design as described above. For each zone of intervention (urban or rural), a control zone was defined in which characteristics that might affect nutritional status were as comparable as possible. So-called "baseline surveys" were carried out in the intervention zones and in the control zones, to assess the nutritional situation at the beginning and to study the comparability of the groups. Repeating identical surveys using an identical methodology in all the zones two years after the programme is implemented, will make it possible to assess whether the trend in the nutritional situation in the intervention zone differs from that in the control zone. Furthermore, the cross-sectional surveys carried out to evaluate the process will show whether comparability between the groups has been maintained.

[3] Integrated health centres.

[4] National health development programme

3.3 Practical aspects

Once the design of the study has been established, all the problems of choice of indicators and potentially confounding variables, definition of the target population and samples, delimitation of the intervention and control zones still need to be solved.

Before reviewing these aspects, it should be pointed out that many nutritional studies have been carried out in Congo over the past five years, both in Brazzaville and in rural areas. These provided the basic data necessary to set up our programme which allowed us to limit our field surveys when carrying out our feasibility study.

Choice of indicators

The essential aspect of malnutrition in Congo is the fairly high prevalence of growth retardation. The main hypothesis underpinning our study is that this is to a large extent due to inappropriate complementary feeding practices and to the low energy density of porridge. It was thus logical that the main indicator of programme outcome should be the anthropometric height-for-age index. Weight is also measured to estimate the possible effects of the programme on other indicators, for example, wasting. It was the height-for-age index that was chosen, however, when calculating the number of subjects required.

As mentioned above, this index should preferably be expressed in Z-scores. The mean value of the index in the different groups and the percentage of subjects below two standard deviations will be considered simultaneously. The first index will be the most effective for comparison, while the second will make it possible to verify, together with a study of its distribution, whether the programme has benefited the most at risk subjects.

Choice of covariates

The prior studies available enabled us to identify the main variables, essentially of a socioeconomic nature or concerning complementary feeding practices, whose effects on the nutritional status of young children in Congo have been observed. All these variables were collected in baseline surveys in order to verify the comparability of the groups. Other potential confounding factors were also included in the study for safety's sake even though previous studies had not demonstrated their role in Congo. The same questionnaire will be duplicated for the surveys at the end of the programme and the principal variables are also being studied in surveys evaluating the process so as to verify continued comparability of the groups.

Definition of the target population

The target population for the programme ranges from children aged 4-9 months who either receive complementary food or partake of family food depending on their age. In view of the type of programme, however, this group has only been recommended; an analysis of the trend in complementary feeding practices during the intervention will make it possible to define more precisely what age group was actually influenced by the programme.

The target population for the impact evaluation has been chosen from a much broader range, because a growing prevalence of stunting until the age of 2 years has been observed in Congo (even if it is not possible at this stage to know whether the effects of inappropriate

complementary feeding practices extend to this age). Moreover, since the proposed duration of the pilot programme is two years, the age group 4 to 27 months was chosen as the target population for the impact evaluation. The whole generation surveyed at the end of the study will therefore be concerned with the intervention. We are nevertheless aware that this choice would involve a risk of diluting the possible effects of the intervention if it was more limited in duration.

Delimitation of the study zones

— Urban strategy

For the urban strategy, the intervention zone was defined on the basis of the integrated health centres (CSI). In Congo, the national health development plan (PNDS) envisages the conversion of dispensaries into CSI or the creation of CSI in areas where health coverage is deemed to be inadequate. One of the principles of the PNDS is that each CSI should be responsible for a health district that is carefully delimited. The population of the district is recorded and benefits from certain public health activities. It was therefore quite logical to define the study zones on the basis of the districts, and this was done with the exception of a few small clusters which differed greatly from the rest of the zone. This decision was taken because of the very small size of the population in these clusters and to preserve socioeconomic homogeneity.

The choice of CSIs for the study was governed by the following considerations:

- intervention zone sufficiently far from the control zone to limit contamination phenomena;
- districts with comparable socioeconomic levels to obtain the best possible initial comparability;
- same state of progress in implementing the PNDS, and as a corollary support for this implementation should come from the same development agency;
- avoiding interference by other nutrition research programmes also taking place in Brazzaville.

Taking into account all these considerations, the choice was very limited. Finally, two of the oldest urbanized districts of Brazzaville were chosen: Bacongo and Poto-Poto, in which the implementation of the PNDS is supported by UNICEF and where the operation of three and two CSIs respectively was planned in 1993.

