National Library of Canada
Acquisitions and Bibliographic Services Branch
395 Wellington Street
Ottawa, Ontario
K1A 0N4

Bibliothèque nationale du Canada
Direction des acquisitions et des services bibliographiques
395, rue Wellington
Ottawa (Ontario)
K1A 0N4

Your file Votre référence

Our file Notre référence

NOTICE

The quality of this microform is heavily dependent upon the quality of the original thesis submitted for microfilming. Every effort has been made to ensure the highest quality of reproduction possible.

If pages are missing, contact the university which granted the degree.

Some pages may have indistinct print especially if the original pages were typed with a poor typewriter ribbon or if the university sent us an inferior photocopy.

AVIS

La qualité de cette microforme dépend grandement de la qualité de la thèse soumise au microfilmage. Nous avons tout fait pour assurer une qualité supérieure de reproduction.

S'il manque des pages, veuillez communiquer avec l'université qui a conféré le grade.

La qualité d'impression de certaines pages peut laisser à désirer, surtout si les pages originales ont été dactylographiées à l'aide d'un ruban usé ou si l'université nous a fait parvenir une photocopie de qualité inférieure.

Canada

DETERMINANTS OF GROWTH FAILURE IN A GROUP OF ZAMBIAN CHILDREN

by

DENIS GERARD ALLARD

Thesis submitted to
the School of Graduate Studies and Research
in partial fulfilment of the requirements for the
M.Sc. degree in Epidemiology

University of Ottawa

ISBN 0-315-80039-9

UNIVERSITÉ D'OTTAWA
UNIVERSITY OF OTTAWA

ACKNOWLEDGEMENTS

I would like to thank Dr. Felicity Savage King and Dr. Ian McDowell, the owners of the raw data-set used in this study, who have graciously made it available to me for my project. My thesis supervisor, Dr. McDowell, has always been encouraging and helpful during the course of this research. I wish also to thank Dr. S. Raman for his advice on some aspects of statistical analysis methodology.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

Chapter 1

1.0 INTRODUCTION

1.1 Growth as a Measure of Health:

For many decades now, growth monitoring has been used in clinical paediatrics as an instrument to assess the development and wellbeing of the individual child. Anthropometric standards permit comparison of the child with the 'norm' for his age and sex. Anthropometric monitoring has also been used extensively in epidemiologic studies because growth parameters are thought to be good proxy measures of health in children. Indeed, the growth rate reflects the effects of successive, transient episodes of nutritional deprivations, infections, and metabolic disturbances to which children may have been exposed (42).

Although growth failure in the developing world derives from poor nutrition and illness, these are themselves the product of a complex interaction of variables from social, economic, demographic, cultural and physical aspects of the human environment (11,22,25,36). The intricate interaction of all these factors has not been thoroughly studied yet and research in the field has not provided an accurate perception of the relative importance of each determinant. Many of the factors operative in some cultures (such as unhygienic practices, delay in proper treatment, special

consideration given to a child because of its gender) have not proven to be as important in other cultural contexts (6,21,29). Most of the past studies (8,15,24) have been limited in their scope, including few variables in their analyses, and differing in the choice of variables used as proxies for socio-economic status and for dietary practices. The majority of researchers have not studied, in a comprehensive fashion, the progression from demographic and socio-economic factors, through intermediaries such as feeding patterns and morbidity, to outcomes such as growth failure and mortality, but rather have limited their analyses to only a part of the equation (2,6,9,33). Furthermore, most of the research has looked at cross-sectional data from which, although potential associations could be sought, no causal inference could be drawn (16,31). Only in the last decade have some longitudinal studies started to unravel causes and effects (2,3,9,20,32). However, the great majority of these studies looked specifically at malnourished children and not at a general population.

1.2 Background:

Let us first review here, in chronological order, some of the important research done in recent years on the determinants of growth failure.

In 1956, Gomez et al. (16) studied 733 children hospitalized in Mexico for malnutrition defined by low weight-for-age. The authors were among the first to define malnutrition in relation to the median of the U.S. standard percentile of weight-for-age. The majority of malnourished children thus categorized were found to suffer from dehydration. Bronchopneumonia (Odds Ratio (O.R.) = 9.32), diarrhea (O.R. = 2.27), and a lower weight-for-age (O.R. = 1.73) appeared as notable predictors of death in their children.

Kanawati et al. (23), in 1973, reported on a case-control study of 57 failing-to-thrive children and 48 controls. The index of thriving was made up of the sum of scores given to four indicators of growth: weight, length, head and mid-arm circumferences. The cases differed significantly from controls (in descending order of statistical significance) by the following factors: the late introduction of supplementary solid food, a lower income, a lower education of both parents, a lower rate of vaccination, higher rate of prior hospitalization, early weaning, higher rate of chronic gastro-enteritis and pertussis, a lower rate of water supply within the house, a larger family size, an earlier introduction of bottle feeding, and a lower protein intake.

In 1975, Martorell et al. (30) studied longitudinally 716 Guatemalan children aged 2 weeks to 7 years. They looked at

the influence of morbidity (determined by diarrhea, fever, respiratory illnesses) on weight and height velocities. The authors found diarrhea to be an important determinant of slower weight ($p = 0.025$) and height ($p = <.005$) increments. Fever and respiratory illness, for their part, showed no effect on the growth velocity.

Johnston et al. (21) in a 1980 survey of Mexican children found family size, number of preschool siblings, and occurrence of death among siblings to be inversely associated with the growth of the children.

Dixon et al. (12) in a 1982 study of twenty malnourished children in Kenya identified bonding failure and illness as significant detrimental factors in malnutrition.

In 1984, Victora et al. (45) trying to answer whether prolonged breast feeding was detrimental to growth determined that the prevalence of malnutrition was lowest in those children breast-fed for 3 to 6 months, but that nutritional status appeared to get worse in those breast-fed longer.

In the same year, Oyedeji (39), in a cross-sectional survey of 88 malnourished Nigerian children, found that important associates of malnutrition were: history of bottle feeding

(71% of cases), evidence of poverty (54%), single mother (50%), recent diarrhea (33%), and recent measles (30%).

Lovel (29), again in 1984, in a survey of 151 mothers about causes of death in young Pakistani children, determined that 60% of the deaths were in females, mainly less than two years of age. Toddler girls were markedly malnourished when compared with boys. If a girl was born, the parents were quicker to seek another pregnancy which may bring about an earlier weaning. The parents were also less likely to seek medical care for a girl. These behaviours are however heavily determined by cultural factors and may not be applicable to an African context. To stress the point, Linhares et al. (1986) in a growth survey of Brazilian children found that, contrary to the previous Pakistani survey, underprivileged boys were always more severely affected than girls were, especially on height, when compared to more well-to-do counterparts.

In 1985, Afifi (2) with a study on 170 Egyptian children, showed that an early introduction of solid foods had a favourable effect on their length ($p = .045$). The author quotes other researchers who found a similar positive association between this factor and a reduction in mortality rate among infants after the first 6 months of life. A history of death among siblings had a statistically

significant negative correlation with growth velocity ($p = .006$).

That same year, Crean et al. (8) found an increased risk of death in the children when there were a lower education of the mother, poor living conditions, a low birth weight, a decreased access to medical care, the late introduction of solid food, a lack of pure water, and a lower immunization rate. Unfortunately, the authors do not report on the strength of the associations.

Excler et al. (14) in 1985 looked at 59 Togolese children hospitalized with malnutrition. Five factors were studied: diet, infection, familial demography, economic status, and socio-cultural environment. The authors concluded that poverty, ignorance and food taboos were related to a lower variety of weaning foods and that conjugal disorders produced child-mother break-up and fathers' bonding failure. A decline in breast feeding led to earlier malnutrition and a shorter breast feeding period increased the risk of dying. Morbidity appeared to play an important role in malnutrition (especially from diarrhea and from measles), and more so after two years of age. Economic status was an influential variable, especially in an urban setting.

Pickering (40) in a study of 493 Gambian children six to thirty-six months of age done in 1985, discovered that seasonality (rainy season) had an influence on the rate of diarrhea. No influence of the mother's educational level was apparent on diarrhea or on nutritional status. The quality of housing, possession of a refrigerator, better sanitation, proximity of piped water, presence of the father in the household, and a low maternal parity all showed a positive association with height-for-age. The author does not provide quantitative information on the strength of these relationships.

Martorell et al. (32), in a publication in that same year, reports, from a study of 758 children from Honduras, that breast-feeding solely was associated with better growth than mixture-fed children who themselves had better growth than exclusively bottle-fed children. Breast feeding was also associated with a decreased incidence of diarrhea but not of respiratory illness. Mothers who breastfed longer were more likely to be poorer and less educated. This economic association may explain the apparent negative relationship of prolonged breast-feeding found previously by Victora et al (45).

In that same period, Bindon (6) in a study of more than 6,000 Samoan children found, contrary to other studies, that

breast-fed babies were consistently smaller on all measurements, after correction for age and sex. The group's means were, however, still higher than the 50th U.S. percentile. Earlier introduction of solid foods had a positive influence on height but not on weight. Growth standards in this Samoan population are obviously quite different from children of other developing countries, and so these findings may not be comparable.

The next year, Victora et al. (44) looking at the role of social and environmental variables capable of altering Brazilian children's growth, found a positive, significant relationship of height-for-age and weight-for-age with better education of the parents, more so with the father's. Family income, number of siblings, and single parent family were the other socio-economic factors identified as worthy of note. Environmental factors such as piped water, sewage disposal, type of housing, and degree of crowding were found to be heavily confounded by economic status.

Also in 1989, Victora et al. (46) in a case-control study of 170 South-American infant deaths and 340 neighbourhood controls showed that infants who received milk supplement in addition to breast-feeding were 4.2 times as likely to die from diarrhea as children who received breast milk only. For children never breast-fed, the risk was 14.2 times

greater. The authors controlled for changes in feeding habit brought on by the diarrheal episode. The study has shown a very strong protective effect of breast feeding against diarrheal deaths in infancy. The authors also found that infants from homes without piped water were at a four times greater risk of diarrheal mortality than those with piped water inside the home. However, this effect may be indirectly dependent on economic status, as already commented above.

The whole area of cultural beliefs and attitudes regarding nutrition of infants and toddlers has received some attention from anthropological researchers in Third World countries and deserves a quick review here as some of these factors may be of considerable importance in influencing the type and the timing of food administration. In many cultures, food taboos may play a role to varying extent. Ojofeitimi (1985) (37) classified the reasons for food aversions of low income, illiterate pregnant Nigerian women into four groups: health, tradition, economy, and religion. He found that more than two-thirds of mothers strongly avoided milk, cowpea seeds and 'bournvita' for fear of having big babies which may lead to a difficult labour and a cesarian section. In Egypt, Sukary-Stolba (1987) (41) discovered that 'light' foods are given to rural toddlers and 'heavy' foods are withheld by mothers for fear of their

negative impact on the digestive system. The author argues that understanding the culturally perceived values of food may provide an insight into the overall context of the problem of toddlers' malnutrition. As Olango & Aboud (1990) (38) found in their study of attitudes and practice in the treatment of diarrhea in mothers of Ethiopian children under 5, the prognosis of a disease can often be influenced by cultural beliefs that dictate the type of treatment provided. In their survey, the authors found that over 50% of mothers restricted the child's fluid intake and 70% stopped or decreased food intake at the first sign of diarrhea in the child. This detrimental behaviour was based on the rationale that the less went in the mouth, the less would be lost at the other end.

The interaction of "modern world" nutrition practices with the more traditional beliefs and behaviours of a specific community may lead to effects that are unpredictable, quite heterogenous, and evolving rapidly during the transition period from "old" ways to "new" ones. This phenomenon is often seen during rapid urbanization of a society. In order to understand malnutrition and infant health in a particular community, Dettwyler (1987) (10) thought that knowledge of both the beliefs and the practices associated with infant feeding in that community is essential. In a study to elucidate such facts in a Malian community, he found that

weaning took place at an average 20.8 months. He established that breast-feeding and weaning practices affected the growth and development of infants during the first two years of life. However, in contrast to many other populations, a number of infants in this Malian community showed improved growth after weaning. The author comments that some traditional beliefs about infant feeding are changing under the pressure of urban norms, while others remain resistant to change, with varying effects on infant health. A study done in Nigeria by Odebiyi (1985) (36) found that educated mothers in good jobs spent less time at home and admitted to using more bottle feeding whereas illiterate mothers practised forced hand feeding. While the children from educated mothers suffered neglect and were deprived of advantages of breast feeding, the children of the illiterate endured undue exposure to unhygienic conditions of farms and market places and bore the consequences of forced hand-feeding. Igun (1982) (18) found that in Nigeria a group of 250 non-literate, low-income mothers were most influenced in their infant feeding practice by media advertisement and by the detrimental example of elite mothers favouring bottle-feeding.

Cederblad & Rahim (1986) (7) demonstrated that socio-economic changes over a 15 year period in urban Sudan, as witnessed by improved housing, food and sanitation,

increased accessibility of schools and health facilities, and a new acceptance of female education, was accompanied by an improved somatic health and nutritional state of the older children. More children also survived their preschool years. However, socio-economic status, as an indicator of the family's financial ability to provide food and medical care, could not account for the variation in nutritional status, in an investigation done by Dettwyler (1986) (9) in West Africa. The investigator documented several fundamental beliefs regarding infant feeding: 1) a child does not need to eat solid food before 8 months of age; 2) if a child is hungry, he will eat, if he does not want to eat, he should not be forced to eat; 3) only the child himself knows when he is hungry and when he is full. In his research, ranking of "maternal attitude" (towards the said beliefs) and growth performance correlated significantly. The author concluded that the cultural belief system regarding infant feeding and the variations in implementation of this system reflected in maternal attitudes, play an important role in determining the nutritional status of the child and also secondarily through their effect on medical care. He also pointed out that these variations occurred within the same general cultural framework. Similarly, Kusin (1985) (25) concludes that the mother's perceptions about what infants should eat and her ability to make suitable weaning foods, should be considered

when attempting to understand the background of growth faltering, such as by a too rapid introduction of weaning foods.

Another facet that is often disregarded is the possible influence of the psychological make-up of the child. This may sometime play a role in altering feeding practices and should probably be taken into account more often by researchers. Devries did so (1984) (11) and noted that in East Africa, in 1974, mortality was greater for infants with easy temperaments.

In summary, previous research has identified many factors as potential causes of malnutrition in children. The late introduction of solid foods, early weaning from the breast, early introduction of bottle feeding, and low protein intake were often identified as detrimental to growth and seem to exert their effect on it mainly through an increase in morbidity. Conjugal disharmony, low income, and low education of the parents are also significant factors. A lower rate of immunization, high prior morbidity (especially diarrhea, measles and pneumonia), and death among siblings are seen as forerunners of a poor growth rate. Large family size, many young siblings, and a high parity of the mother are detrimental as well. A low economic status is linked to a low quality of housing, the absence of a refrigerator,

poor sanitary facilities, and poor access to treated water; this state of affairs brings about an increase in infectious disease and subsequent malnutrition. General attitudes towards gender of the child appears to influence the quality of feeding practice and the seeking of medical care for treatment of illness. This is more prevalent among the lower socio-economic groups. However, the direction of the prejudice varies with different cultures.

Other cultural factors that can strongly influence either negatively or positively the growth patterns of young children are food taboos, traditional attitudes towards infant feeding, and influences on ancestral practices of behaviours brought about by a rapid socio-economic development.

Predictors of death in the children are similar to the ones for growth but in the reverse direction, with a stronger influence of prior morbidity, low birth weight, and poor feeding practices (early weaning, bottle feeding, late solid food supplementation). Innate personality factors in the child, such as an easy temperament, as well as low education of the mother, and low economic status with its influence on sanitary and nutritional conditions were also shown to be associated with a higher rate of mortality.

From this overview of the research done on predictors of malnutrition and of death in young children, it can be seen that the strength of the association between the predictors and the outcomes measured was not always mentioned or even calculated. This was obviously not possible with some studies which only looked at a very limited number of malnourished subjects, without a control group. Furthermore the underlying pathways of action of some factors were rarely specified. There is a need to consider the effect that consolidated measures of all these potential intervening variables may have on the outcomes of poor growth and early demise in order to derive effective remedial solutions. The present study will attempt to do just this, using a set of data on African children.

This research project is a longitudinal study of factors affecting the growth of a cohort of Zambian preschoolers. During the period of data collection (1971-72), infant mortality in Zambia was approximately 120/1,000, and life expectancy at birth about 45 years. These estimates are interpolated from figures given by the World Health Organization for 1960 and 1980 (49). These numbers are distressing by industrialized world standards but are quite within range of the average developing country at the time of data collection.

Chapter 2

2.0 AIM AND OBJECTIVES OF THE STUDY

2.1 Aim:

This study will attempt to determine, in the context of a developing country (Zambia), the impact that many demographic, socio-economic, and environmental factors may have on young African children's diet, on their morbidity, and, in turn, on their growth pattern and their mortality.

2.2 Rationale for the Design:

The analysis will use data collected in 1971-72 in a longitudinal study of 1342 children in Lusaka. The children, between one month and five years of age, were followed for two years. During that period, extensive information was collected on socio-economic and morbidity background at the baseline survey, while growth, some illness data, and feeding patterns were recorded at six months intervals. The analysis is based on models of the ways various demographic, socio-economic, environmental, and other factors may interact to produce their effect on a child's growth and of how they might precipitate an early death (figures 1 & 2).

Bairagi (3) has suggested that some factors, such as economic status, appear to "affect nutritional status mainly

through two intermediate variables, food intake and disease". The author states, however, that this appears to vary between different populations; for example, income was a major determinant of child nutrition in Bangladesh but not so in Kenya. These two intermediate variables were included in my initial growth model (figure 1) and were derived from original variables measured longitudinally during the course of the study, but prior to the growth outcomes being considered. As indicated by the heavier effect arrows in figure 1, the bulk of the influence of the other factors on growth was assumed to pass through these two intermediate variables. Prevention practices (comprising vaccinations and 'well baby visits' to the village dispensary) were also included in the initial models as it was suspected that they may have a favourable influence by improving feeding practices and by reducing morbidity. As suggested by Gerein (15), visits to the under-5 baby clinics with its monitoring of the child's height and weight might be expected to provide prompt feedback to the mother about how well she was doing in feeding her child and caring for it and might thereby have enhanced her feeding practices. Immunizations would be expected to decrease morbidity from tuberculosis, whooping cough, diphtheria, tetanus and poliomyelitis. Some data on the kind of food present in the household and the frequency with which it was consumed was available to me for

half of the sample and will be used also for some of the analyses.

It makes sense that the factors increasing infant mortality would be similar to those impairing growth. Unfortunately, because the components of the intermediate variables (Feeding Practices and Current Morbidity) were recorded over the two-year follow-up period, whereas the death of the child may have occurred at any time after the start of the study, it was not possible to include these intermediate variables in the analysis of mortality and thus they do not appear in the model (figure 2), although ideally they should have. Male gender was included in the model as a protective factor to test my impression that a boy is generally assigned a greater worth than a girl in African society so that a greater effort will be exerted to ensure his survival.

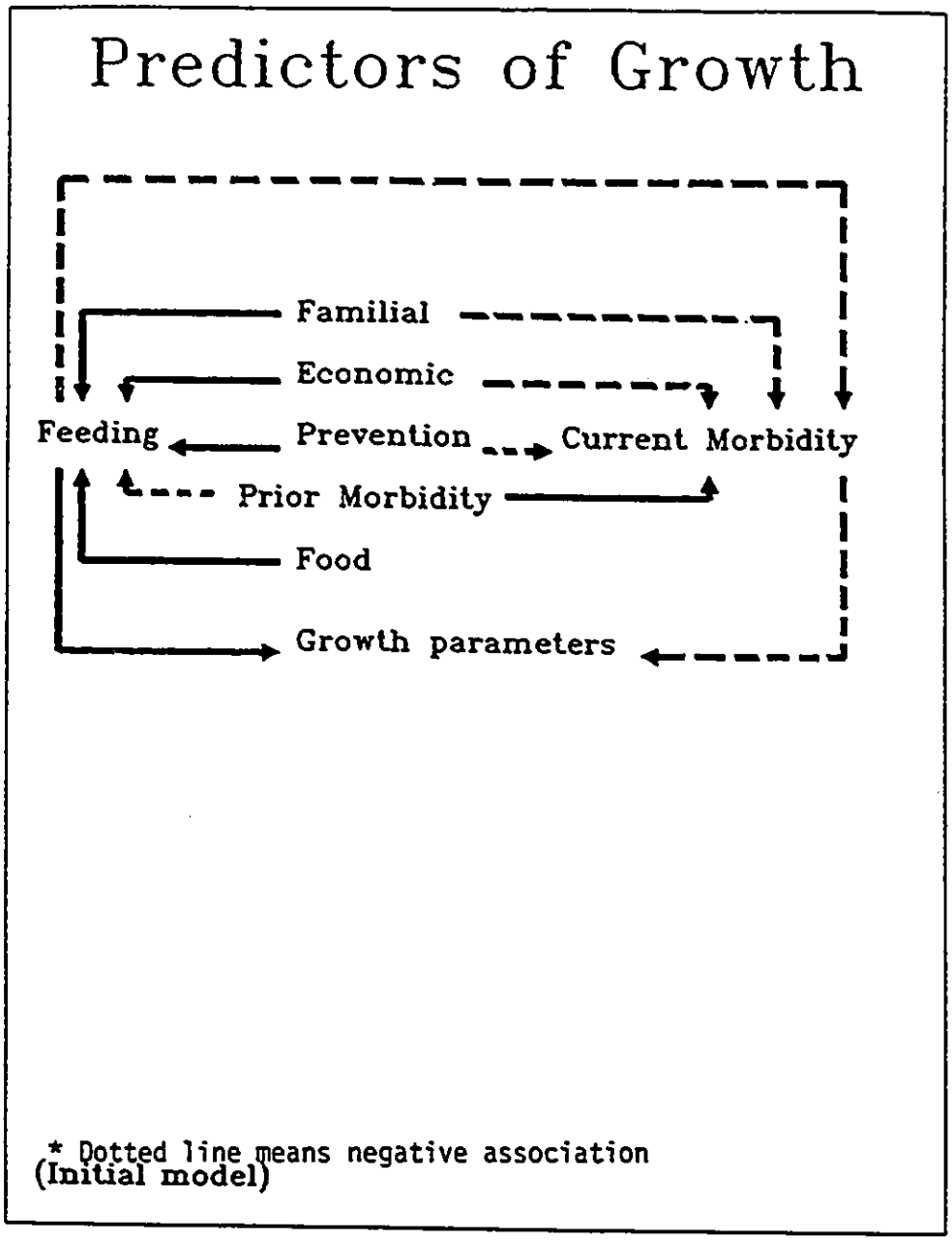


Figure 1 - Hypothesized Model of Predictors of Growth

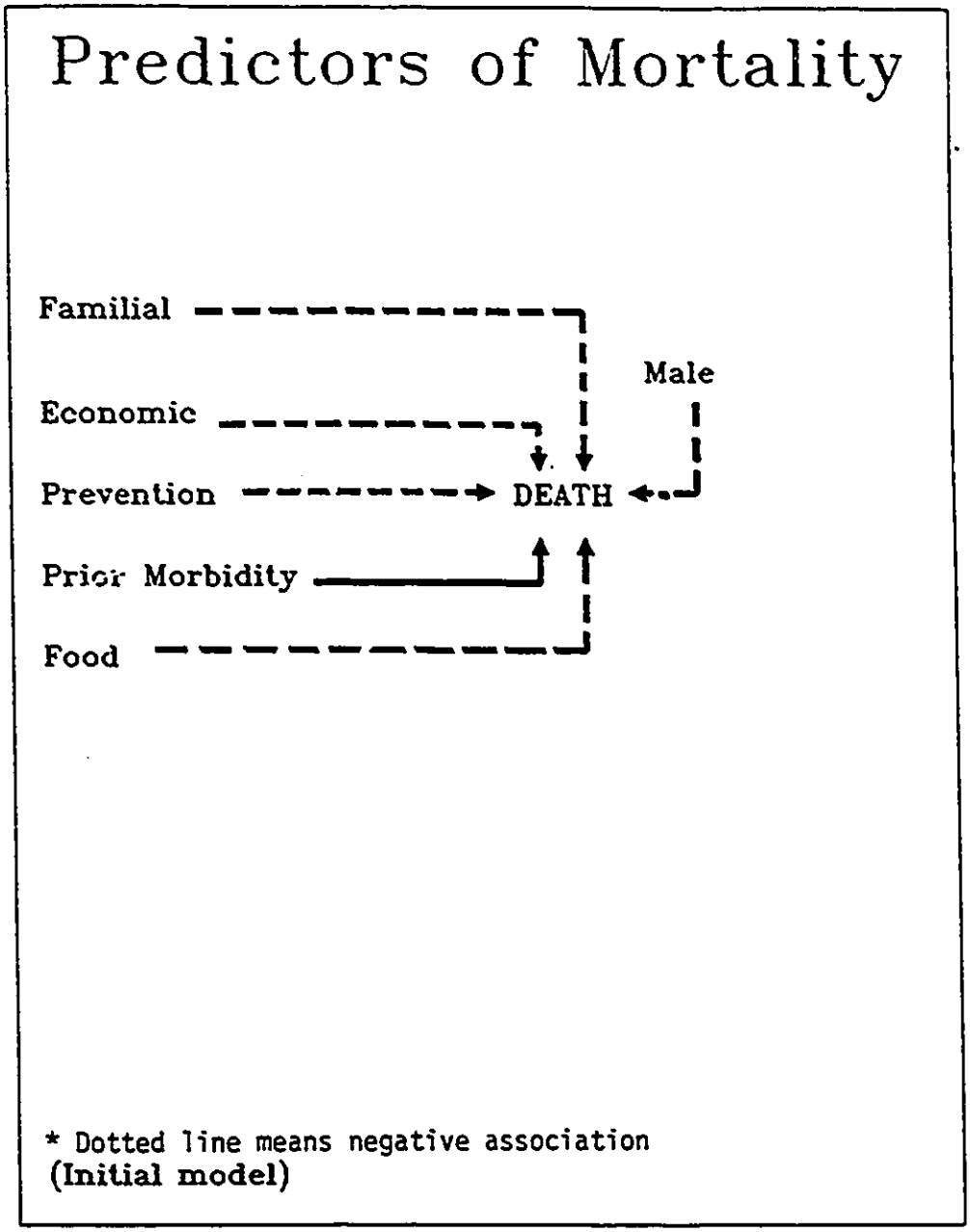


Figure 2 - Hypothesized Model of Predictors of Death

2.3 Objectives:

The objectives of this study are:

1) To develop adequate indicators of growth, using data from anthropometric assessment.

2) To develop composite indexes of familial circumstances, of economic level, of feeding practices, of food availability, of preventive health care, of prior morbidity, and of current morbidity.

3) To calculate path coefficients (see p. 77) for all the factors in the growth models and thereby to test the models.

4) To test and refine the models to help plan interventions that would prevent growth failure and reduce premature mortality in such a paediatric population.

5) To study subgroups of children with particular growth characteristics (e.g.: stunting, wasting, and underweight), and various age groups, in order to generate hypotheses about the most important determinants.

6) To assess what factors (as measured by the indexes developed in (2) above) are associated with a higher mortality.

2.4 Conceptual definitions:

Variables included in this data set are grouped into various indexes of the causal factors discussed above. These are defined as follows: **Feeding Practices** depict the type of supplementary food used (e.g. bottle milk, solid food) and their overlap with breast feeding, if any. **Prior Morbidity** represents illnesses which occurred before the beginning of the study. **Current Morbidity** means sickness occurring during the course of the study. **Familial Circumstances** denotes environmental characteristics of the household, composition of the family, education, and health of the parents. **Economic Status** stands for income and employment of the parents. **Food Availability** incorporates the kind of foodstuff seen in the household and frequency of consumption. **Prevention** includes the number of immunizations received and the frequency of visits to the under-5 health clinic.

2.5 Hypotheses:

1) **Infant feeding practices will be better in:**

- children who are healthy (and thus feed more easily)
- stable families
- wealthier families
- families with access to a more varied and abundant source of food
- mothers who visit the medical clinic more often

2) **Current infant morbidity will be lower in:**

- stable families
- infants with a healthy past
- wealthier families
- families that practice prevention (immunizations)
- families with access to a more varied and abundant source of food
- infants who have more types of food introduced before weaning

3) **Growth will be faster in children who:**

- have few infectious diseases
- have more types of food introduced before weaning

AND, THROUGH THEIR EFFECT ON THE TWO PRECEDING FACTORS,

- have access to a more plentiful food supply which is richer in protein content
- live in stable families
- live in wealthier families
- have more immunizations and well-baby clinic visits

4) **Mortality is more likely in children who:**

- have an unstable family
- have a poorer family
- have a low variety of food that is lacking in protein
- have mothers practising less prevention
- are ill more often
- are of female sex.

Chapter 3

3.0 METHODOLOGY

3.1 Study Design:

The data analyzed in this study was collected during a research project which involved the follow-up, over a two-year period (1970-1972), of a cohort of African children consisting of all the preschoolers from 1 month to 5 years of age living in two contrasting suburbs of Lusaka, the capital of Zambia in South-Eastern Africa. One of these suburbs was an approved municipal suburb, and the other an unauthorized squatter settlement. Altogether, 1342 children, belonging to 589 families, were followed longitudinally. Only families with at least one child under age 5 were included in the study.

It was thought that questions to the mother related to nutrition of the child might have changed her feeding practices. To examine this possible bias, half of the cohort was used as a control group and questions related to availability of various foods in the household were asked of these mothers only near the end of the period of observation. Food availability data for the control group is not available.

3.2 Data Collection:

Extensive baseline information was collected on each household and on each participating child during the initial interview. The variables measured covered some feeding practices, past morbidity of the child and of parents, certain health practices, socio-economic information, demographic and environmental data and, for half of the sample, availability of certain foods. At each of four follow-up six-monthly visits, the children were weighed and measured at home by the same pediatrician and by a research assistant. Dietary and morbidity histories were taken at these times and have been used to create indexes of feeding practices and of current morbidity.

Age accuracy:

Verifying age is critical when assessing growth. Age was determined by asking the mother both the child's age and date of birth. This was confirmed by a written record of birth for 53% of the children. In another 32% of cases, it was possible by clinical and other means to ascertain the child's age to within one month, and for another 10%, to within 6 months. In order to use as many of the cases as possible but at the same time to avoid much distortion of results from misclassification on growth parameters, I devised a weighting system to reflect the varying certainty of the child's age. If the birth date was confirmed by a

certificate, then the weight coefficient is 1.0; if the pediatrician could ascertain the age at +/- one month, then the weight is 0.8; if the age was estimated +/- two months, then the weight coefficient used is 0.5; for within +/- three months, the coefficient is 0.3; within five months it is 0.1; and within one year, it is 0.05. This means that during a regression analysis, a case with a weight coefficient of 0.5 will be worth only one half as much as a case with confirmed birth date ("weight" in this context means "relative value"). Although the differences in weight may look rather arbitrary, they were determined so that the larger the risk of age determination being wrong, the smaller the weight given to that case during the analysis. This permitted using the 15% of the children who did not have very precise age determination without letting them influence the analytical results to an undue degree.

Anthropometric measurements:

Supine length was measured for children of all ages, using a locally made wooden board with a fixed foot piece and movable head piece. Weight (in kg) was ascertained using a Salter hanging spring balance, the accuracy of which was checked every week. Arm circumference, although measured and available for analysis will not be used here as it was not found to be of superior value to weight for age and weight for length as a predictor of malnutrition (34).

Missing Data:

During the initial data collection, the intention was, after the baseline survey, to monitor each child four times over a two-year period, but in many instances this was not achieved owing to departure of the family, death of the child, or to the birth of another child into the family. Although strenuous attempts were made to locate absent children, there were fewer than four sets of anthropometric and health data for many of them. In fact there were 3294 sets of anthropometric observations in all; an average of 2.5 for each child. The first and the last measurements provided the lowest numbers of missing cases (about one third of the original sample of 1342); for this reason, the data on the children's weights and heights used for our analysis were derived from the last (fourth) measurements only and most of the indexes were derived from data collected at the beginning of the study except for Feeding Practices and Current Morbidity which were derived from a combination of data collected during all of the follow-up periods with missing information being replaced by the overall sample mean or by the most frequent (modal) value for the sample. This would tend to narrow the variance in some elements of my data but would permit more cases and more variables to be used in the analysis. I realize that by averaging the missing values, the usefulness of the information obtained

from the component variables of the two longitudinal indexes may be weakened. However, because of the already small number of variables available as components of these two indexes, it was important to try and keep as many of them as possible in the analyses to preserve the semi-normality of the two indexes' distributions (see Appendix C). Indeed, dropping the component variables with missing data may have brought a gain in precision of what the indexes were measuring but would have caused a larger departure from normality of the index's values towards a more explicit ordinal distribution since there would have been less possible values available to the two indexes. Using longitudinally collected data for the intermediate variables was necessary in order to maintain the temporal sequence that would permit causal inference. Various methods to address the problem of missing information in survey data is discussed in more depth in section 3.4, while discussing the development of indexes (p.44).

3.3 Operational Definitions:

Dependent variables:

Growth Outcomes: the raw anthropometric data comprise weight, and length as already noted. For purposes of data analysis, three indicator values of weight and length were calculated for each child. These are **Height-for-Age, Weight-for-Age,** and **Weight-for-Height.** These parameters

have been used in many previous studies and are accepted ways of monitoring human growth. Bairagi (3) advises that "to investigate the effect of a long-term factor such as family income or mother's education on child nutrition, weight-for-age or height-for-age should be used" as opposed to weight velocity which is more affected by short-term factors.

Using the American CDC "Anthropometric Software Package" (version 3.0) an attempt was made to derive a set of dependent variables expressed as percentiles of the U.S. NCHS standard growth curves by sex and age. These will be referred to as Height-for-Age Percentile (HAPCT), Weight-for-Age Percentile (WAPCT), and Weight-for-Height Percentile (WHPCT). Waterlow et al. (47) argue that the American NCHS growth standards may not be appropriate for comparison with other populations. Bhandari and Mandowara (4) state that U.S. growth standards may be suitable for Indian children up to the age of six months only and not thereafter. They argue that the North American standard will wrongly classify a large proportion of Indian children as malnourished. "From physiological view point also normal growth depends upon an orderly sequence of genetic, racial, constitutional, environmental, nutritional and endocrinal factors and Indian children have differed with American counterparts for generations in all these respects." Their arguments appear

valid for African children as demonstrated in a study published in 1984 by Ijsselmuiden (19) of 658 South African black pre-school children. The author found that 29.8% of the children were below the 3rd percentile of weight-for-age of the 1976 NCHS standards; this went up to 38.2% for the 1-2 year age-group. In the current study, a large proportion of the sample ended up being classified within the lower first percentile of the NCHS standard. The spread thus obtained was markedly skewed to the right and did not provide an ideal pattern (i.e. one approaching a normal distribution) for the use of a parametric method for data analysis.

To help assess further the degree and nature of the difference between the sample of Zambian children and the American standard, the mean length and weight for age and sex groups **at baseline** were plotted against the NCHS anthropometric curves (figures 3 & 4). For curves of Height-for-Age and Weight-for-Age, and within each gender group, age categories were as follows: one for each month interval from age 1 month to age 24 months; one for each two month interval from age 24 to 36 months; one for each four month interval from age 36 to 64 months. There were thirty-four age categories and an average of 15 cases per category. In an attempt to smooth the plotted curve, the mid-point was used for locating the category mean on the abscissa. For

example, the location "38 months" was chosen to place the mean of the 36-40 month male category on the "X" axis.

In the case of Height-for-Age and Weight-for-Age, the two curves were fairly close up to age 6 months, then diverged, the Zambian children resuming a course parallel to but below the American standard. For Weight-for-Height, the Zambian and American children did not differ significantly.