— Rural strategy

Here, the problem is very different because the intervention zone was defined very early on. It is the Kukuya plateau, situated approximately 400 km north of Brazzaville, for which basic data on nutritional status and feeding practices were already available. The programme has been specially designed for this zone, particularly as far as food technology is concerned.

The main difficulty was choosing a study design and then a control zone. The first solution envisaged was to divide the plateau into two zones (one intervention zone and one control zone); but this turned out to be impossible for reasons of acceptance by the population and also because of the small number of people available. It was therefore decided to take the

neighbouring plateau of Djambala as a control zone whose population is very similar from the ethnic standpoint. Although the proximity could be a source of contamination, but this should be negligible given the type of programme. The implementation of the PNDS, which covers the whole of Congo and could be a significant confounding factor, had fortunately been programmed for the same period for the region as a whole. However, this choice raised the sampling problems described below.

Sampling

— Sample size

This is the first aspect to be taken into account. As data on growth retardation in the population involved in the study was available, we calculated "in reverse" in order to show, for different sample sizes, the hypotheses from results of the intervention to consider a statistically significant result (Table 1). This gives a rough idea and makes it possible to provide a rapid estimate of the chances of success of the study. Subsequently, a "straight" calculation can be made to define the size of the sample according to the hypotheses of results finally adopted. This was not done in our study, on the one hand because the hypotheses of results were not easy to make and, on the other hand, because we quickly decided to make an exhaustive study among the target population. Table 1 shows that, for example, at the intersection of the 30% column under initial prevalence and the row 500 subjects by group, the value is "-8.0". This means that if the true prevalence of stunting in the population subject to intervention is 30% at baseline and that the prevalence dropped by 8% after the programme, a figure of 500 subjects in each group will be required to show the true difference between the two population sectors compared, at the 5% threshold and with a power of 90%.

It should be understood that this does not mean that a difference of at least 8% must be noted in the samples for the study to be conclusive. It means that if the true difference (which remains unknown) between the population sectors (and not the samples) is 8%, the fluctuations in the sample mean that, with normal distribution of the variable, carrying out 100 surveys of 2 x 500 persons will statistically show a significant difference at the 5% threshold 90 times (i.e. power = 90%).

In our study, we considered that in urban areas, where the initial prevalence of stunting was estimated at 15.0%, 1500 children per group was desirable. This would show, under preset conditions, a true difference of 3.6% in the population. Taking into account that the population's involvement in the programme will certainly not be full, this would already be a satisfactory result. In rural areas, on the other hand, the initial prevalence is estimated at around 30.0% and a much higher level of participation can be expected. Under these conditions, a decrease of 8.0 to 10.0% in prevalence can be envisaged and 500 subjects per group appeared to be sufficient.

— Inclusion/exclusion criteria

The only criterion that appeared to be important concerns the length of residence in the designated zone. It is preferable to include in the evaluation only children who have been subject, or might have been subject, to the intervention. To do so, these children must have lived in the zone since they were born. We have nevertheless allowed for a temporary absence tolerance of 1 month before the age of 9 months, and 3 months between 9 and

27 months. Naturally, the same criteria apply to the control zone in order to maintain comparability. In rural areas, on the other hand, where the population is much less mobile, the application of this type of criterion was not seen as necessary.

— *Composing the samples*

Here, the limit was the size of the population in the study zones. In view of the problems faced in delimiting the study zones, they were necessarily of a modest size.

In urban areas, mainly for reasons of homogeneity, it was only possible to use the health districts corresponding to two CSIs in the Poto-Poto area and three CSIs in Bacongo. The population living in these zones was estimated to be around 30 000 in each area, equal to a potential of around 2000 children aged 4 to 27 months. Taking into account the high rate of exclusion foreseen, an exhaustive survey was the only way of reaching the objective fixed of 1500 children per group. The sample was therefore composed of all the children aged 4 to 27 months who met the criteria of residence in the zones. In order to be as exhaustive as possible, it was necessary to draw up plans of the districts, compound by compound, and to undertake considerable work on a census. Finally, less than 1% of the subjects could not be surveyed and less than 2% refused to participate.