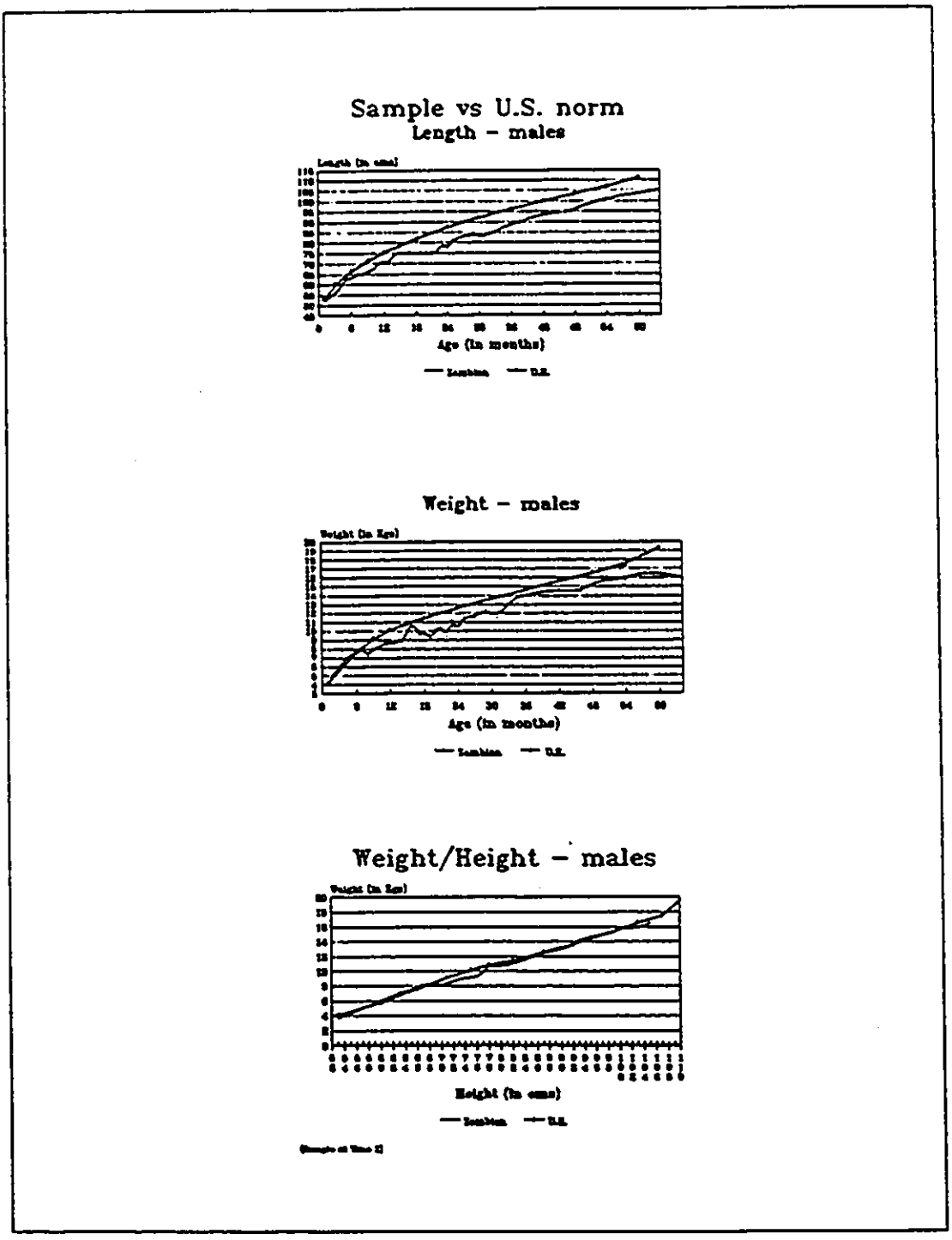


Figure 3 - Growth Curves: Comparisons with NCHS Standard - Males

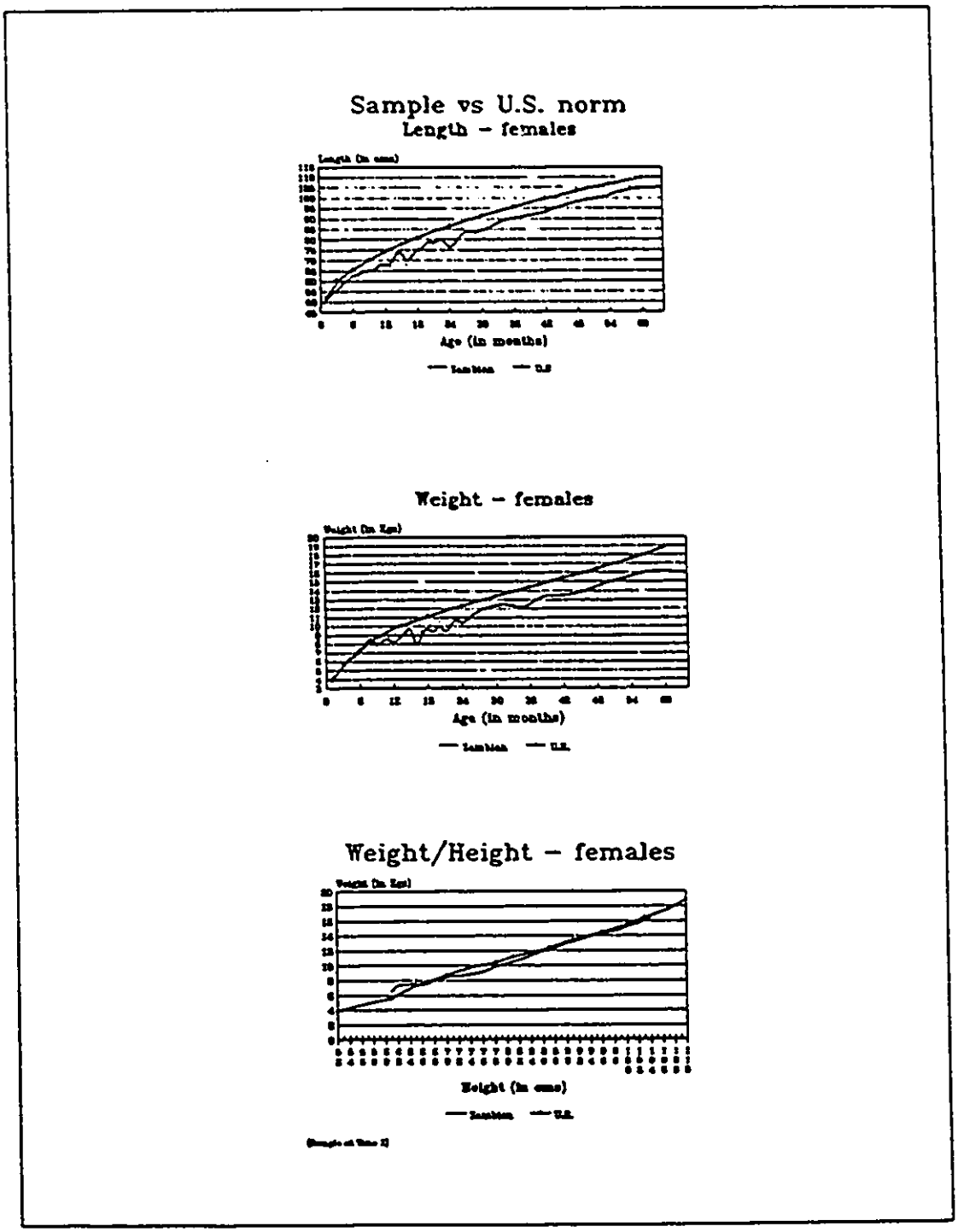


Figure 4 – Growth Curves: Comparisons with NCHS Standard – Females

Bairagi (3) refers to a study by Yarbrough C, Habicht JP, Malina RM, et al. which showed that both weight and height of Guatemalan children at each age from birth to seven years were distinctly lower than that of North American children but that weight-for-height of the two groups were comparable. This suggests that stunting (small stature) is probably the initial factor and that a reduced body weight is an adaptation to a short stature in order to reach an ideal proportion between height and weight. Johnston et al. (21), in a study of the growth of Mexican children, comment on a similar finding, suggesting "that the parents of failing children were themselves chronically malnourished during their own growth. This indicates that the etiology of chronic malnutrition, indicated here by growth failure, involves a significant generational aspect". Among the three growth deficiencies, stunting might be the least amenable to direct intervention if, as suggested, it depends more on factors having affected parental growth rather than on current, more direct environmental factors.

Because of the above caveats, I decided to derive a second set of outcome variables that compare the Zambian children among themselves, rather than to the U.S. standards. These used **Z-scores** (standardized to a mean of 0 and a standard deviation of 1) for Height-for-Age (ZHA4), Weight-for-Age (ZWA4), and Weight-for-Height (ZWH4). The Z-scores were

calculated within age and sex groups based on the mean and the standard deviation of each age and sex stratum (similar to the ones used for the anthropometric curves in figures 3 & 4) of the sample of Zambian children available at the end of the study (Time 4). Using the Z-score outcome variables allowed comparison of the sample to its own distribution and did not produce a bias by means of an outside comparison (such as the U.S. standard). It indicates relative growth among the Zambian sample.

Mortality Outcome: This is the eventuality of death occurring during the course of the study. It was analyzed as a dichotomous variable.

3.4 Analytical Methods for Growth Outcomes:

The data set has a large number (>74) of predictive variables. These include many dichotomous and several scaled (ordinal or continuous) variables. The dependent growth parameters are continuous. Various options were considered for analyzing this:

a) **Logistic regression:** it handles dichotomous and scaled predictors, but would require dichotomizing growth outcomes. Although using dichotomous predictors might help increase accuracy of exposure assessment, it would lose useful information concerning the growth outcomes. This may bring

about misclassification of outcome as well if the cut-off points are not chosen with care.

b) **Analysis of Variance:** it handles dichotomous determinants and scaled outcomes, but does not handle scaled predictors. It prefers a balanced design which is not the case with most observational data such as this one (as opposed to experimental data). It is usually difficult to pinpoint precisely the source of significant differences between strata, especially when many predictors are used. It then becomes complex to interpret the contribution of main effects and interaction terms to the overall result.

c) **Analysis of Covariance:** this might adjust for some of the confounders by treating scaled predictors as covariate and, in doing so, might simplify interpretation of the effects. However, in my analysis, I am interested not only in the effect of one predictor, but in the effects of many. Adjusting for the scaled predictors will lose information about their effect which is unacceptable.

d) **Multivariate Linear Regression:** it handles scaled outcome, and scaled predictors. Binary determinants can be included with the creation of dummy variables; however, this technique can be tedious when many nominal predictors are used.

From all these procedures, Multivariate Linear Regression (MLR) was chosen as the best method to test the proposed

growth models and to study the interactions between their components.

In view of the desirability to maintain a ratio of 10:1 between the sample size (especially the smaller subgroups) and the number of variables used in the analysis, and because of the large number of nominal predictors whose handling would prove cumbersome, it seemed preferable to try and reduce the number of independent factors. To achieve this, some conceptual considerations made it appealing to form index-type variables that combined several predictors. These considerations and the resulting indexes are discussed below (p. 41). It was not felt necessary to consider the age and gender of the child in the analysis on growth outcomes since the dependent variables were already adjusted for age and gender, growth being expressed relative to other children of the same age and sex.

Path Analysis was used once the regression coefficients were available. This procedure is applied to causal models and used "to combine quantitative information given by the correlations with such qualitative information as may be at hand on causal relations to give a quantitative interpretation" (24). The technique was useful as well to help partition the direct and indirect ways by which predictors affect growth. The necessary assumptions for

this type of path analysis are that: 1) the relations among the variables in the model are linear, additive, and causal; 2) the residuals are not correlated among themselves, nor are they correlated with the variables in the system; 3) there is a one-way causal flow in the system; 4) the variables are measured on an interval scale (24).

Suitability of Growth Indicators:

In preparation for multivariate analysis of the data, the distributions of the growth outcome indicators were examined. As previously stated, the three outcome variables derived as percentiles from the U.S. standard growth curves (HAPCT, WAPCT, WHPCT) were found to be far from normally distributed. More than 20% of the sample was located below the 1st percentile of the U.S. standard for Height-for-Age and more than 10% for Weight-for-Age; predictably Weight-for-Height was distributed closer to a normal pattern with only 60% of cases falling below the 50th percentile (slight skew to the right). After doing a natural logarithmic transformation of these percentile values, the distributions of the first two variables behaved somewhat better but still not good enough for an analysis based on the assumption of a normal distribution. The Weight-for-Height distribution moved farther from the normal when logarithmically transformed. The distributions of Z-score-derived variables of the Height-for-Age (ZHA4), Weight-for-Age (ZWA4), and

Weight-for-Height (ZWH4) within the sample group were much closer to a normal distribution (see Appendix D).

Cogill (quoted in Bairagi (3)) had demonstrated in his thesis that it made no important difference whether the indicators were expressed as percentiles of the NCHS standard or as Z-scores. In my sample, the U.S. Percentile and Z-score indicators are highly correlated (Pearson coefficients for Height: 0.76; for Weight: 0.84; for Wt/ht: 0.79). However, the U.S. standard-derived indicators being far from normally distributed, the better behaved Z-scores were used instead in the analysis, as recommended by Waterlow et. al (47). Nevertheless, because of the closeness to normality of the WHPCT variable, this variable was also analyzed and the results are compared to the Z-score of Weight-for-Height (ZWH4) to check for robustness of the model for that growth parameter.

Independent variables - Indexes:

At the stage of developing this thesis, I had expectations that general characteristics, such as financial means, familial stability, food availability, preventive measures, feeding practices, and general morbidity would influence growth. However, these are abstract or, at best, vague concepts that cannot be measured by a single variable in the available data set. Choosing only a few original variables

as proxies of these concepts may provide a better defined measure of exposure but will narrow its scope too much. Therefore, I took the fundamental decision to create index variables to represent the determinants described in my hypotheses stated above. These indexes combine a number of individual variables from the raw data, and seek to provide a more comprehensive reflection of the processes involved in these concepts.

Seven indexes were formed to represent:

1) The desirable feeding practice of overlapping various foods with breastfeeding before weaning the child (feeding practices).
2) The degree of morbidity experienced by the child during the study (current morbidity).
3) The degree of morbidity experienced by the child prior to the beginning of the study (prior morbidity).
4) The availability and frequency of various foodstuffs in the diet of the household (food availability).
5) The characteristics of the familial environment such as level of education and health of the parents, number of sibs, presence of environmental amenities, etc. (familial circumstances).
6) The level of economic wellbeing as measured by income and by type of job of the parents (economic status).

7) The amount of preventive practices followed by the parents as measured by attendance at well-baby clinics and number of vaccines administered to the child (prevention).

Rationale for content of Indexes: As mentioned above, concepts such as morbidity, prevention, socio-economic status etc. are quite vague and may be interpreted in various ways by different people. They do not always mean the same thing in other contexts. Furthermore, their setting in a distinct demographic milieu might warrant a different way to define them. When reviewing previous research in the field, it became evident that no simple method of deriving indexes has been conceived, especially when using a secondary data set with a fixed variety of possible components (numbering 132 in my data set). For example, Afifi (2) in his analysis of factors affecting the growth of Egyptian children used income, livestock, land status, education, family structure, source of water, ownership of latrine and radio, presence of adequate ventilation and of flies as components of a socio-sanitary indicator. Crean et al. (8), in a similar analysis of patterns of infant feeding and malnutrition in Indian children, chose overcrowding, house construction, presence of electricity, ownership of electronic equipment (radio, T.V., etc.) and ceiling fan as items of a socio-economic

score indicator. Both of these authors attributed scores to each of these items based partly on previous experience with such populations and partly on common sense. Abramson (1) in his excellent text on "Survey Methods in Community Medicine" mentions the fact that composite scales are often weighted arbitrarily, on the basis of the apparent respective importance of the items. Supporting myself on these authors and on previous empirical knowledge gained during a long stay in Africa, I chose seventy-four of the available variables to represent the somewhat abstract concepts that I was attempting to measure.

Development of Indexes:

Most of the original continuous variables used to form the indexes were categorized, based on meaningful cut-off points (such as already known or accepted differences in risk; for example, mother's age <15 years and >35 years, or parity >5 and bearing in mind the number of cases per category, and the general distribution of the sample within that variable. Unless stated otherwise, missing values of the variables forming most of the indexes were replaced either by their mean value (in the case of continuous variables) or by their 'modal' value (in the case of categorical variables) to maximize the number of cases used in the analysis although I realized this will cause a loss of variance in the sample and may bring about, consequently, an underestimation of the

size of the effect of the predictors on growth and on mortality, especially if data are not missing at random.

Missing values in data set:

The number of missing values was considerable, ranging from a minimum of one to a maximum of 58 per case with a mean of twelve. Appendix K provides the frequencies of missing values for most variables used in the construction of the indexes. Appendix L shows results of a comparison of cases with 5 missing values or less with the rest of the sample which had more missing values. The results tend to show a potential for bias in replacing missing values by an "average" value since some of the responses on variables used for some of the indexes were found to be significantly different between the two groups, especially for components of the familial (social) index and for components of the two indexes on infant nutrition.

Missing survey data usually falls into one of two categories: 1) total non-response (when one of the individuals in the sample has not provided any answers) and 2) item non-response (when only some answers to specific questions are missing). Total non-response arises because of: refusals, inability to participate, not-at-homes, and untraced elements while item nonresponse originates because of item refusals, "don't knows", omissions and answers

deleted in editing. Different adjustment methods are needed for these two distinct circumstances. Kalton & Kasprzyk, in a paper published in 1986 (22), review in detail the various methods available to the researcher in compensating for missing data. I will summarize here some of the procedures they describe. For more details, the reader is encouraged to refer to their very informative article.

WEIGHTING:

As a rule, weighting adjustments are used for total nonresponse. The essence of all weighting adjustment procedure is to increase the weights of specified respondents so that they represent the nonrespondents. In the present case, not having at my disposal information on the population from which the study sample came from, I would have been limited to sample information to use in weighting. An effect of sample weighting adjustments is to increase the variance of the survey estimates, i.e. there is a loss of precision from weighting. The smaller the weighting classes, the more unstable their response rates, and consequently, the larger the variation in the weights. There is therefore a trade-off to be made between bias reduction and variance increase when one attempts to stratify weighting classes with more refinement.

Another method, that of weighting with response probabilities has not seen wide application. This method assumes that all population elements have opportunities of responding to the survey. One then estimates the response probabilities for responding elements. These elements are then given nonresponse adjustment weights that are in inverse proportion to their response probabilities. One approach for estimating response probabilities is to regress status (1 for respondents, 0 for nonrespondents) on a set of variables available for both respondents and nonrespondents, using a logistic regression. This method is most appropriate for situations of partial nonresponse when some data are collected for a sampled element but a substantial amount of data is missing, as for instance when the nonrespondents are lost after the first wave of a panel survey.

IMPUTATION:

In order to retain all survey responses for elements with some item nonresponses, the usual adjustment procedure produces analysis records that incorporate actual responses to items for which the answers were acceptable and imputed responses for other items.

Deductive imputation, when edit checks constrain missing response to only one possible value, is the ideal form of

imputation. To a limited degree, that methodology was used with the present data set during data cleaning. Various other methods can be used for the purposes of imputation. The overall respondent mean can be attributed to all missing responses on the item. A more refined method divides the sample in classes according to values of other variables and attributes the class mean of respondents to the nonrespondents within the class. Random imputation can also be used when a respondent is chosen at random and his response is given to a nonrespondent. This procedure may be refined as previously by dividing the sample into classes first, based on other variables whose responses are available for the whole sample. Derivatives of this latter method are called "hot-deck" methods where nonrespondents and respondents are matched in terms of the auxiliary variables, and nonrespondents are assigned values from matched respondents. Regression on a set of auxiliary variables for the variable for which imputations are required is a useful way to derive an equation that can be predict the values of the missing responses. Without careful modelling, there is a serious risk of poor imputation and the imputed value not being a feasible one. A variant method, called "predictive mean matching", makes use of regression and of the "hot deck" process, and avoids the problem by matching the nonrespondent with the respondent with the closest predicted value. Overall mean,

class mean, and regression methods are obviously not suitable for use with categorical or discrete variables. A notable advantage of the "hot-deck" methods is that they always give feasible values since the values are taken from respondents.

Imputation can have harmful effects on all analyses of relationships, often attenuating the associations between variables. If the relationship between two variables is to form an important part of the survey analysis, the one should be used as an auxiliary variable in imputing for missing values of the other. This suggests that when one record has several missing related values, they should be taken from the same donor. Imputation can also distort estimation of standard error, generally causing underestimation with overstatement of confidence levels. Methods of multiple imputations, using either an average of the imputed values derived or a randomly selected one have been devised to address these problems and further reduce the loss of precision in the survey estimates. A major inconvenience with the use of multiple estimation, is the additional computer analysis needed.

In view of the large amount of missing values in many variables in the present study sample, and of the numerous 'auxiliary' variables available for stratification or

matching, anything more than an overall means or modal imputation procedure would have demanded considerable computer time and human labour without the guaranty of any gain of power in data analysis. One should be reminded that none of the cases in the data set had complete information on all variables.

The resulting composite indexes were computed as follows:

INDEX OF FEEDING PRACTICES (Feedindx):

Fifteen original variables were used in its computation. Since the introduction of supplementary foods prior to weaning was judged desirable by various researchers, this was represented by counting the number of different feeding methods used concurrently with breast feeding (overlap) by the mother. Its score is equal to the sum of the values of all the "overlaps", for baseline and each of times 1, 2, 3, and 4. An overlap of breast feeding with bottle feeding or with porridge feeding had a value of "1"; an overlap of breast feeding with 'nsima' (flour-based food) received a value of "2" while an overlap of breast feeding with solid foods was worth "3". Those children with missing data on supplementary foods at baseline and who were not yet weaned at the start of the study were given modal values, which correspond to an overlap on porridge and nsima, but not on bottle feeding. Cases with missing data at Time 2 and Time 3 were also given modal values for those same items by

considering no overlap (value of "0") for bottle feeding or for porridge (frequency above 90% in non-missing cases) but an overlap for nsima and for solids (frequency of 60 and 70% in non-missing cases).

INDEX OF CURRENT MORBIDITY (Morbid2):

Eight original variables were used in its computation. This represents illnesses experienced by the child during the period of follow-up. Its score comes from a count of the following morbidity variables: 'diarrhea' for each of times 1,2,3,4; 'other symptoms of disease' for each of times 1,2,3,4. The absence of the symptom was given the value "0"; the presence of it received the value "2"; the unknown (missing data) received the midpoint value "1".

INDEX OF PRIOR MORBIDITY (Morbid1):

Fifteen original variables were used to represent illnesses experienced by the child prior to the beginning of the study. A higher value was given to the illnesses that were found, in previous research or from clinical experience, to be more important. The score was the sum of the values (in parentheses):

Ever been hospitalized (no "0", yes "1")

Ever had skin disease (no "0", yes "1")

Ever had ocular disease (no "0", yes "1")

Ever had ear/nose/throat affection (no "0", yes "1")

Ever had hematologic problem (no "0", yes "2")

Ever had respiratory ailment (no "0", yes "2")

Ever had neurological illness (no "0", yes "2")

Ever had surgery (no "0", yes "2")

Ever had intestinal parasites (no "0", yes "2")

Ever had enteritis (diarrhea) (no "0", yes "3")

Ever had measles (no "0", yes "2")

Ever had home-made remedies (no "0", yes "1")

Ever had remedies from witch-doctor (no "0", yes "1")

Ever had difficulty with lactation (no "0", yes "2")

Born premature or from twin pregnancy (no "0", yes "1")

INDEX OF FOOD AVAILABILITY (Nutridex): (see Appendix M for details)

Seven original variables were used to represent food variety and adequacy. The score first sums the frequencies with which selected foods were recorded in the diet, multiplied by the value attributed to each type of food. The food value was based on the relative protein content of the particular food as follows: meat "2"; fish "2"; milk "2"; beans "1.5"; staples "1". Second, the values for presence of a garden (value "2") and amount of money spent on buying food (value from "1" to "5") were added to the total.

INDEX OF FAMILIAL CIRCUMSTANCES (Familial):

Seventeen original variables were used in its computation. This index represents a mixture of social factors, environmental elements and health status of other members of the household. Its score is derived from the sum of values (in parentheses) of the following variables (some of the variables were weighted, as documented in Appendix M):

Location of residence (squatter town "0", legal suburb "1")

Number of older children in house (none "0", less than three "1", more than three "2")

Number of younger children in house (several "0", one "1", none "2")

Health of father (dead "0", sick "1", healthy "2")

Health of mother (dead "0", sick "1", healthy "2")

Age of mother (high risk "0", low risk "1")

Parity of mother (more than 5 "0", 5 or less "1")

Previous deaths of sibling (some "0", none "1")

Status of marriage (single parent "0", unstable relationship "1", stable "2")

Education of father (actual grade achieved, grade 1 to 12+)

Education of mother (actual grade achieved, grade 1 to 12+)

Father's alcohol intake (a lot "0", little "1", none "2")

Mother's alcohol intake (a lot "0", little "1", none "2")

Source of water in house (by lorry "0", well "1", from tap "2")

Handicapped child in household (yes "0", none "1")

Location of birth of case (not in hospital "0", in hospital "1")

Identity of relatives living with child (grand-parents "0", father only "1", mother only "2", both parents "3")

INDEX OF ECONOMIC STATUS (Econdx):

Six original variables were used in its computation. It represent the various sources of income of the household. Its score is obtained from the sum of values (in parentheses) of the following variables:

One half of the value given to the father's job category (low "0" to high "8")

One half of the value given to the mother's job category (low "0" to high "5")

Father's income category (very low "0", low "1", average "2", high "3", very high "4")

Mother's allowance category (very low "0", low "1", average "2", high "3", very high "4")

Mother's income (none "0", some "1")

Mother's other financial support (none "0", some "1")

INDEX OF DISEASE PREVENTION (Previndx):

Five variables were summed:

One half of the number of visits to the under-5 clinic; twice the combined total of smallpox vaccinations, DPT

vaccinations, poliomyelitis vaccinations, and BCG vaccinations. BCG immunization was allotted a value of "0" for none; "1" for late; and "2" for early. Vaccines were thus given four times more weight than clinic visits.

Suitability of Indexes:

A look at the distribution of the indexes revealed that most of them were not quite normally distributed with the exception of the Familial Circumstances Index which had a Z-value of .995 with a p-value of .275 (two-tailed) on a Kolmogorov-Smirnov Goodness of Fit Test. The Economic Status index has a distribution curve fairly close to normal and the values of the Prior Morbidity index similarly approach normality, although skewed to the right. The histograms of the distributions of indexes themselves may be found in Appendix C.

Refinement of the Indexes: The indexes above were derived based on elements found significant by previous research and based on intuitive sense. Before use in the analysis, it was desirable to ensure that the component parts did, indeed, contribute meaningfully to the composition of each index. I used the Scale Reliability Analysis procedure (SPSS-PC V. 3.0 package) and Cronbach's Alpha statistic in trying to eliminate components of the initial indexes that were less correlated with the others and might reduce, or

even negate, the effect of other better internally correlated components. Any component whose removal would increase the value of the Alpha statistic by even a small amount was removed from the index and any component whose removal would decrease the Alpha score substantially (e.g. 0.10) was kept in; other components whose removal would not affect the value of the Alpha statistic were either kept in or removed depending on their empirical importance in defining the concept being measured. Resulting indexes are described in Appendix B. The following data show the internal consistency of the indexes after revision as measured by Cronbach's Alpha.

Index	Original Index Alpha	Reduced Index Alpha
Prior Morbidity	.4580	.4382
Familial	.3472	.3877
Economic	.2906	.4753
Prevention	.7913	.8354
Food	.6369	.7456

The object of such an exercise was to arrive at simpler indexes with better internal consistency and hopefully a better measure of exposure. For most, the alpha varied little as a result of this procedure. Subsequent analyses proved that the new set of reduced indexes did not, in most cases, explain any more of the variance of the growth

outcome indicators than the original indexes. Consequently, only the regression results obtained with the original indexes will be presented. The 'longitudinal' indexes of Feeding Practices and Current Morbidity were not submitted to this reliability exercise as each one had only two kinds of components, both of which were considered important.

Regression on Intermediate Factors:

To test hypotheses number 2 (predicting feeding practices) and number 3 (predicting morbidity), multiple linear regressions used as dependent variables the Feeding Practices index (FEEDINDX) and the Current Morbidity index (MORBID2) measured during the course of the study (Times 1,2,3,4). All the indexes made up of variables measured at the start of the study were used as independent predictors of these two intermediate outcomes. A stepwise method was employed, with a p-value of .10 used as the entry and removal criteria. This means that, if the partial correlation coefficient of the variable being evaluated with the outcome of interest had a degree of statistical significance smaller than 0.10 after adjusting for the effect of the independent variables already in the equation, that variable was included; after its inclusion, the significance of the corresponding partial correlation coefficients with the outcome of all the other variables already in the equation was recalculated and if any was

greater than 0.10, the coinciding variable was excluded. These criteria were chosen instead of the usual p-value of 0.05 because Greenland (17) counselled "markedly increasing the significance levels of the tests" to compensate for the tendency of "conventional significance-testing algorithms (such as stepwise regression)" to "underselect confounders and so yield invalid point and interval estimates". "Confounders", as used here, means: other independent variables of interest. In this thesis, the emphasis is not so much on specific hypothesis testing but rather on model building so it was important for me to study the effects of other independent variables, even if not quite statistically significant (at the 5% level) in this data set.

Regression on Growth Outcomes:

To test the inferences implied by hypothesis number 1 (on growth) similar multiple linear regressions were done, this time using each of the three Z-score growth indicators and also WHPCT, in turn, as a dependent variable and all the developed indexes as predictor variables. The intermediate (longitudinal) indexes FEEDINDX and MORBID2 were first introduced in a stepwise fashion and then, in order to check for any residual effect, the other indexes made up of variables measured at the beginning of the study were introduced, again in a stepwise fashion. Entry and exit

criteria used were the ones in effect in the previous regressions.

As stated previously, the Food Availability index (NUTRIDEX) was available only for half of the sample; consequently, that part of the sample was run another time on all the above regressions after adding the Food Availability index to the other independent predictors. The results are provided in separate tables ('small sample').

Although Leigh (25) advises against the use of a stepwise method and the use of the 'amount of variance explained' as a help to decide whether an independent variable should be included in a model, the use of the direct method (all variables forced in), with significance set at a probability level of 0.10 gave essentially similar results.

Refinement of the model:

Goodness of fit on the model will vary for different groups so it is important that it be tested on malnourished subgroups as well as on different age categories, as these groups may be differentially affected by morbidity and by variation in socio-economic level.

Gomez (16) and Jelliffe (20) have defined cut-off points for malnutrition that are well accepted, using the American

standard percentiles. There are no similar, sanctioned, cut-off points for Z-scores used within the sample. For the purpose of defining malnourished subgroups, I used the cases falling within the lower 10% of the distribution of Z-scores in each growth parameter for Height-for-Age, Weight-for-Age, and Weight-for-Height after noting that their distributions (Appendix D) were close to normal. A similar lower 10% cut-off was used by Victora et al (44) in their study of Brazilian children. The resulting subgroups will permit study of three growth retardation categories:

Stunted: Height-for-age less than 10th percentile of the Z-scores distribution for that indicator.

Underweight: Weight-for-age less than 10th percentile of the Z-scores distribution for that indicator.

Wasted: Weight-for-height less than 10th percentile of the Z-scores distribution for that indicator.

Children defined as Stunted (N = 86), Underweight (N = 77), or Wasted (N = 114) were analyzed separately. Most of the published studies to date were done on populations of malnourished children and not on a general population. In order to compare my findings to theirs, it became important to see if predictive factors significant for these

malnourished subgroups were different from the ones found for the entire sample of Zambian children. Unfortunately, because of the small number of cases available for such analyses and because of the desirability to maintain a case/variable ratio of at least 10:1 in order to ensure reliability in results of correlation, it was not possible to analyze separately the subgroup containing the Food Availability variable. This variable was not used for the analysis.

For age-group analysis, the chosen age categories were: 'less than 6 months of age' and 'more than 6 months of age'. This age cut-off was chosen based on the different fit of the growth curves for these two sub-groups when compared with the curves of their American counterpart. This was evident from figures 3 & 4 where the younger group (less than 6 months) followed more closely the N.C.H.S. growth standards than their older compatriots. Lopez de Romana et al. (28) found a similar age-related pattern among their sample of Peruvian children as their compared growth curves demonstrate. Factors that may be operative among the less than six months old may be different from the ones that are important for the older group. This possibility deserved investigating.

3.5 Analysis of Mortality:

Discriminant analysis appeared the best method of testing the model for predictors of mortality. At this point, I had formed scaled indexes of predictor variables and my outcome was dichotomous (death/survival). Although logistic regression also handles dichotomous dependent variables, it works best when predictors are categorical. Categorizing the indexes would mean losing precision in exposure assesment. With discriminant analysis, the following assumptions about the data should be met in order to minimize the probability of misclassification: 1) each group must be a sample from a multivariate normal population, and 2) the population covariance matrices must all be equal. Box's M test, based on the determinants of the group covariance matrices is a good test of these two assumptions as the test is also sensitive to departures from multivariate normality and will be used to verify the validity of the assumptions.

The analysis sought to differentiate between the survivors and the children who died. The predictors used were: the indexes of Prior Morbidity, Economic Status, Familial Circumstances, degree of Prevention, and gender. Gender was determined by some researchers to have an influence on the quality of health care, or on the amount of food a child was given. The direction of this effect varied with the

culture studied. My initial model postulated that being male would have a protective effect in an African culture. For the discriminant analysis, the intermediate outcomes of Feeding Practices and of Current Morbidity were not used as data on their components would stop being collected after the child's death. The index of Food Availability was again available for only half of the sample so the analysis was repeated after its inclusion. Independent variables were included in the equation by the stepwise algorithm if their partial correlation coefficient maintained a p-value of less than 0.10. As shown in Appendix N which displays the distributions of the index variables used and a Kolmogorov-Smirnov goodness-of-fit test for the total sample and by group (alive, dead), only Familial Circumstances (Socsandx), and the dead group of the Economic Status are significantly similar (at p-value of 0.05) to a normal distribution. Nevertheless, it was decided to go ahead with the discriminant analysis while keeping in mind that the derived equation obtained would probably be less than ideal.

Individual Variable Analysis:

In order to raise hypotheses that could be tested in a subsequent study and to support findings of the discriminant analysis, twenty one original variables and dichotomized composite indexes were studied regarding the strength of their individual association with the likelihood of death in

the children. To compensate for multiple testing, only the associations having a p-value of at least 0.002 should be considered statistically significant, so that the overall level of significance is <0.05. This is a conservative way of bringing down to 5% the possibility of a chance association. Although a Chi-square distribution is not similar to a "t" distribution, a p-value of .002 approaches the p-value of the t-statistic (3.02) recommended by Bonferoni's table in Miller's book (35) when 20 t-tests are performed on a sample of more than 120 subjects for an alpha-error level of 5%.

Chapter 4

4.0 RESULTS

4.1 Description of the Sample:

Table 1 lists the means and standard deviations of some of the continuous and ordinal variables used to form the indexes and of the percentiles of the NCHS growth standards. The average family had five to six children, with poorly educated parents and a mother in her mid-twenties. The median age of the children at the start of the study was 19 months of age. The average child was weaned at about one year of age and had been introduced to bottle feeding at five months, had received his first porridge at six months, and had eaten his first nsima (a more solid, flour-based food) at about nine and one-half months of age. Height-for-Age and Weight-for-Age Percentiles were low by U.S. standard but body proportion (Weight-for-Height Percentile) was comparable to the North American child.

TABLE 1

Means & Standard Deviations
of Some Original Variables of the Sample

Variable	Mean	St Dev.	Median	No.
No. older sibs	4.3	2.0	4	1342
No. younger sibs	2.4	0.9	2	1342
Mother's age	27.3	6.7	27	924
Mother's parity	5.3	2.7	5	1341
Father's educ.(gr.)	3.8	2.9	4	1215
Mother's educ.(gr.)	1.7	2.3	0	1319
Age of child T1 (mo)	22.6	18.1	19.3	1173
Age at weaning(yr)	1.1	0.4	1.1	767
Age at bottle (yr)	0.4	0.4	0.2	493
Age 1st porridge	0.5	0.2	0.5	801
Age 1st 'asima'	0.8	0.4	0.7	767
H/A %tile at T1	16.9	20.3	8.4	1088
W/A %tile at T1	27.2	26.7	17.7	1172
W/H %tile at T1	52.2	29.9	53.5	1080
H/A %tile at T4	14.7	19.4	6.5	976
W/A %tile at T4	20.6	21.6	12.4	1007
W/H %tile at T4	43.5	26.2	41.0	976

Categorical variables of the sample which were used as components of the indexes were distributed as follows. About 49.4% of the children lived in the approved city suburb, the rest in the squatter town. Most (97.7%) fathers were alive and well, 1.1% were sick, 1.2 % had died; for mothers, these proportions are 96.6%, 3.2 %, 0.2% respectively. Sixty-seven percent of mothers fell within the 21-34 age group and sixty-eight percent had a parity of 5 or less. Thirty-five percent of families previously experienced the death of a child. Seventy-one and one half percent of marriages were stable (good relationship), 16.4% were unstable (fights, or just living together), 12.1% were single parents. Four percent of fathers were unemployed, 52.7% were labourers, the rest was diversified.

Most (78.6%) mothers were housewives with no other employment while 13% were sales-ladies (at the market or door to door). Nearly 75% of them had no personal income. Nearly half of households had a vegetable garden and 15.4% were fed by a relative or friend. Fifty-nine percent of fathers drank a fair amount of alcohol daily, while 32.9% reported never drinking it; for mothers, the proportions were 7.4% and 77.8% respectively. Based on the baseline household survey, half of households obtained their drinking water from a well, 47.4% from a nearby tap, and 2.4% via a lorry. Few (1.8%) of the families had a handicapped child. Fifty-one percent of the children in the study were males. Nearly 2.4% had been born as premature singleton and 4.2% as twins. Forty-three percent of study children were born in hospital, most of the others were born at home. Eighty-seven percent of study children were living with both parents at the time of the study, 8.5% with the mother only, 0.7% with the father only, and 3.4% with grand-parents. Seventeen percent of mothers had experienced difficulty with lactation. Twenty-two percent of the children were ever hospitalized for an illness and 9.9% had had a clinical diagnosis of malnutrition in the past. Sixty-three percent of the children had measles prior to the study; 16.9% had suffered from a neurological ailment; 3.9% from an ENT problem; 22.5% from a skin condition; 9.7% from a respiratory illness; 5.9% from an eye infection; 6.6% from a

blood problem; 7.8% from intestinal parasites; and 17.2% from enteritis. Almost 86% had never received any of the DPT vaccines. About seven percent were treated by the local witchdoctor and 16% had received home-made remedies.

4.2 Regressions in general:

Throughout the presentation and discussion of results on correlations and regressions, the use of the terms "positive" and "negative" will refer only to the direction of the association and do not necessarily imply a favourable or a detrimental effect in terms of what is beneficial to the individual.

The technical graphs (Normal probability plots) in Appendix I show that the models should be appropriate for the data as the residuals ('errors') are independent, normal values since the expected values are linearly correlated with the observed ones and their standardized mean is around the "0" mark with a reasonably even dispersion about that mean. A possible exception is the Feeding Practices index which follows a step pattern rather than a smooth line.

TABLE 2

Regressions on Intermediate Variables - Large Sample [1]
(Regressions Weighted By Accuracy of Birth Date)

Dependent Variable = Feeding Index (N = 938)

Variable	B[2]	Beta[3]	p-value	Cumulat. Variance (R^2)
FAMILIAL	-.03211	-.05753	.0782	.00224

Dependent Variable = Current Morbidity Index (N = 938)

Variable	B	Beta	p-value	Cumulat. Variance
FEEDINDX	.23905	.16300	.0000	.02698
MORBID1	.10448	.12312	.0001	.04126
FAMILIAL	-.05187	-.06338	.0479	.04425

TABLE 3

Regressions on Intermediate Variables - Small Sample
(Regressions Weighted By Accuracy of Birth Date)

Dependent Variable = Feeding Index (N = 414)

Variable	B	Beta	p-value	Cumulat. Variance (R^2)
PREVINDX	.03301	.10580	.0314	.00879

Dependent Variable = Current Morbidity Index (N = 414)

Variable	B	Beta	p-value	Cumulat. Variance
NUTRIDEX	-.00978	-.19363	.0000	.03337
FEEDINDX	.27719	.18637	.0000	.06636
MORBID1	.08543	.10527	.0000	.07520

[1] Large and Small Samples refer respectively to the analyses of all cases, or of the subset for which data on food availability was collected.

[2] Unstandardized Partial Regression Coefficient

[3] Standardized Partial Regression Coefficient

4.3 General Sample:

Predictors of Feeding Practices: As seen in table 2, the only variable having an effect (negative) on the Feeding Practices was the Familial Circumstances. Even then, the variance it explained was negligible (0.2%). When only the half of the sample on which Food Availability variables were obtainable was analyzed (table 3), the Prevention index was the only one that reached significance, with a positive effect but again, did not explain much of the total variance (0.9%).

Predictors of Current Morbidity: As depicted on table 2, Feeding Practices had the largest positive effect on Current Morbidity (increased it), followed by Prior Morbidity whereas the Familial Circumstances had a smaller and negative effect as shown by the beta weights (standardized partial regression coefficients). The total variance explained, although not very large (4.4%) was nevertheless much larger than with Feeding Practices as the dependent. When the smaller sample is analyzed (table 3), Food Availability appears to be the most important variable with a negative effect, followed by Feeding Practices and by Prior Morbidity, both with positive effects. The variance explained was improved by including Food Availability (7.5%).

Zero-order Correlations with Growth Indicators:

The Pearson correlations between independent and outcome variables were low:

Independent	Dependent Variables		
	ZHA4	ZWA4	ZWH4
FEEDING INDEX	-.046	-.034	-.042
CURRENT MORBIDITY	-.013	-.040	-.045
PRIOR MORBIDITY	-.150	-.189	-.169
FAMILIAL CIRCUMSTANCES	.095	.066	.039
ECONOMIC INDEX	.082	.086	.084
PREVENTION INDEX	.030	.039	.070
FOOD AVAILABILITY	.113	.082	.046

Nevertheless, because of the large sample all of the coefficients with Prior Morbidity reached the .001 level of significance and those with Economic Status reached the .01 level of significance. Prior Morbidity shows the highest correlation of all of the indexes. Economic Status has the second strongest association with growth as measured by any of the three outcome indicators. Finally, better Familial Circumstances seem to exert a positive effect on the child's height. The values for Food Availability do not reach a high level of significance because of a smaller (half) sample size. A better Food Availability appears to have a positive effect on height and weight but not so much on the ratio of "weight for height".

TABLE 4

Regressions on Growth Outcome Variables - Large Sample
(Regressions Weighted By Accuracy of Birth Date)

Dependent Variable = Zscore Height for Age at Time 2 years
(N = 911)

Variable	B[1]	Beta[2]	p-value	Cumulat. Variance (R^2)
MORBID1	-.05086	-.14928	.0000	.02140
FAMILIAL	.03100	.09382	.0042	.02915

Dependent Variable = Zscore Weight for Age at Time 2 years
(N = 938)

Variable	B	Beta	p-value	Cumulat. Variance
MORBID1	-.06425	-.18713	.0000	.03468
ECONDX	.03287	.08156	.0110	.04031

Dependent Variable = Zscore Weight for Height at Time 2 years
(N = 911)

Variable	B	Beta	p-value	Cumulat. Variance
MORBID1	-.06436	-.16012	.0000	.02756
ECONDX	.03890	.082262	.0115	.03281
PREVINDX	-.01275	-.05453	.0977	.03467

[1] Unstandardized Partial Regression Coefficient

[2] Standardized Partial Regression Coefficient

TABLE 5

Regressions on Growth Outcome Variables - Small Sample
(Regressions Weighted By Accuracy of Birth Date)

Dependent Variable = Zscore Height for Age at Time 2 years
(N = 405)

Variable	B[1]	Beta[2]	p-value	Cumulat. Variance
MORBID1	-.04511	-.13805	.0048	.01914
FAMILIAL	.03446	.10671	.0387	.03420
NUTRIDEX	.00199	.09821	.0482	.04191
ECONDX	.03399	.08856	.0916	.04885

Dependent Variable = Zscore Weight for Age at Time 2 years
(N = 414)

Variable	B	Beta	p-value	Cumulat. Variance
MORBID1	-.06347	-.19853	.0000	.03595
FAMILIAL	.03805	.11961	.0130	.04943
NUTRIDEX	.00171	.08582	.0748	.05447

Dependent Variable = Zscore Weight for Height at Time 2 years
(N = 405)

Variable	B	Beta	p-value	Cumulat. Variance
FEEDINDX	-.05797	-.08280	.0897	.00554
MORBID1	-.07063	-.18423	.0002	.03764
FAMILIAL	.03641	.09610	.0490	.04453

[1] Unstandardized Partial Regression Coefficient

[2] Standardized Partial Regression Coefficient

Table 4 displays the results of the regressions on the growth indicators Z-score of Height-for-Age (ZHA4), Z-score of Weight-for-Age (ZWA4), and Z-score of Weight-for-Height (ZWH4). Table 5 does the same for the smaller sample containing the Food Availability index.