Table 1
Hypotheses of the necessary differences between the two groups, according to the size of the sample, in order to achieve a significant result, with a Type I error of 5% and a power of 90%

		Initial prevalence (%)							
n/group	*mean Z-score*	12.5	15	17.5	20	22.5	25	27.5	30
250	*-0.38*	-7.3	-8.1	-8.8	-9.3	-9.9	-10.4	-10.8	-11.2
500	*-0.27*	-5.5	-6.0	-6.5	-6.9	-7.2	-7.6	-7.9	-8.0
750	*-0.22*	-4.5	-5.0	-5.4	-5.7	-6.0	-6.3	-6.5	-6.7
1000	*-0.19*	-4.0	-4.3	-4.7	-5.0	-5.2	-5.5	-5.7	-5.8
1500	*-0.15*	-3.3	-3.6	-3.9	-4.1	-4.3	-4.5	-4.6	-4.7

In the intervention zone in the rural environment, i.e. the Kukuya plateau, a large-scale nutrition survey had already been carried out in April 1992. The method used had been two-stage cluster survey with a sampling fraction of 0.4. Since the total population was around 16 000, in the end just over 400 children from 4 to 27 months were included in the sample, and 25% of them came from the small town of Lékana. For obvious reasons of acceptance, it was impossible to carry out a further survey just before the programme at the beginning of 1993. We therefore had to accept this sample, slightly less than the desired figure of 500 subjects. In the control zone, also with around 16 000 inhabitants, the problem was that almost 60% of the subjects lived in the town of Djambala. Reproducing the same survey

design as in the intervention zone (two-stage cluster survey) would have resulted in a sample that retained roughly 60% of "urban" subjects, which was not proportionate to the 25% in the survey on the Kukuya plateau. Hence, we had to stratify our sample according to residence in order to respect the 25-75% distribution between the "centre" and the "periphery". The two-stage cluster survey was duplicated for the town of Djambala, but for the periphery the calculation gave an estimate of a total of 350 to 400 children aged 4 to 27 months. Once again, sampling was useless and the study of children in villages outside Djambala was exhaustive.

3.4 Limitations

Evaluating a "package"

Our study aims at evaluating the impact of a programme whose strategies have been defined according to prior analysis of the nutritional situation. The simultaneous implementation of the national health development programme (PNDS), however, obliged us to adapt the intervention planned. The evaluated impact will therefore be that of a "package" programme at the level of complementary feeding practices combined with a much broader health programme, in comparison with the impact of the health programme alone. This raises a number of problems:

- Risk of bias due to a different rate of progress of the PNDS in the zones; this is unfortunately what happened at the rural level, the Lékana plateau having received greater attention in the PNDS than the neighbouring plateau. However, we will have to wait for the final evaluation, before ascertaining whether the level of the covariates adopted was in fact altered to any significant extent in one zone compared to the other.
- Risk of diluting the effect of nutritional intervention: if the impact of the nutritional programme is less than that of the health programme, the difference in trends in the intervention and control zones may not appear with statistical significance.
- Problem of interpretation: we do not have any elements that allow us to assume the independence of the impacts, at the anthropometric level, of each of the programmes (the reverse is more likely). Consequently, any conclusion regarding the effect of the nutritional programme alone is impossible.

Comparability of the groups

The initial comparability of the groups is satisfactory in urban areas even though, as expected, differences already existed in the two districts in Brazzaville. On the other hand, the difference in methodology when composing the samples in rural areas, added to the fact that the two zones were not surveyed the same year, raises a problem. Even if the analysis confirms the initial comparability of the variables which were chosen, there remains a question concerning the secular trend in the year separating the two surveys. If it was insignificant, there is no problem. But if it showed a deterioration in nutritional status, this may obfuscate the real impact of the programme. If it led to an improvement in nutritional status, on the other hand, this could exaggerate the effect of the programme. It is necessary therefore to try to estimate this secular trend.

The continued comparability of the groups during the course of the study has to be verified. At the rural level, as pointed out in the preceding paragraph, it is feared that there may be a

confounding factor caused by the earlier involvement of development agencies implementing the PNDS on the Lékana plateau in comparison with the Djambala plateau. At the urban level, the situation is much worse because of the important socio-political events that have shaken the Congolese capital since the study began. These greatly hindered the progress of the programme because the production and supply of Vitafort flour were disrupted for several months. Above all, however, these events caused important population movements in both the intervention and control zones, including transfers from one district to another. The impact evaluation in urban areas therefore had to be abandoned because it would obviously not provide results that could be interpreted.

External validity

At the rural level, the involvement of the various actors in the practical implementation of the nutritional programme whose impact we are trying to evaluate has been significant. If the evaluation proves conclusive, despite the possible problems of interpretation described above, it will be necessary to consider whether the programme can be replicated on a larger scale. Although the strategy assessed was designed to be integrated in the national nutrition policy, the human and financial resources utilized on a pilot scale, as well as the motivation of the participants, are often important factors in the success of a programme.