Height-for-Age: Prior Morbidity had the most effect (negative) followed by Familial Circumstances (positive). The variance explained was only 2.9%. When the small sample was analyzed, Prior Morbidity retained its prominence as a negative influence; Familial Circumstances showed a positive effect while Food Availability and the Economic Status followed in order of importance, each with a positive influence. The variance explained increased to 4.8%. None of the intermediate variables of Feeding Practices and Current Morbidity entered the equation.

Weight-for-Age: Prior Morbidity had the most influence, again a negative one, followed by the positive effect of Economic Status. The variance explained is low, at 4%. With the smaller sample, the Familial Circumstances substitutes for the Economic Status while the Food Availability enters the equation, in third place, with a positive influence. The explained variance improved to 5.4%. The two intermediate variables were again not significant enough to enter the equation.

Weight-for-Height: Prior Morbidity had a negative influence, followed by a positive influence of Economic Status and a smaller, negative effect of Prevention. The explained variance was 3.4%. Using the small sample, Prior Morbidity retained the dominating influence (negative) but the Familial Circumstances came second (positive) and Feeding Practices third (negative). The explained variance improved to 4.4%. As mentioned in the methods section, the distribution of the U.S. Percentile for Weight-for-Height (WHPCT) being closer to normal than the other indicators, it was decided to try it in the regression to see if it gave similar results to the Z-score of Weight-for-Height. Indeed, the regression gave very similar results when the large sample was used. This adds robustness to the conclusions. Prevention was substituted for Familial Circumstances as the second influential factor when the small sample was used. This is more consistent with the results for the large sample. The results of regression on WHPCT may be found at end of Appendix F.

Given the above findings, one had to consider amendments to the initial growth models. The revised models are depicted in figures 5 and 6, together with their path coefficients which will be discussed next.

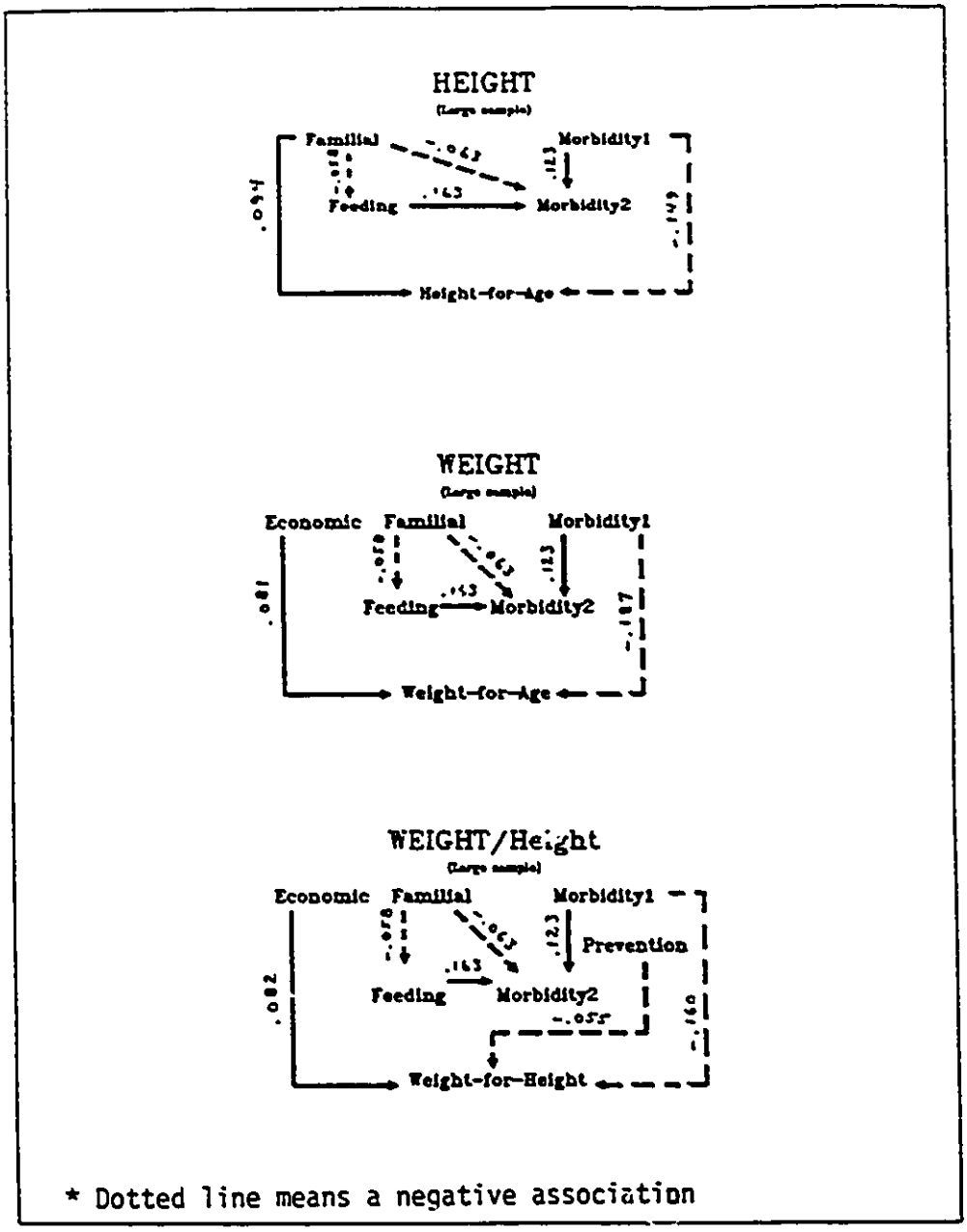


Figure 5 - Derived Models of Predictors of Growth - Large Sample

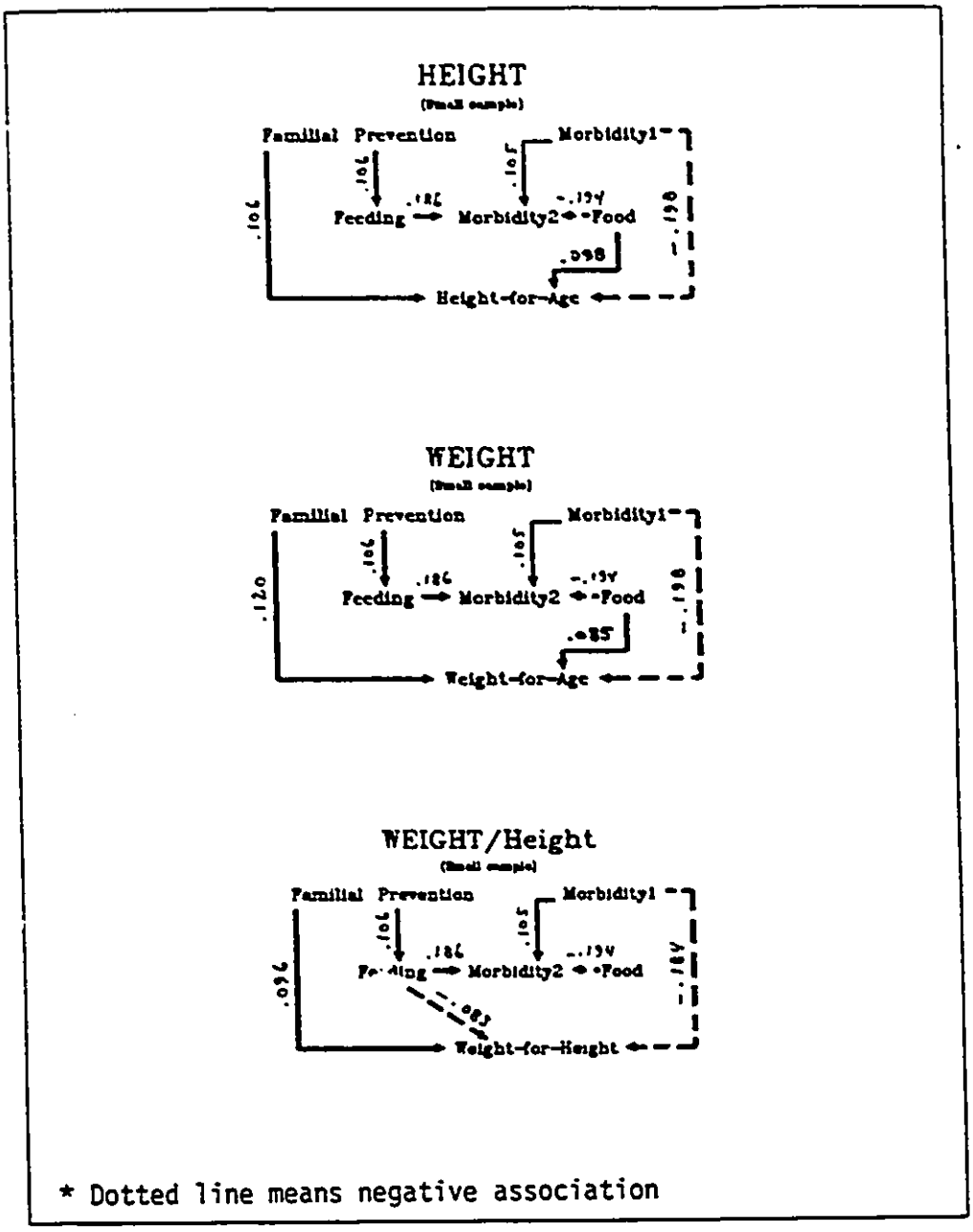


Figure 6 - Derived Models of Predictors of Growth - Small Sample

4.4 Path Analysis:

A path coefficient is the fraction of the standard deviation of the dependent variable for which a designated factor is directly responsible. Path coefficients were calculated for the derived models, using the beta-weights as noted on figures 5 and 6. In almost all cases, the path coefficient is very close if not identical to the Pearson zero-order correlation coefficient. This finding would tend to validate the derived models. However, as can be seen, the indirect components of the calculated path coefficients contribute minimally to the total value and their inclusion into or exclusion from the models does not cause any significant change in the final path coefficients. One could thus simplify the final models without the loss of important information as depicted in figure 7. Possibly, the only exception to this rule is the effect of Familial Circumstances on Current Morbidity through feeding practice which has an indirect path coefficient value of -.011 when done on the whole (large) sample. Although the elimination of some of the postulated but not substantiated effects or paths of action was carried out, it is quite possible that some of these hypothezised effects or paths have been masked due to imputation of missing values by their general mean or their modal values.

In the interest of brevity, in table 6 numbers were given to the variables. The corresponding values are as follows:

Familial Circumstances = No.1
Prior Morbidity = No.2
Economic Status = No.3
Prevention = No.4
Feeding Practices = No.5
Current Morbidity = No.6
Height-for-Age = No.7
Weight-for-Age = No.8
Weight-for-Height = No.9
Food Availability = No.10

Correlations coefficients (r) and Path coefficients (P) are represented by conventional notations discussed in Kerlinger & Pedhazur (24); for paths, the first number indicates the dependent variable, the second, the independent variable.

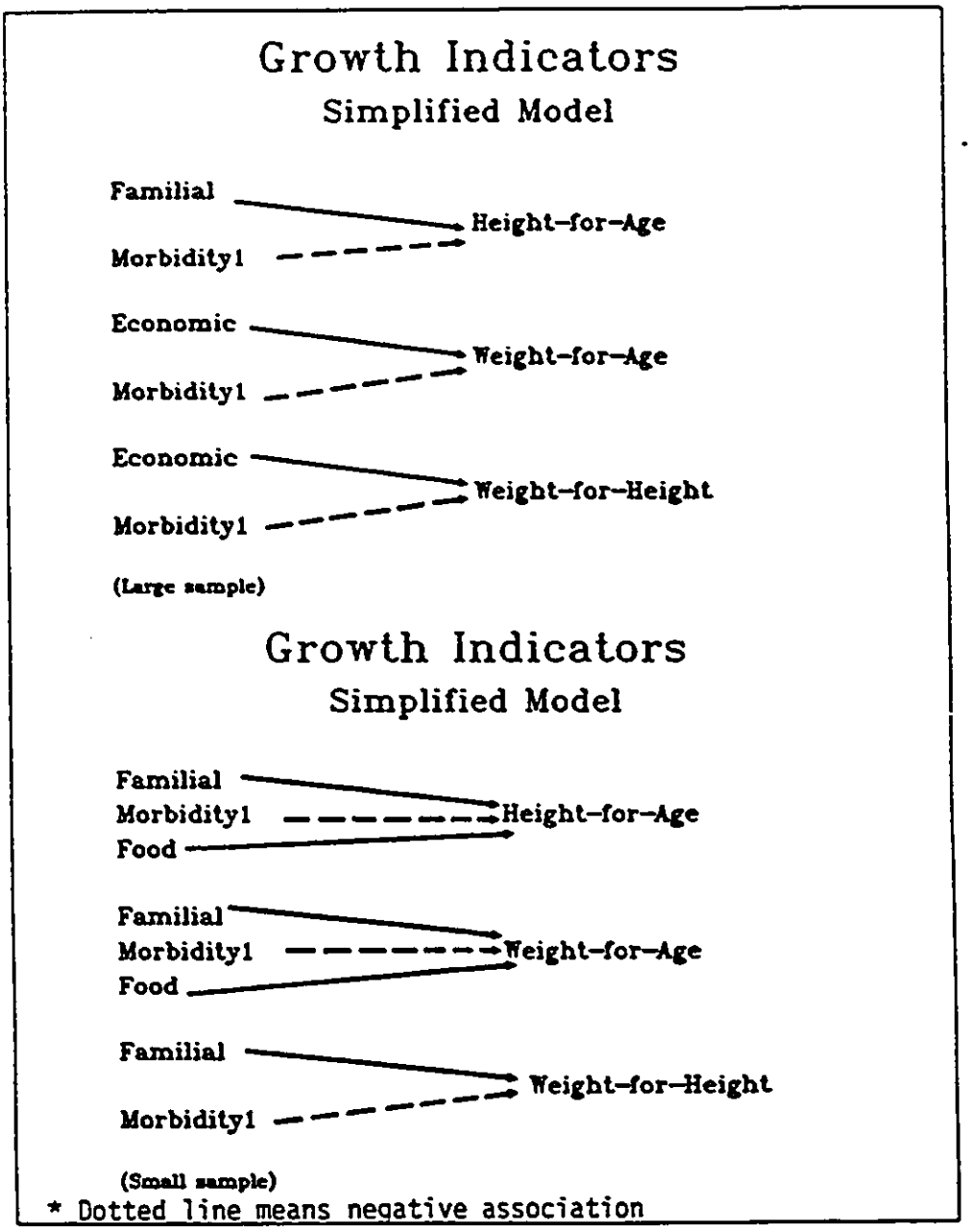


Figure 7 - Simplified Models of Predictors of Growth

TABLE 6
Path Coefficients

LARGE SAMPLE:
r1,5 = P5,1 = -.058
r1,6 = P6,1 + (P5,1 * P6,5) + (P7,1 * P7,2 * P6,2) = (-.063) + (-.011) = -.074
r2,6 = P6,2 + (P7,2 * P7,1 * P6,1) = (.123) - (-.001) = .124
r1,7 = P7,1 + (P6,1 * P6,2 * P7,2) + (P5,1 * P6,5 * P6,2 * P7,2) = (.094) + (.0013) = .095
r2,7 = P7,2 + (P6,2 * P6,1 * P7,1) + (P6,2 * P6,5 * P5,1 * P7,1) = (-.149) - (.0008) = -.150
r2,8 = P8,2 = -.187
r3,8 = P8,3 = .081
r2,9 = P9,2 = -.160
r3,9 = P9,3 = .082
r4,9 = P9,4 = -.055

SMALL SAMPLE:
r4,5 = P5,4 = .106
r2,6 = P6,2 = .105
r4,6 = (P5,4 * P6,5) = .0197
r10,6 = P6,10 = -.194
r1,7 = P7,1 = .106
r2,7 = P7,2 = P7,2 + (P6,2 * P6,10 * P7,10) = (-.138) + (-.002) = -.140
r4,7 = (P5,4 * P6,5 * P6,10 * P7,10) = -.0004
r10,7 = P7,10 + (P6,10 * P6,2 * P7,2) = (.098) + (.003) = .101
r1,8 = P8,1 = .120
r2,8 = P8,2 + (P6,2 * P6,10 * P8,10) = (-.198) + (-.002) = -.20
r4,8 = (P5,4 * P6,5 * P6,10 * P8,10) = -.0003
r10,8 = P8,10 = .086
r1,9 = P9,1 = .096
r2,9 = P9,2 + (P6,2 * P6,5 * P9,5) = (-.184) + (-.002) = -.186
r4,9 = (P5,4 * P9,5) = -.0088
r10,9 = P9,10 = -.083

4.5 Malnourished Sub-Groups:

Children defined as Stunted, Underweight and Wasted were analyzed separately as previously stated (p. 58).

Zero-order Correlation Coefficients on Growth Indicators among malnourished children:

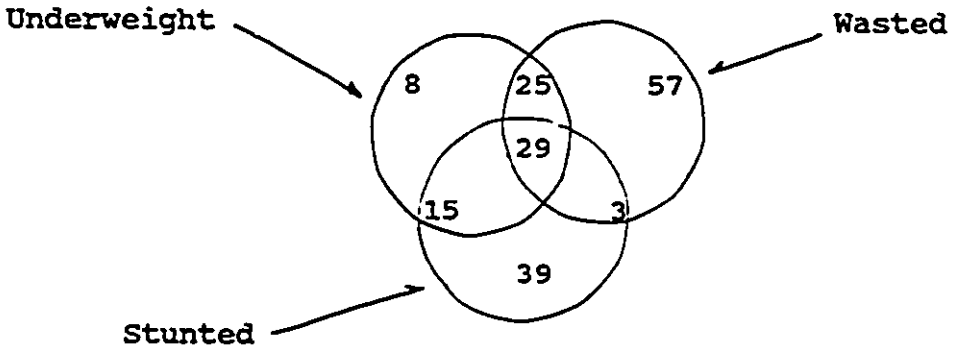
Independent	Dependent Variables		
	ZHA4 (N = 86) (Stunted)	ZWA4 (N = 77) (Underweight)	ZWH4 (N = 114) (Wasted)
FEEDING INDEX	.295	.195	.181
CURRENT MORBIDITY	.166	.142	.213
PRIOR MORBIDITY	-.115	-.215	-.150
FAMILIAL CIRC.	.096	.068	-.004
ECONOMIC INDEX	.077	.157	.056
PREVENTION INDEX	.024	.157	-.002
FOOD AVIALABILITY*	-.028	-.020	-.202

* Not used in the regression analysis.

As can be seen from the preceding zero-order correlations on the growth outcome variables, Feeding Practices was associated positively, albeit weakly in absolute terms, with the growth indicator in all of the subgroups. Prior Morbidity had an association with Weight-for-Age among the Underweight sub-group while Current Morbidity was correlated with Weight-for-Height among the Wasted sub-group. These correlations differ from those obtained from the general sample of children. In fact they show a relatively stronger but reverse relationship of intermediate variables (Feedindx and Morbid2) with the growth indicators. This suggests a threshold of weight-for-age and height-for-age below which

some factors such as Feeding Practices and Current Morbidity become influential, with an effect on that malnourished group opposite to their milder influence on the general population of Zambian children.

Looking at the simple correlations among the predictors themselves (see Appendix G), one realizes that the coefficient values are not consistent between the three sub-groups. The only consistency of association lies between Feeding Practices and Current Morbidity, between Familial Circumstances and Prevention, between Familial Circumstances and Economic Status; all being positively correlated. This general lack of consistency could be, in part, due to subgroups made up of different individuals. Some of the children belonged to more than one of these subgroups but many fitted only in one as shown by the following Venn diagram:



The variance explained by regression analyses varied around 5% to 12%; this is higher than with the general sample. (The actual variance and regression coefficients can be found in Appendix G).

Predicting feeding practice: No constancy of association came to light when the regression on the three subsamples were examined so no conclusion can be drawn.

Predicting current morbidity: Feeding practice entered the equation on two occasions with a low p-value (.0086 and .0067) and positive beta-weights thus suggesting that feeding practice may have a positive effect on current morbidity; this might mean that the more foods are introduced early before weaning, the higher is the likelihood of Current Morbidity in these malnourished sub-groups of children. The introduction of other foods may bring with it the danger of contamination with enteric pathogens and this would increase the likelihood of diarrhea thus increasing the score of the Current Morbidity index. Whether bottle feeding is the villain here instead of the introduction of all foods cannot be teased out because of the way the Feeding Practices index was compounded. This may have been remedied by creating two separate feeding indexes, one dealing with bottle feeding, the other with introduction of solid foods. Unfortunately, proceeding thus

would have created two new indexes with less components than the original feeding index and consequently, further from normal in their distribution since the range of possible values for these new indexes would have been more limited (thus less continuous).

Stunting: The feeding practice (p-value .0058) had a positive correlation with the growth of this sub-group of children and explained 7.6% of the variance in Height-for Age. One should remember that feeding practice is correlated with Current Morbidity (r = .282) which might also be an important predictor even though it did not enter the final equation. The difference in height among this stunted group is thus due, at least in part, to feeding.

Underweight: Prior Morbidity was the only significant factor with a p-value of .0874 and a negative correlation. It explained 5% of the variance of Weight-for-Age in this sub-group. Feeding practice almost entered the equation (p = .131) and should be considered a meaningful positive factor in view of its zero-order correlation (r = .195) with the growth outcome. This suggests that thinness is due, to some degree, to feeding and morbidity.

Wasting: Prior Morbidity (p = .038) was a negative predictor and Current Morbidity (p = .008) a positive predictor of

Weight-for-Height, explaining 6.5% of its variance. One should remember that feeding practice has a zero-correlation coefficient of .181 with the growth outcome and probably did not enter the equation because of its correlation with Current Morbidity (r = .228) and Prior Morbidity (r = -.115), although it might also be an important positive predictor. As might be expected, wasting is partly the result of prior morbidity.

Comments:

Published research, mostly done on malnourished cases, stresses the significant influence of feeding practices, morbidity and socio-economic status on the growth of prepubertal children. Analysis of the general sample in this study did not support the importance of Feeding Practices on growth indicators but analysis of the malnourished sub-sample did, suggesting a treshold effect. Feeding practice was common to both the general sample and the malnourished sub-groups as a positive predictor of Current Morbidity. Prior Morbidity was the only common predictor of the three growth outcome indicators (Height-for-Age, Weight-for-Age, Weight-for-Height) for the general sample as well as among the three malnourished sub-groups. Feeding practice has a weak direct influence on the three growth indicators but also exerts an influence on the latter

outcomes through Current Morbidity which it increases (figure 8).

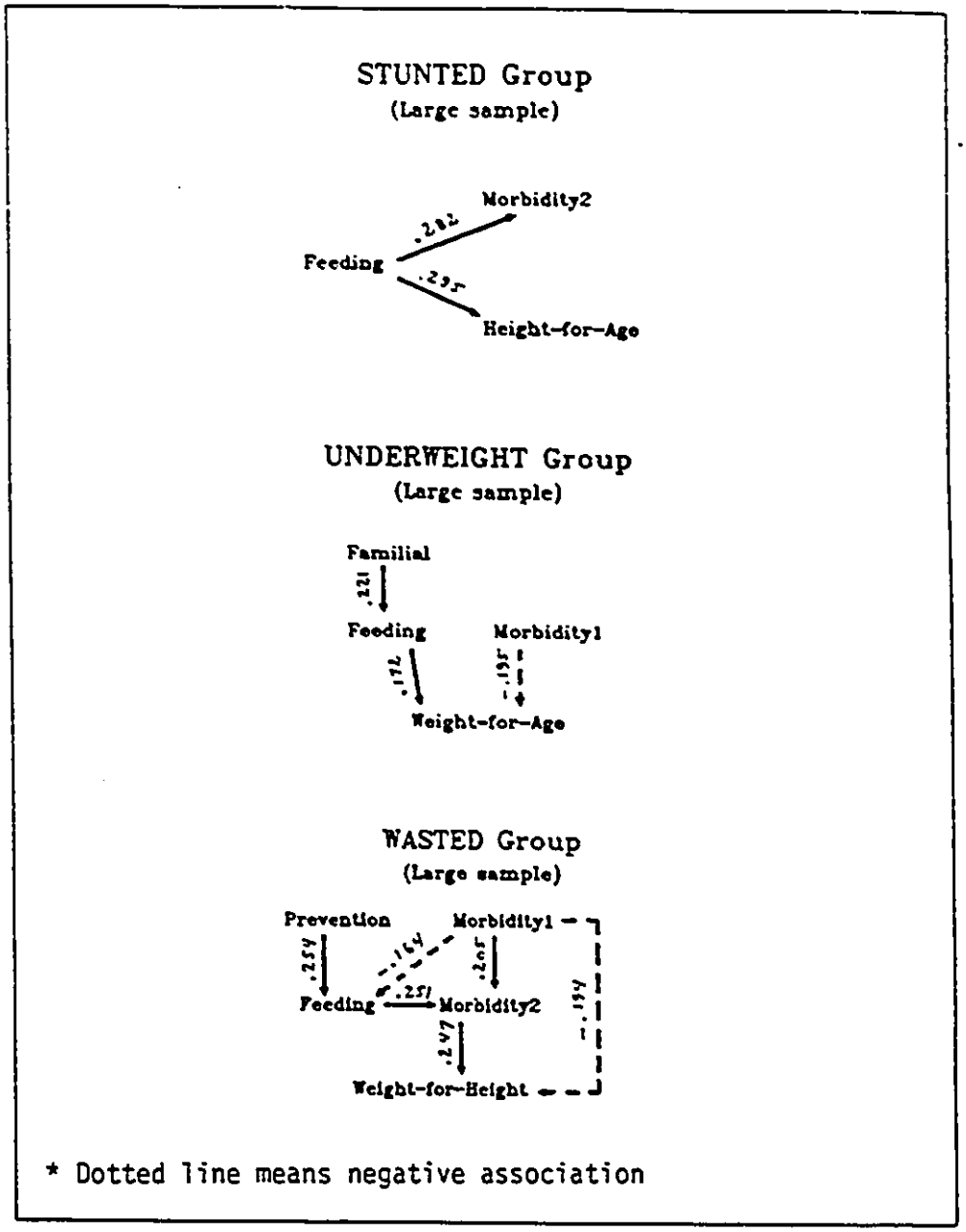


Figure 8 - Derived Models of Predictors of Growth - Malnourished Groups

4.6 Age Specific Sub-Groups:

As specified in the methods section, the general sample was divided into children 6 months of age and younger at the start of the study (28% of the sample), and those more than 6 months old. Regressions were run again, using the same criteria and variables as for the previous regression analyses on the general sample (see Appendix H for tables).

Children less than 6 months:

Predicting feeding practice: No predictor entered for the large sample (N = 248). On the small sample (N = 119), Prevention was the only predictor (positive) retained and not with a high statistical significance. The explained variance was less than 1.9%.

Predicting current morbidity: Prevention was a weak positive predictor on the larger sample (R^2 <1%). Only Food Availability was significant for the smaller sample (R^2 = 4%). Both factors affected Current Morbidity negatively.

Height-for-Age: Prior Morbidity was the most important predictor (negative) on the large sample, followed by Prevention (positive). Prior Morbidity was the only factor retained on the smaller sample. The explained variance was less than 3% in both equations.

Weight-for-Age: Prior Morbidity was the only significant predictor (negative) for both samples with the explained variance at less than 3.6%.

Weight-for-Height: Prior Morbidity was the most important predictor (negative), followed by Economic Status (positive). The factors were identical for both samples with R^2 at 4% for the larger group and 6% for the smaller one.

Children aged 6 months to 5 years:

Predicting feeding practice: Prior Morbidity was the only factor of importance on both samples (N of 663 and 293 respectively), presenting a positive association perhaps due to an earlier introduction of other foods or of bottle feeding when the child became ill. The explained variance did not exceed 1% in both samples.

Predicting current morbidity: Prior Morbidity was the most important (positive) predictor on the larger sample, followed by feeding practice (positive) and by Economic Status (negative). On the smaller sample, feeding practice appeared as the most important, followed by Prior Morbidity and Food Availability (negatively correlated). R^2 was 5% for the larger group and 12% for the smaller one.

Height-for-Age: Prior Morbidity was the most important (negative) predictor on the large sample, followed by Familial Circumstances and by Economic Status, both positive (R^2 = 3.5%). On the smaller sample, the Economic Status came first, followed by Prior Morbidity and by Familial Circumstances (R^2 = 4.5%).

Weight-for-Age: Prior Morbidity was strongest, having a detrimental influence, followed by Familial Circumstances (positive) on both samples. Food Availability was a third place, positive factor, on the smaller sample. The variance explained did not exceed 4.5% for the larger group and 7.6% for the smaller one.

Weight-for-Height: Prior Morbidity was the most important (negative) predictor, followed by Economic Status (positive). The factors were identical for both samples (R^2 was 3.2% for large group and 5.3% for smaller one).

Comments:

Prevention may have some positive effect on Feeding Practices and Current Morbidity in the younger group but not in the older one. The explanation for the effect of Prevention on Feeding Practices is not obvious. It is possible that health teaching at the clinics encouraged early introduction of supplementary foods. As for its

negative effect on Current Morbidity, the good advice given at the clinic and vaccinations may have prevented diarrhea and other diseases. Prior Morbidity appears to have been the major determinant of Feeding Practices for the older group of children suggesting that illness encouraged the mother to introduce other foods into the child's diet, perhaps due to the child's reduced ability to breastfeed or to loss of energy while ill. Prior Morbidity also predicted Current Morbidity in the older group but not in the younger group. Looking at the growth indicators themselves, Prior Morbidity appears to be the major predictor (negative) overall in all age groups with Economic Status having a positive but lesser influence on weight-for-height in the younger group; Familial Circumstances do the same for the older group. Better Familial Circumstances appeared to improve, as well, the growth indicators height-for-age and weight-for-age in the older age group but not in the younger one. One should remember that Economic Status was correlated with Familial Circumstances (.20 to .40), so that each may still be of importance when the other one is present although it may have been excluded by the step-wise method of variable selection. The two age groups differ in their predictors, except for Prior Morbidity which affects both. The older group is quite similar to the general sample.

Table 7 (p. 94) gives a summary of all results on growth.

a) PREDICTING FEEDING PRACTICES: There was no consistent predictor of Feeding Practices across all the various groups studied. Even between the stunted, the underweight, and the wasted there were differences. An overlap of various foodstuff with breast feeding was not directly influenced by their availability in the household.

b) PREDICTING CURRENT MORBIDITY: In general, Feeding Practices and Prior Morbidity correlated positively with Current Morbidity, with the exception of the group of children less than 6 months old and of the underweight sub-group. Familial Circumstances and Food Availability correlated negatively with Current Morbidity in all groups except the malnourished. In these malnourished subgroups, the analyses suggests that increasing the availability of high protein food stuff might decrease this type of morbidity (mainly diarrhea).

c) PREDICTING GROWTH INDICATORS: Prior Morbidity was consistently a negative predictor except for the height in the stunted group, where Feeding Practices were the only retained predictive factor. In fact, the overlap of various foods with breastfeeding seemed to be an important favourable predictor for the three indicators of growth in the malnourished subgroups but not in the other groups. Familial Circumstances and/or Economic Status were also important positive predictors of growth except among

children aged less than 6 months and in the wasted sub-group whose growth did not appear to be as much influenced by these two factors. Food Availability was a positive predictor of weight-for-age for children more than 6 months of age and in the general sample in which it also predicted height-for-age. As a rule, the percentage of variance explained by the chosen predictors was low but improved (reached 12%) when subgroups of the general sample were analyzed.

TABLE 7

Summary of Predictive Factors - Growth Model

	Subgroups			
Outcome	General Sample	<6 months	>6 months	Malnourished
Feeding	Familial -	—	Morbid1 +	(Familial +)
				(Prevention +)
				(Morbid1 -)
Current	Feeding +	Prevention -	Morbid1 +	(Feeding +)
Morbidity	Morbid1 +	Food Av. -	Feeding +	(Morbid1 -)
	Familial -		Economic -	
	Food Av. -		Food Av. -	
Height	Morbid1 -	Morbid1 -	Morbid1 -	Feeding +
	Familial +	Prevention. +	Familial +	
	Food Av. +		Economic +	
Weight	Morbid1 -	Morbid1 -	Morbid1 -	Morbid1 -
	Economic +		Familial -	Feeding +
	Food Av. +		Food Av. +	
Wt/Ht	Morbid1 -	Morbid1 -	Morbid1 -	Morbid2 +
	Economic +	Economic +	Familial +	Morbid1 -
	Prevention -			?Feeding +

+ = Positive predictor of growth

- = Negative predictor of growth

— = No predictor achieved statistical significance

? = doubtful statistical significance

Note: From the smaller sample results, only the Food Availability factor is included when it was significant enough to enter the regression equation. The rank of the factors listed within each outcome denotes the relative size (importance) of the partial regression coefficients of the predictors, except in the case of the Food Availability factor.

Comments:

Hypothesis #1 on predictors of feeding practices was not confirmed since the results were quite inconsistent across the various groups studied, perhaps for the reasons just cited. It may be that the components of the Feeding Practices index are influenced differently in various sub-populations. Only in some of the malnourished subgroups did better Familial Circumstances have a positive effect on feeding practices, and did a higher degree of prevention have a positive effect on feeding practices, as suggested by Gerein (15). Hypothesis #2 was verified, mainly in the general sample and in the children older than 6 months of age. Prior Morbidity had a direct positive effect on Current Morbidity while Familial Circumstances, Economic Status, and Food Availability had negative effects, as suggested they would. Prior Morbidity was consistently the most influential predictor of low values of all growth indicators while Familial Circumstances alternated with Economic Status as the main favourable predictors of growth indicators, usually with about half the influence of Prior Morbidity. Food Availability had a moderate influence on height-for-age and on weight-for-age but not on weight-for-height. In the general sample, contrary to expectations and as demonstrated by path analysis, most of the effects of Prior Morbidity, Familial Circumstances, Economic Status, Prevention, and Food Availability on the growth indicators

were direct and did not operate through Feeding Practices or Current Morbidity. This negates hypothesis #3, except for the malnourished sub-groups. This finding seems to disagree with Bairagi's assertion (3) that most factors influence growth through the intermediate factors of morbidity and feeding practices, but it brings credibility to the impression I had that most of the previous findings regarding predictors of growth apply mainly to the malnourished. The way my Feeding Practices index was constructed may also have contributed to this disagreement. It may have been an error on my part to group the introduction of bottle feeding, found to be a direct (positive) predictor of malnutrition by various authors (3,24,34), and the introduction of solid foods, found by many researchers (2,7) to be a negative predictor of malnutrition, in the same index, thus reducing the value of the index as a measure of exposure to either influence.

Chapter 5

5.0 Mortality:

5.1 Discriminant Analysis:

There were 155 deaths among the children. Unfortunately, seventeen deaths were excluded from the analyses because of missing data for some of the predictor variables, leaving 138 usable cases for the large sample and 58 for the small one. Classification in discriminant analysis is very sensitive to the initial probabilities specified for each category of the outcome variable. When prior probability of death for the model was defined as similar to its proportion within the sample (about 11%), the accuracy of classification of the dead children was null (i.e. 0%) while the one for the overall sample was 89%; in other words, the function was very good in predicting correctly as "alive" the children who in fact would survive but was useless in predicting the outcome of most interest, i.e. death of the child. This situation sometimes occurs when the number of cases in each group involved in the discrimination analysis is very disproportionate (almost 10:1 here). The desired result, in the current study, was not to minimize the overall misclassification rate but to identify most cases susceptible of dying. In an effort to improve the prediction for a death outcome, the criterion of prior

probability of death was raised to 50%; with this tactic, predictive ability of the resulting equation for death of the child was improved to 69% (but down from 89% to 55.1% for survival), suggesting a better predictive value for death when the prevalence of death before age 5 reaches 50%; however, this is rather extreme and highly unlikely to ever happen. The accuracy of prediction was further enhanced by running the analysis using the dead children and an equal number randomly chosen from the survivors. If the larger "donor" sample of the live children had been homogeneous, the performance on classification should have been essentially the same as the one obtained by setting a prior probability of death of 50%. The further improvement in classification suggests poor consistency within the study sample, perhaps caused by hidden biases introduced by replacing missing values in the data set, as suggested earlier in this paper. Using the derived equation, almost 75% of the dying children were correctly identified. However, only 60% of the survivors were correctly classified, leaving 40% of false positive (i.e. falsely predicted to die) and an overall predictive accuracy of 68%. Since a change in specification of prior probability of events affects only the classification of cases but does not alter the rest of the output of the discriminant analysis (e.g. coefficients), only the results of this latter analysis will be presented.

As shown in table 8 for the large sample, Prior Morbidity appears to be the most important predictor of likelihood of death among the children but, surprisingly its influence is negative. This is followed by the Economic Status, also negatively correlated. Degree of Prevention was in third place, with a positive correlation and Familial Circumstances in fourth place, negatively correlated. The other variable, gender, did not enter at the chosen p-value of .10 and may be less important than was postulated. Gender was excluded as well in the analysis of the small sample containing the Food Availability variable (table 9). In that sample, Prior Morbidity remained the most important predictor, with Economic status in second place and Food Availability taking the third place with a positive correlation. Prevention and Familial Circumstances did not attain statistical significance in that second analysis. The degree of relatively high simple correlation between the Economic Index and the Familial Index, and similarly between the Morbidity Index and the Preventive one may have influenced their respective order of inclusion during the stepwise analysis of the larger sample and may well have contributed to their exclusion in the final function for the smaller sample. In the latter situation, one should then not necessarily conclude that familial or preventive factors had no significant role.

During the analysis, statistical significance of the discriminant function was tested by a Chi-square on Wilk's Lambda (ratio of within-groups sum of squares to the total sum of squares). Wilk's Lambda represents the proportion of the total variance explained by the differences among the groups. Eigen values (ratio of between-groups to within-groups sum of squares) were also examined to check on the discriminant ability (ideally as close to 1.0 as possible) of the discriminant function. As previously mentioned in the methods, section 3.5, Box's M test was used to verify that the assumptions of normality of the groups' multivariate distributions and of equality of their covariance matrices had not been violated. The following table lists the three values of Lambda, Eigen, and Box's M for the larger and smaller samples:

	Eigen	Lambda	Sig	Box's M	Sig
Large Sample	.1529	.8673	.000	202.68	.000
Small Sample	.1578	.8637	.000	17.58	.009

The Eigen values are not very high, forecasting a poor discriminating ability, as already noted. Lambda remains fairly high despite being statistically significant (due to high N). Box's M points to violation of assumptions which suggests that caution should be exercized when interpreting the results of the discriminant analyses, as the equation would not be optimal even though this type of analysis may

still perform reasonably well under less than ideal conditions.

TABLE 8

Discriminant analysis - Mortality-Large sample

N = (136 ALIVE + 138 DEAD) = 274

Pooled Within-Groups Correlation Matrix

	SEX	MORBID1	PREVINDX	FAMILIAL
MORBID1	.12466			
PREVINDX	.08545	.18880		
FAMILIAL	-.02468	-.06355	.04557	
ECONDX	.07218	-.00432	.04080	.26651

Step	Action Entered	Vars In	Wilks' Lambda	Sig.	Std. Regr. Wt
1	MORBID1	1	.90578	.0001	.86142
2	ECONDX	2	.88362	.0001	.50572
3	PREVINDX	3	.87453	.0001	-.27002
4	FAMILIAL	4	.86735	.0001	-.25906

Classification Results:

Actual Group		No. of Cases	Predicted Group Membership: Alive	Dead
ALIVE	0	136	82 60.3%	54 39.7%
DEAD	1	138	34 24.6%	104 75.4%

Percent of "grouped" cases correctly classified: 68%

TABLE 9

Discriminant analysis - Mortality-Small sample

N = (59 ALIVE + 58 DEAD) = 117

Pooled Within-Groups Correlation Matrix

	SEX	MORBID1	PREVINDX	FAMILIAL	ECONDX
MORBID1	.11552				
PREVINDX	.15484	.22360			
FAMILIAL	-.01270	-.04973	.06238		
ECONDX	.05906	-.03933	.01815	.20123	
NUTRIDEX	.08616	.18920	-.06181	.09704	.02847

Step	Action Entered	Vars In	Wilks' Lambda	Sig.	Std. Regr. Wt
1	MORBID1	1	.92610	.0030	.65311
2	ECONDX	2	.88436	.0009	.55358
3	NUTRIDEX	3	.86371	.0008	.42178

Classification Results:

Actual Group		No. of Cases	Predicted Group Membership: Alive	Dead
ALIVE	0	59	32 54.2%	27 45.8%
DEAD	1	58	18 31.0%	40 69.0%

Percent of "grouped" cases correctly classified: 62%

5.2 Tests of Association between the outcome "Death of the Child" (all 155 occurrences) and 21 original variables:

Results of univariate analysis, obtained by way of contingency table calculations and shown in table 10, confirmed the association of some of the components of Prior Morbidity with the likelihood of death that had been found by previous research. The strong association of Prevention with increased mortality was confirmed by the crosstable analysis which found an odds ratio of 10 for a dichotomized version of the Prevention index, casting doubt as to what it really measured, prevention or a proxy for sickness. The other variables examined were all statistically non-significant at the 5% level, using the chosen cut-off of $p < .002$ as justified in the methods section.