4. CONCLUSION

Evaluation of the impact of a nutrition programme is relatively complex from the methodological point of view and, as we have shown, constraints in the field make it even more difficult to carry out. Finally, the implementation of such a study is basically a matter of compromise between what is methodologically acceptable and what is possible in practice.

Because of the time needed to carry out the study, an impact evaluation is rarely of use to the programme itself, except when replicating the pilot intervention on a larger scale. Therefore, it is important to judge whether the type of programme evaluated improves the nutritional status of the target population. In addition, the impact evaluation is accompanied by a "process" evaluation that will provide valuable information to allow the programme to be adapted to another situation and integrated into a national policy.

Providing the resources to evaluate the impact of a programme logically appears to be essential for the further development of nutritional intervention policies. At the conclusion of a programme in the area of nutrition, however, it is only rarely that one can see to what extent the objectives fixed have been reached. This is basically due to the methodological problems of impact evaluation, its cost, and the duration of the study, which makes it vulnerable to large-scale uncontrolled modifications of the overall context.

In conclusion, two requirements are essential:

- Plan impact evaluation of a programme at the same time as the programme itself so as to define a study design that is best adapted to the special characteristics of the programme and to budgetary constraints.
- Explore and develop all the so-called "qualitative" evaluation methods which, although they cannot replace the quantitative results of an impact study, are less expensive and easier to implement.

REFERENCES

Bouyer J, Hémon D, Cordier S, Derrienic F, Stücker I, Stengel B, Clavel J (1993) *Epidémiologie: Principes et méthodes quantitatives* [*Epidemiology: principles and quantitative methods*]. Paris, Editions Inserm.

Haaga J, Clark L, Edmonston B, Habicht JP, Kenrick C, Kurz K, Mason J, Test K (1985) *Evaluating Effects of Child Feeding Programs (Report of Workshop)*. Ithaca, New York, Cornell University.

Jenicek J, Cléroux R (1983) *Epidémiologie [Epidemiology]*. Edisem, Québec.

Klein RE, Read MS, Riecken HW, Brown JA, Pradilla A, Daza CH (1979) *Evaluating the Impact of Nutrition and Health Programs*. New York, Plenum Press.

Mason J, Habicht JP, Tabatabai H, Valverde V (1984) *Nutritional Surveillance*. Geneva, World Health Organization.

Moukolo A, Tchibindat F, Trèche S, Martin-Prével Y, Pezennec S, Gami N, Louyat de Dibantsa Y (1999) Nutrition education and technology transfer: a pilot intervention on the Kukuya plateau. In: Dop MC, Benbouzid D, Trèche S, de Benoist B, Verster A, Delpeuch F, eds: *Complementary feeding of young children in Africa and the Middle East*. Geneva, World Health Organization, pp. 165–173.

Rumeau-Rouquette C, Bréart G, Padieu R (1985) *Méthodes en Epidémiologie (3è éd.) [Methods in epidemiology]*. Paris, Flammarion Médecine-Sciences, Paris.

Tchibindat F, Trèche S (1999) Congo: Vitafort: flour for the preparation of energy dense gruel. In: Dop MC, Benbouzid D, Trèche S, de Benoist B, Verster A, Delpeuch F, eds: *Complementary feeding of young children in Africa and the Middle East*. Geneva, World Health Organization, pp. 383–392.

WHO (1983) *Measuring change in nutritional status*. Geneva, World Health Organization.

WHO/NHD/99.3
WHO/AFRO/NUT/99.4
Distr.: General
Original: English

Complementary feeding

of young children in Africa and the Middle East

IRD
Institut de recherche pour le développement

World Health Organization
Geneva

COMPLEMENTARY FEEDING OF YOUNG CHILDREN IN AFRICA AND THE MIDDLE EAST

MC Dop[1], D Benbouzid[2], S Trèche[1], B de Benoist[2], A Verster[3] and F Delpeuch[1]
Editors

[1] Laboratoire de Nutrition Tropicale, IRD
Institut de Recherche pour le Développement (formerly ORSTOM)
Montpellier, France
WHO Collaborating Centre for Nutrition

[2] Department of Nutrition for Health and Development
World Health Organization, Geneva, Switzerland

[3] WHO Regional Office for the Eastern Mediterranean
Alexandria, Egypt

World Health Organization
Geneva, 1999

Part of this document was published in French as "L'alimentation de complément du jeune enfant" by Orstom Editions, 1995 (ISBN 2-7099-1289-9). The English translation is published with the permission of Orstom.