Other highly statistically significant associations with death of the child were found. Living in the squatter town had more than twice the risk of living in the better suburb. Prematurity increased four-fold the risk of death and being a twin, elevated it by five-fold. A more specific measure of morbidity such as measles was found to be predictive of death in the child. However, enteritis was protective; again difficult to explain. In the latter case, some unaccounted for beneficial cultural factor (attitude and/or behaviour) may have been operative during diarrheal episodes.

TABLE 10

Measures of Associations on Likelihood of Death

Variable (Case/Reference)	p-value	O.R.	C.I.(95%)
Kind of Town (Worse/Better)	<.00001	2.24	1.50-3.20
Prematurity (Prem./Term)	<.00001	4.20	2.60-6.60
Pregnancy (Twin/Singleton)	<.00001	5.20	2.80-9.60
Prevention (High/Low)	<.00001	10.00	5.30-20.0
Measles (Yes/No)	.00004	2.40	1.50-3.80
Enteritis (Yes/No)	.0016	0.36	0.18-0.70
Prior Morbidity (High/Low)	.0056	0.54	0.30-0.83
Treat. witchdoct. (Yes/No)	.0668	1.81	0.96-3.36
Food Availability (High/Low)	.0840	0.50	0.23-1.08
Death of sibs (Some/None)	.1178	1.33	0.93-1.90
Mother's Allowance (High/Low)	.1268	0.75	0.52-1.07
Mother's Education (High/Low)	.1624	0.64	0.34-1.17
Economic Status (High/Low)	.2451	0.73	0.44-1.21
Past Diag. of Malnut. (Yes/No)	.3073	1.45	0.74-2.78
Gender (Male/Female)	.4327	0.86	0.60-1.22
Age mother (Average/Extremes)	.7030	1.09	0.74-1.61
Father's Income (High/Low)	.7244	0.93	0.65-1.32
Familial Circ. (High/Low)	.7354	0.90	0.55-1.45
Place of birth (Hosp./Home)	.8193	0.94	0.66-1.34
Diff. lactation (Yes/No)	.9025	0.99	0.59-1.66
Parity (High/Low)	.9424	0.99	0.68-1.45

Comments:

Hypothesis #4 on predictors of mortality was confirmed somewhat except for gender which did not appear to be as important as anticipated. A high Economic Status and high Food Availability had a favourable effect on survival. It should be noted, however, that Food Availability was not statistically significant on univariate analysis. The negative correlation of Prior Morbidity with the likelihood of death and the positive association of Prevention found in the discriminant analysis are difficult to explain. The negative association of Familial circumstances gives similar difficulty. The fact that Prior Morbidity and Prevention are intercorrelated ($r = .19$), and similarly between Economic status and Familial Circumstances ($r = .27$) may be partly responsible for this 'reversal' of signs. In discriminant analysis, signs (direction of action) of correlation coefficients may be difficult to interpret if predictors are intercorrelated. The apparent protective action of Prior Morbidity found by discriminant analysis is confirmed however by the univariate analysis, although not quite statistically significant by the multiple comparison criterion (p-value must be $<.002$). The fact that some of its components such as measles and prematurity show an effect in the opposite direction suggests a lack of homogeneity of effect among the index's components. This is plausible in view of the Cronbach alpha coefficient of less

than 0.45 for this index. Both Crean (8) and Excler (14) had found that morbidity and low immunization rates increased the likelihood of death. The positive association of Prevention may be explained by the mothers bringing their children to the under-five clinic more readily when they were ill than when they were in apparent good health, especially after the age of regular immunizations, as suggested by the regression analyses on age-specific groups. This suggests a lack of validity of the Prevention index in measuring "real" prevention in some subgroups of children. The negative association of Familial Circumstances with the likelihood of survival found by discriminant analysis was not supported by the univariate analysis where it did not reach statistical significance. In view of the above results, I was forced to reappraise the original model for predictors of death. Because of the difficulty in assessing the direction of action of some of the predictors, I decided not to specify the direction of effect but only to list the retained predictors especially if one takes into account the fact that assumptions of ideal conditions for a discriminant analysis does not seem justified. The final model is shown in figure 9.

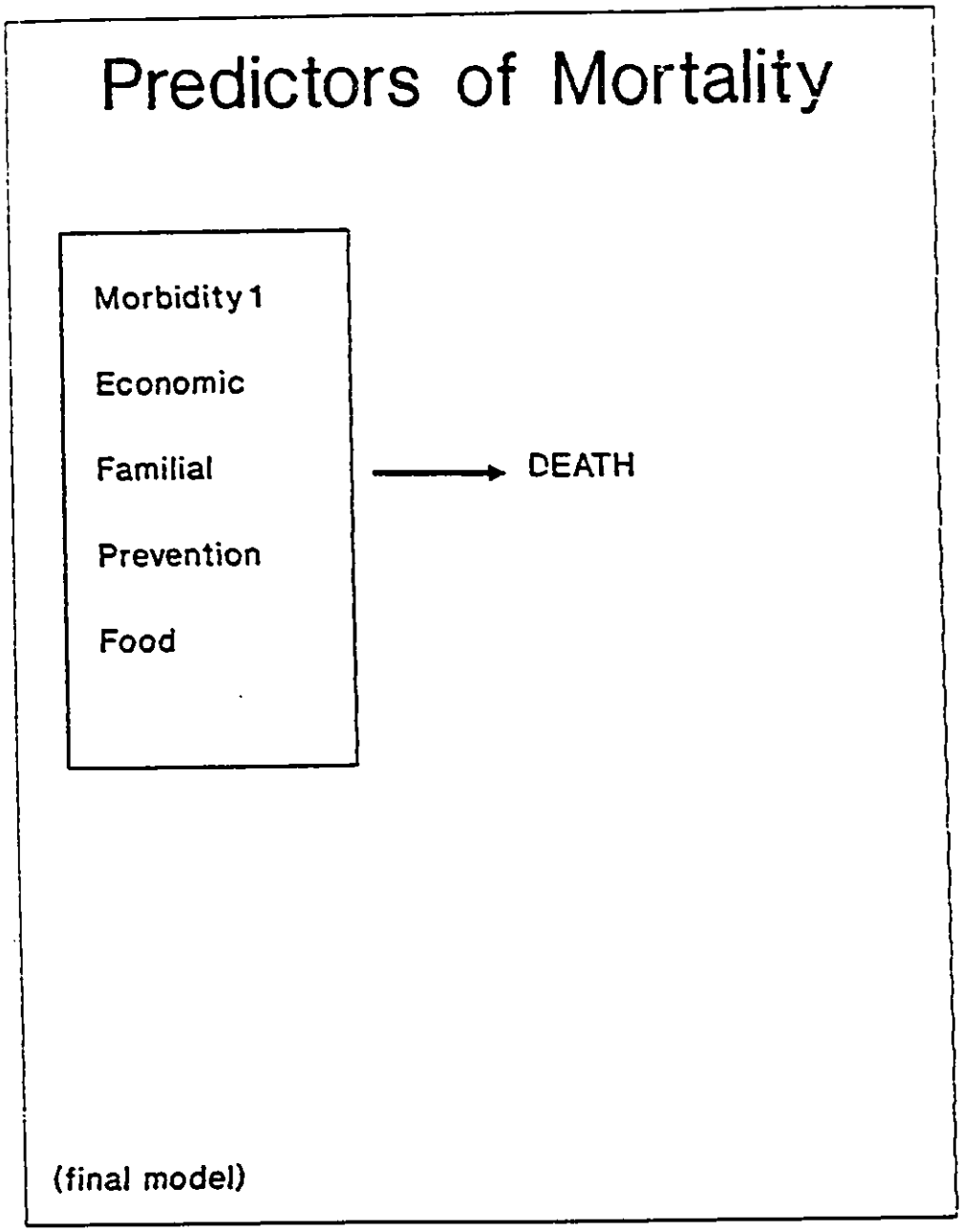


Figure 9 - Derived Model of Predictors of Death

The degree of prediction achieved by the derived discriminant equations is not very impressive and might have improved had one been able to use data on feeding practices and current morbidity. This was not possible however as the two indexes were using data collected longitudinally during the study period and part of this information would have stopped being collected after the death of the child. Nevertheless, when viewed in the light of the results that could be expected by chance, the final derived function is at least an improvement over that of tossing a coin. However, in real life, the proportion of dying children observed would be much less than the 50% used for the equation (it was about 11% in this data set) whereupon the equation would do a worse job of identifying the children at risk of dying. In fact, when the prior probability of 0.10 was chosen, none of the dying children was identified as being at risk. This indicates that, if the degree of prevalence of early mortality existing in this sample (11%) is representative of the population, the factors retained by the discriminant equation will not be useful in identifying children at high risk of premature death in order to target them for corrective interventions.

Chapter 6

6.0 GENERAL DISCUSSION

Comments on results of analyses on specific groups were included immediately after their presentation in order to make it easier for the reader.

6.1 Explained Variance:

Overall, it seems that the variables chosen for the analyses of growth have not managed to account for much of its variance. The models derived explained about 3% of the variance of height-for-age, 4% of weight-for-age, and 3.4% of weight-for-height. This percentage improved when the half-sample with the addition of the Food Availability index was used; the values attained then were 4.9%, 5.4%, and 4.4% respectively. This increased to 7-12% for some analyses done in subgroups.

Although the variance explained by the growth models is low, it is within the range found in other studies of predictor variables of growth. For example, Martorell et al. (31) obtained an explained variance of less than 10% on height-for-age and on weight-for-height; on the weight-for-age indicator, it reached 17%. Bairagi (3) in a similar study had managed to explain only 4.2% of the variance of height-for-age, 6.9% of weight-for-age, and 3.2% of weight-for-

height. Afifi (2) was able to explain only 12% of the variance of length. Why is there such a large unexplained variance?

a) Methodological Issues:

1) The growth indicators chosen were derived from anthropometric measurements taken with a locally made length ruler and a Salter spring balance. Although, the fact that the Z-scores of the growth indicators were close to being normally distributed is reassuring, the calibration or the precision of the instruments may not have been ideal thus causing an increase in variability (random error) and resulting in dilution of the measured effect. This may also have caused misclassification of cases into malnourished categories with a similar consequence: dilution of effect. The inaccuracy of age determination may also have had a similar effect, although an attempt was made to compensate for this by weighing the cases accordingly. 2) The fact that there was an average of two children per family in the sample may have reduced the variance between them in the indexes that were composed of variables common to both children (Economic Status, Familial Circumstances, Food Availability). This may have contributed to a reduced explained variance for these indexes similarly, albeit to a lesser extent, to the effect that imputation of missing value with the overall mean or the modal value has had. 3)

The indexes, derived from the original variables, were not as well distributed as were the growth indicators and may not have been, for some, good summary measures of supposedly related exposures. As previously discussed, when an attempt was made to enhance the intercorrelation of the items of the indexes derived from initially collected variables, Cronbach's alpha measure of correlation exceeded .50 for only two of the indexes (Prevention and Familial Circumstances). The regressions attempted with the 'reduced' indexes did not explain any more of the variance.

One may have tried a comprehensive approach and used as many of the collected variables as possible. However, this is somewhat mindless, and turns the analysis into an exploratory, hypothesis-forming exercise. There is also a limit as to how many of these variables may be included in a multivariate regression analysis. Most statistical texts recommend a ratio of about 10-15 cases to one independent variable. In my study, after exclusion of the cases with missing data and division into subgroups, there were times when only 77 children were available for regression analysis. This meant that a maximum of seven predictors could be used. It made sense then, in order to take advantage of the information provided by many of the collected variables, to regroup them into larger compounded indexes. Had most of these variables been continuous

instead of nominal in nature, an ideal way of compounding them would have been to use the Principal Component Analysis procedure as did Afifi (2). The latter procedure could still have been implemented using as input the indexes themselves to try and develop predictors that would achieve a larger proportion of explained variance. The new set of predictive factors derived from Principal Component Analysis does not always make sense intuitively or conceptually and would have made it difficult to test the stated hypotheses.

During the course of the analysis, because of the small sample size of some of the groups studied, it became doubtful that I always had enough power to stay within an acceptable level of beta-error (type II). I decided to verify this by constructing the following table, using incremental correlation coefficients, and the sample sizes used for the majority of the analyses.

Sample size	Correl. Coeff.	95% Conf. Int.	Power
911	.05	- 0.01478, 0.114775	50.0%
911	.06	- 0.00470, 0.124704	50.0%
911	.07	0.00538, 0.134620	56.0%
911	.08	0.01547, 0.144522	67.0%
911	.09	0.02558, 0.154412	77.0%
911	.10	0.03571, 0.164288	86.0%
911	.13	0.06616, 0.193840	97.0%
911	.15	0.08652, 0.213477	99.5%
911	.18	0.11716, 0.242834	99.9%
911	.20	0.13766, 0.262340	99.99%
405	.05	- 0.04715, 0.147150	50.0%
405	.06	- 0.03704, 0.157043	50.0%
405	.07	- 0.02692, 0.166916	50.0%
405	.08	- 0.01677, 0.176770	50.0%
405	.09	- 0.00660, 0.186604	50.0%
405	.10	0.00358, 0.196419	52.0%
405	.12	0.02400, 0.215991	68.0%
405	.13	0.03425, 0.225747	75.0%
405	.14	0.04451, 0.235484	81.0%
405	.15	0.05479, 0.245202	86.0%
405	.18	0.08586, 0.274238	95.5%
405	.20	0.10650, 0.293497	98.3%

From the table above, and supposing that a minimum power of 80% should be achieved, one can see that for the large sample (N = 911), any correlation with a coefficient less than 0.1 may be subject to an unacceptable risk (>20%) of beta (type II) error, i.e. of rejecting a correlation despite there being one. Although a coefficient of 0.07 was statistically significant at $p = 0.05$, a similar correlation still had a 44% chance of being missed. For the small sample (N = 404), the same would be true for a correlation coefficient below 0.14. Although statistical significance is reached at coefficients beyond 0.1, a similar correlation had a 48% chance of being ignored. For even lower sample sizes such as in the subgroups, the coefficients would thus have to be much higher to be acceptable. This power problem, realized only at the end of my study, casts doubt

on many of the conclusions that can be drawn from the results of the analyses, especially the ones done on smaller sample sizes such as the age-dependent subgroups and the malnourished subgroups and underlines even more the likelihood that important effects, although present in the population under study, may have gone undetected.

Univariate analysis was also considered and indeed was used (see table 9 and Appendix J). The difficulty with this method was the increased probability of getting a positive finding by chance alone when doing multiple comparisons. As well, univariate analysis does not permit, compared to a multivariate technique, to control for the intervening influences of confounders unless provision for this has been made at the study design stage by matching on them. Stratification, another method of controlling for confounders, would not have been practical in this study. Too many multilevel strata would have been required, resulting in small numbers of cases within each cell and making interpretation of results difficult.

As previously discussed, the assumptions needed for an ideal performance of a discriminant analysis did not appear to be met by the sample. Thus, caution should be exercized when assessing the validity of the derived discriminating function.

b) Conceptual Issues:

1) One index derived from longitudinally collected variables, that of Feeding Practices, had components which may have been antagonistic, especially on their influence on Current Morbidity. It would be of interest, in further analyses, to differentiate the influence of bottle feeding from the effects of other foods, perhaps by reversing the sign value (+ --> -) of bottle feeding as a component of the index since, as demonstrated by some authors, bottle feeding may be detrimental to growth and survival whereas an overlap of breast feeding with other types of food may be beneficial. However, generalizations such as "bottle feeding was bad" may be too simple. Bottle feeding may introduce infectious elements which result in enteritis or other diseases. Conversely, illness in a young child may weaken his ability to suck at the breast and have necessitated the introduction of bottle feeding which requires less sucking effort. It is not possible, with the data available, to decide which came first, the illness or the bottle feeding.

2) The weights given to many of the various components of the indexes, although based on the best available information, remained the same whether I analysed their predictive ability on height-for-age, weight-for-age, or

weight-for-height, thus assuming that the various components had equal bearing on the three outcome. This assumption may not have been valid. For example, sickness of the mother may have mattered more for the height than the weight of the child whereas her educational level may have had the opposite relative importance although both of these variables were included in the same index (Familial Circumstances). The effect of biases that may have been introduced by such inappropriate weighing of components are impossible to assess but would have certainly contributed to a reduction in the proportion of the variance that could be explained.

3) From the findings of Johnston et al (21), it appears that the height and weight of the parents are important factors in determining the eventual height and, to a lesser degree, weight of their children. Weight-for-height would also be affected, obviously. Unfortunately, height and weight of the parents were not available in my data set. Afifi, using these parental parameters, was able to explain 28% of her children's overall size, 24% of their growth velocity, and 12% of their length (2).

4) On the one hand, using a sample representative of the general population increased the external validity of the results and the extent of their applicability to general

health policy. At the same time it may have diluted the effects found because of the greater heterogeneity of the sample in its exposure to predictors and in its levels of outcomes. On the other hand, analyses of smaller groups within the sample, such as the malnourished or the infants, may have provided more homogeneous exposures to factors of interest but less stable and reliable results due to a lower case to predictor ratio.

5) As for the likelihood of death, the inclusion of Feeding Practices and Current Morbidity might have explained more of the variance and permitted a better discriminating equation. However, because death of children could occur at any time during the study, it was not possible to use these two indexes derived from longitudinally collected variables since some of their components ceased to be collected after the death of the children.

6.2 Collinearity:

From the correlation tables (Appendix E), it can be appreciated that inter-correlation among independent variables may have led to their exclusion from the final regression equation, not on the basis of a lack of influence on the growth indicators but rather because of multicollinearity. The relation Familial with Econdx (r =.300), and perphaps Nutridex with Econdx (r = .187) and

Nutridex with Morbid2 (r = .182) might have had such a potential for mutual exclusion. Independent variables that were excluded from the final regression equation where their correlated companion was included may be of importance nevertheless, especially if they had a significant zero-order correlation with the corresponding outcome variable. This was especially prominent for the association of Familial Circumstances with Economic Status. As stated by Victora et al. (44) a problem often associated with all such studies, is that socio-economic status may act as a confounding variable, since environmental factors are also indicators of wealth. Although I have tried to separate the economic factor from other familial variables, I have left in the Familial Circumstances index some environmental variables that are undoubtedly determined by economic factors.

6.3 Comparison of the group having food availability data with the rest of the sample regarding feeding practices:

The design of the original study from which the dataset was obtained included food availability data collection on only half of the sample to check whether this practice would modify mothers' feeding practice of their babies. To test for this potential bias, I compared the groups on 'age at weaning', 'age at starting bottle feeding', and 'age at starting solid food'; the two groups were not statistically

different by the multiple comparison criterion (p-value was higher than 0.08). 'Age at starting porridge' gave the only significant difference at 0.03 (t-score = 2.1883, df 996) but, allowing for the multiple comparison effect that increases the likelihood of finding a difference in testing several variables occurring by chance alone, then that loses significance as well. Hence findings of the regression analyses on Feeding Practices whether done on the whole group or on the half sample (containing the Food Availability variable) should be comparable.

6.4 Comparison of children used in the regression analyses with those excluded because of missing data:

A part of the sample was consistently excluded from the analysis because of lack of data on some variables. The main reason for missing anthropometric data at the last measurement of the study (Time 4) was death of the child. None of the dead children had data available at Time-4 so were excluded from regression analyses on growth indicators. This may mean that the most malnourished children (and the most susceptible of dying) may have been excluded from the growth regressions on malnourished subgroups, thereby underestimating the effects. The remainder of the missing group (without the dead children) has been compared on some of the available variables with the used part of the sample to check for the likelihood of a systematic selection bias

other than death of the child. The tables in Appendix J demonstrate that such a bias was operative. The missing cases differ rather significantly from the used cases on many of the variables compared. These are: age at weaning, marital status, father's alcohol intake, the parents' incomes, the mother's educational level, the mother's allowance and other financial support, and the incidence of diarrhea. In some instances, the missing group had worse (i.e. considered detrimental to health) values than the cases retained but on other variables the situation was reversed so no consistent bias (in only one direction) appeared to be present. The differences create reservations about the generalizability of the conclusions to the rest of the Zambian children. Using this secondary data set, there is no way of knowing the reasons why certain data was collected on some children but not on others; it is thus impossible to determine in what direction the potential biases introduced by the "missing cases" influence the regression coefficients.

Chapter 7

7.0 CONCLUSION

7.1 Growth:

In this study, variables appearing to have an influence on growth are, in order of importance: Prior Morbidity, Economic Status, and Familial Circumstances. The latter two predictors are highly correlated with each other and it is difficult to decide which one is more important, or which one influences the other one. When Food Availability was included in the analysis, it also came out as a factor not to be ignored.

When the malnourished subgroups of stunted, underweight, and wasted children were analyzed, the "intermediate variables" of Feeding Practices and Current Morbidity did appear to have a significant influence on some aspects of the growth of malnourished children. Feeding Practices had a direct favourable effect on height and on weight. In that group of children, Prior Morbidity also played an important detrimental role. It should be remembered, as previously stated, that the children who died during the study were excluded from regression analyses on growth indicators because none of them had antropometric data available at Time-4. Some of the most malnourished children may thus

have been excluded from the growth analyses, thereby underestimating the effects.

While children older than 6 months of age had findings similar to the general sample, younger children's growth was influenced almost exclusively by Prior Morbidity with a possible benefit of preventive measures on their height.

It should be remembered that children not used in the study because of missing data (up to 27% of original sample) differed significantly from the cases used in the analyses. This may mean that the findings above may not be generalizable to all the children of Zambia. The fact that the cases used which had many missing variable values were significanty different on important items from the ones which had only a few missing values (5 or less) (see Appendix L) brings into question the validity of any negative statistical inference since missing value imputation was done similarly for the whole sample thus eliminating potentially significant differences between individuals. From results listed in Appendix L, one can see that this would be especially true for location of residence, stability of marital relationship, parental education, and type of foods available in the home. As well, as mentioned during the discussion of methodological issues, validity of the conclusions drawn from analyses of

the subgroups is questionable in view of often unaccceptable statistical power levels which may have caused relevant determinants to be rejected.

7.2 Mortality:

The analysis of predictors of death in the children was not as satisfactory as the ones on growth. Economic Status and Food Availability which were inversely related to mortality were the only associations that made intuitive sense. The only way to explain the other relationships is to look for potential biases introduced by the study design as already commented upon or to consider the potential influence of cultural beliefs, attitudes, and behaviours unknown to the current investigator and which may have affected the indexes in ways that would be difficult to undertand without a detailed knowledge of the culture of the Zambian tribes comprising the study population. As discussed previously, the study sample did not meet rigourously some of the basic assumptions of discriminant analysis thus providing less than ideal conditions for this type of analysis. Consequently, the derived equations were not very sensitive and even less specific in helping to predict which child may suffer a premature death but were nevertheless better than what would be obtained by chance alone. Introduction of feeding practice and current morbidity data into the equation would probably have improved predictive ability, if

it could have been done, given a different index design. Univariate analysis outlined certain factors as predictors of an early death, these are: being premature, being a twin, living in a squatter town, and getting measles.

7.3 Implications for Public Health:

Recalling the many caveats caused by the large number of imputations necessary for missing values and the exclusion of at least one quarter of the original sample from the analyses, the results of this study are probably best used only to suggest potential areas where further research may be useful.

Nevertheless, until further research provides a better understanding of the processes involved in human growth, and if one is anxious to implement some public health measures, the actions that would appear to be the most useful to better the growth of most Zambian children, healthy and malnourished, very young and older ones alike, are, in order of importance: the treatment and prevention of childhood diseases; the implementation of social measures that will improve the education and the health of their parents; the creation of new economic incentives that should in turn ameliorate environmental conditions of the household such as easy access to treated piped water and to proper disposal of sewage; the provision of the financial means to buy an

adequate quantity and quality of food and to access competent medical care when needed. One should remember that the amount of variance explained by the various analyses is small (less than 7% in most cases) so that the impact of all these preventive and corrective measures on the growth of the children may not be very important.

The relationships of the various indexes studied with the likelihood of an early death appears complex and difficult to interpret. Consequently, the study could not come up with any useful guidelines for preventive action based on these. However, the univariate analyses suggests that prevention of prematurity and low birth weight, and immunization with measles vaccine may help reduce the death toll in this population.

REFERENCES

(1) Abramson JH. Survey Methods in Community Medicine. Churchill Livingstone 1984. p. 113.

(2) Afifi ZEM. Determinants of growth of infants in an Egyptian village: maternal anthropometry, birth interval, solid food and death of siblings. Hum Biol 1985; 57 (4): 649-58.

(3) Bairagi R. A comparison of five anthropometric indices for identifying factors of malnutrition. Am J Epidemiol 1987; 126 (2): 258-267.

(4) Bhandari B, Mandowara SL. Does grading of PEM with Reference to Harvard Standard Need a Change? Indian J Ped 1982; 49(397): 162-5.

(5) Bhaskaram P. Infection and Malnutrition among Poor Children. Indian J Ped 1987; 54: 535-45.

(6) Bindon JR. The influence of infant feeding patterns on growth of children in American Samoa. Med Anthropol Spring 1985; 183-195.

(7) Cederblad M, Rahim SI. Effects of rapid urbanization on child behaviour and health in a part of Khartoum, Sudan -- I. Socio-economic changes 1965-1980. Social Science & Medicine 1986; 22 (7): 713-21.

(8) Crean JTG, Ling SL, Pearson DJ, Shaw SN, Stewart RJ, Wilson JH. Dharavi Project 1985 - An investigation into patterns of infant feeding and malnutrition in the urban slum of Dharavi in Bombay. J Trop Ped 1987; 33(Suppl): 2-29.

(9) Dettwyler KA. Infant feeding in Mali, West Africa: variations in belief and practice. Social Science & Medicine 1986; 23 (7): 651-64.

(10) Dettwyler KA. Breastfeeding and weaning in Mali: cultural context and hard data. Social Science & Medicine 1987; 24 (8): 633-44.

(11) Devries MW. Temperament and infant mortality among the Masai of East Africa. American Journal of Psychiatry Oct 1984; 141 (10): 1189-94.

(12) Dixon SD, LeVine RA, Brazelton TB. Malnutrition: a closer look at the problem in an East African village. Develop Med Child Neurol 1982; 24: 670-85.

(13) El Samani EF, Willett WC, Ware JH. Association of malnutrition and diarrhea in children aged under five years. Am J Epidemiol 1988; 128 (1): 93-105.

(14) Excler JL, Nicolas E, Mojon M. La Malnutrition Protéino-Energétique en Milieu Urbain Africain (Togo). Medecine Tropicale 1985; 45(2): 155-61.

(15) Gerein N. Is growth monitoring worthwhile? Health policy and Planning 1988; 3 (3): 181-94.

(16) Gomez F, Galvan R, Frenk S, et al. Mortality in second and third degree malnutrition. J Trop Paed 1956; 2: 77-83.

(17) Greenland S. Modelling and variable selection in epidemiologic analysis. Am J Pub Health 1989; 79 (3): 340-9.

(18) Igun UA. Child-feeding habits in a situation of sccial change: the case of Maiduguri, Nigeria. Social Science & Medicine 1982; 16 (7): 769-81.

(19) Ijsselmuiden CB. Nutritional status of children under the age of 5 years in northern Gasankulu. A cross-sectional assessment. South African Medical Journal Mar 3 1984; 65 (9): 346-7.

(20) Jelliffe DB. The Assessment of the Nutritional Status of the Community. World Health Organization Geneva 1966.

(21) Johnston FE, Scholl TO, Newman BC, Cravioto J, De Licardie ER. An analysis of environmental variables and factors associated with growth failure in a Mexican village. Hum Biol 1980; 52 (4): 627-37.

(22) Kalton G, Kasprzyk D. The Treatment of Missing Survey Data. Survey Methodology; Statistics Canada June 1986; 12 (1): 1-16.

(23) Kanawati AA, McLaren DS. Failure to thrive in Lebanon - II. An investigation of the Causes. Acta Paediat Scand 1973; 62: 571-6.

(24) Kerlinger FN, Pedhazur EJ. Multiple Regression in Behavioral Research. Holt, Rhinehart & Winston Inc., New York. 1973; 304-33.

(25) Kusin JA. Traditional infant feeding practices: right or wrong? Social Science & Medicine 1985; 21 (3): 283-6.

(26) Leigh PJ. Assessing the importance of an independent variable in multiple regression: is stepwise unwise? J Clin Epidemiol 1988; 41(7): 669-77.

(27) Linhares EDR, Round JM, Jones DA. Growth, bone maturation, and biochemical changes in Brazilian children from two different socio-economic groups. Am J Clin Nutr 1986; 44: 552-8.

(28) Lopez de Romana G, Brown KH, Black RE, Kanashiro HC. Longitudinal Studies of Infectious Diseases and PHysical Growth of Infants in Huascar, an Underprivileged Peri-urban Community in Lima, Peru. Am J Epidemiol 1989; 129(4): 769-84.

(29) Lovel HJ, Sabir NI, Cleland J. Why are toddler girls at risk of death and undernutrition in a slum area of Pakistan? Lancet April 7 1984; 797.

(30) Martorell R, Habicht JP, Yarbrough C, Lechtig A, Klein RE, Western KA. Acute morbidity and physical growth in rural Guatemalan children. Am J Dis Child 1975; 129: 1296-1301.

(31) Martorell R, Leslie J, Moock PR. Characteristics and determinants of child nutritional status in Nepal. Am J Clin Nutr 1984; 39: 74-86.

(32) Martorell R, O'Gara C. Breastfeeding, infant health, and socioeconomic status. Med Anthropol Spring 1985; 173-81.

(33) McDowell I, Hoorweg J. Social environment and outpatient recovery from malnutrition. Ecol Food Nutrition 1975; 4: 91-101.

(34) McDowell I, Savage King F. Interpretation of arm circumference as an indicator of nutritional status. Arch Dis Child 1982. 57: 292-6.

(35) Miller RG. Simultaneous Statistical Reference. Mc-Graw-Hill 1981. p.241

(36) Odebiyi AI. Child-bearing practices among nursing mothers in Ile-Ife, Nigeria. Child: Care, Health & Development Sep-Oct 1985; 11 (5): 325-35.

(37) Ojofeitimi EO. Diet restriction by pregnant women in Nigeria. International Journal of Gynaecology & Obstetrics Apr 1982; 20 (2): 99-103.

(38) Olango P, Aboud F. Determinants of mothers' treatment of diarrhea in rural Ethiopia. Social Science & Medicine 1990; 31 (11): 1245-9.

(39) Oyedeji GA. The Present Day Epidemiology of Severe Protein-Energy Malnutrition in Nigeria. Clin Pediat 1984; 23(11): 623-8.

(40) Pickering H. Social and environmental factors associated with diarrhoea and growth in young children: child health in urban Africa. Soc Sci Med 1985; 21 (2): 121-7.

(41) Sukary-Stolba S. Food classifications and the diets of young children in rural Egypt. Social Science & Medicine 1987; 25 (4): 401-4.

(42) Tanner JM. Physical Development. Brit Med Bull 1986; 42 (2): 131-8.

(43) Underwood P, Margetts B. Cultural change, growth and feeding of children in an isolated rural region of Yemen. Soc Sci Med 1987; 25 (1): 1-7.

(44) Victora CG, Vaughan JP, Kirkwood BR, Martines JC, Barcelos LB. Risk factors for malnutrition in Brazilian children: the role of social and environmental variables. Bulletin of WHO 1986; 64 (2): 299-309.

(45) Victora CG, Vaughan JP, Martines JC, Barcelos LB. Is prolonged breast-feeding associated with malnutrition. Am J Clin Nutr 1984; 39: 307-14.

(46) Victora CG, Smith PG, Vaughan JP, Nobre LC, et al. Infant Feeding and Deaths due to Diarrhea. Am J Epidem 1989; 129(5): 1032-41.

(47) Waterlow JC, Buzzina R, Keller W, et al. The Presentation and Use of Height and Weight Data for Comparing the Nutritional Status of Groups of Children under the Age of 10 Years. Bulletin of WHO 1977; 55: 489-98.

(48) WHO. World Health Statistical Quarterly 1985; 38: 337-40.

(49) WHO. World Health Statistics Annual 1983.

APPENDIX A
ORIGINAL COLLECTED VARIABLES USED

Familial Environment:

Township of residence
Parent's education, employment, income, health, alcohol intake
Mother's age, parity, marital status, financial allowance, other source of support
Number of older siblings, number of younger siblings, presence of handicapped children, history of deaths among siblings
Source of water for household

Child's demographics:

Sex, age, place of birth, prematurity at birth
If applicable: occurrence of death, rank in the family; who living with (parent/s); been to village or not

Diet & Food Availability:

Child: Difficulty in breast feeding, age at weaning, age started bottle, age started porridge, age started solids.
Household: Cultivated garden; number of times seen eating meat, fish, beans, good staple, milk; amount of money spent on food.

Morbidity:

Types of illness, occurrence of measles, past history of malnutrition, number of hospitalizations and diagnoses, traditional medical treatment.

APPENDIX B
COMPONENTS OF REDUCED INDEXES

* = multiplied by

Morbid1: (8/15 items)

Ever-hospitalized + [ever-CNS * 2] + [diarrhea * 3] + rxhome + measles + diff-lact + prematurity.

Previndx: (4/5 items)

[Under5-clinic / 2] + [dpt * 2] + [polio * 2] + [bcg * 2].

Nutridex: (4/7 items)

[meat * 2] + [fish * 2] + staples + [milk * 2].

Familial: (15/17 items)

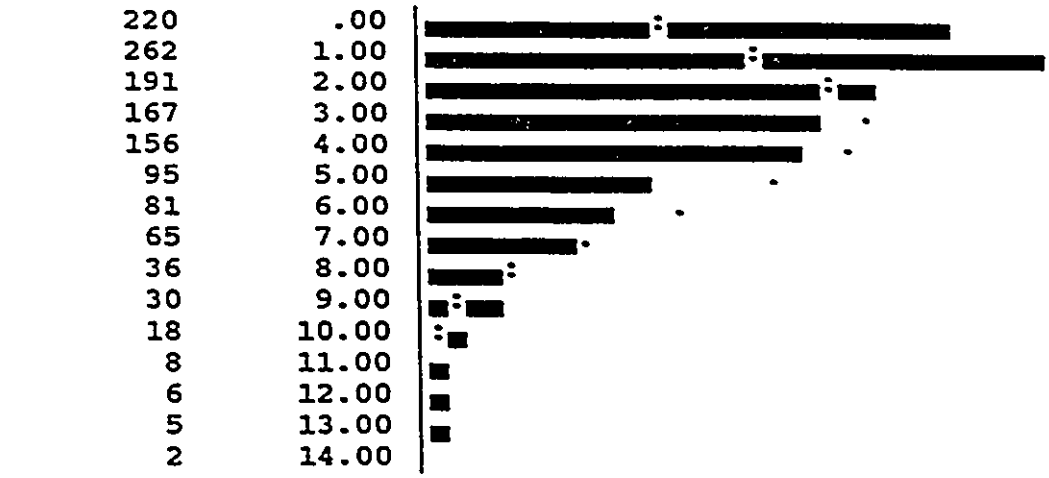
Town + nyoung + [healthpa * 2] + [healthma * 2] + parity + deaths + [marriage * 2] + [edpa * .3] + [edma * .3] + drinkpa + drinkma + water + handicap + born + living-with-who.

Econdx: (4/6 items)

[jobpa / 2] + father-income + allows + mother-income.

APPENDIX C
DISTRIBUTION GRAPHS FOR INDEXES[1]

MORBID1

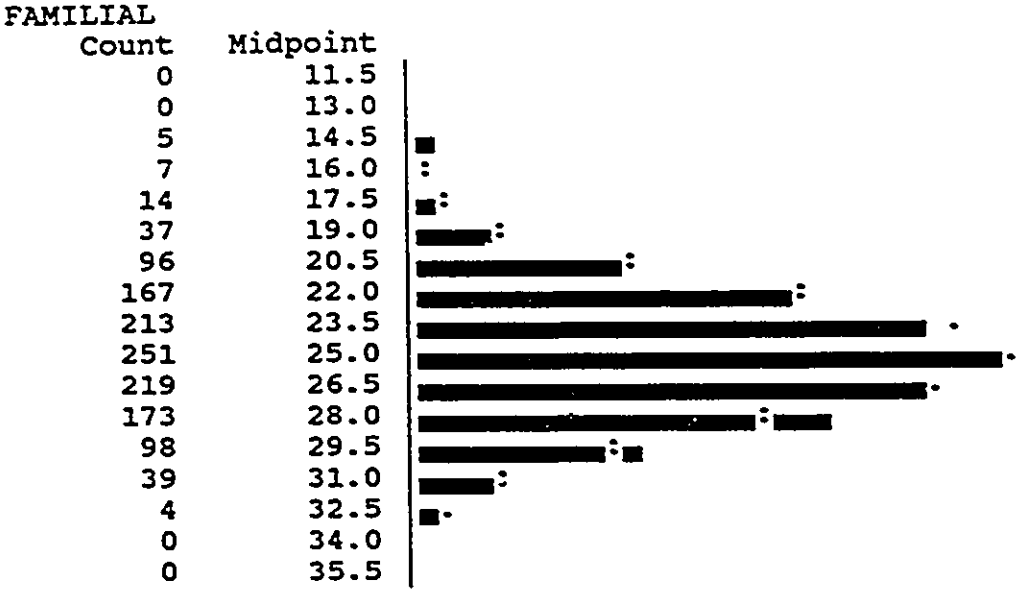


Percentile	Value	Percentile	Value	Percentile	Value
1.00	.000	10.00	.000	50.00	2.000

Valid Cases 1342 Missing Cases 0

[1] Dots on graphs show outline of normal distribution

FAMILIAL

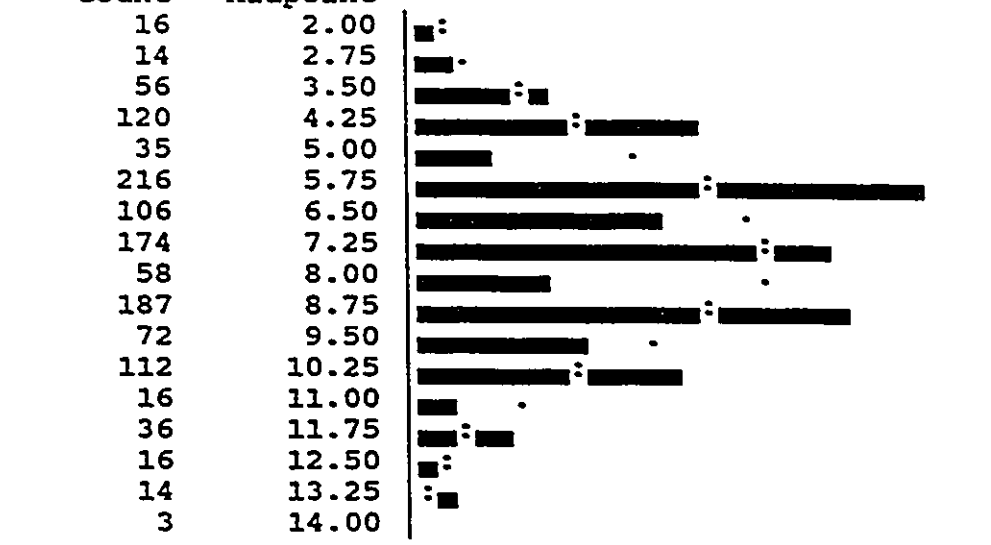


Percentile	Value	Percentile	Value	Percentile	Value
1.00	16.972	10.00	20.864	50.00	25.000

Valid Cases 1323 Missing Cases 19

ECONDX

Count	Midpoint
16	2.00
14	2.75
56	3.50
120	4.25
35	5.00
216	5.75
106	6.50
174	7.25
58	8.00
187	8.75
72	9.50
112	10.25
16	11.00
36	11.75
16	12.50
14	13.25
3	14.00



I....+....I....+....I....+....I....+....I....+....I....+....I

0 80 160 240

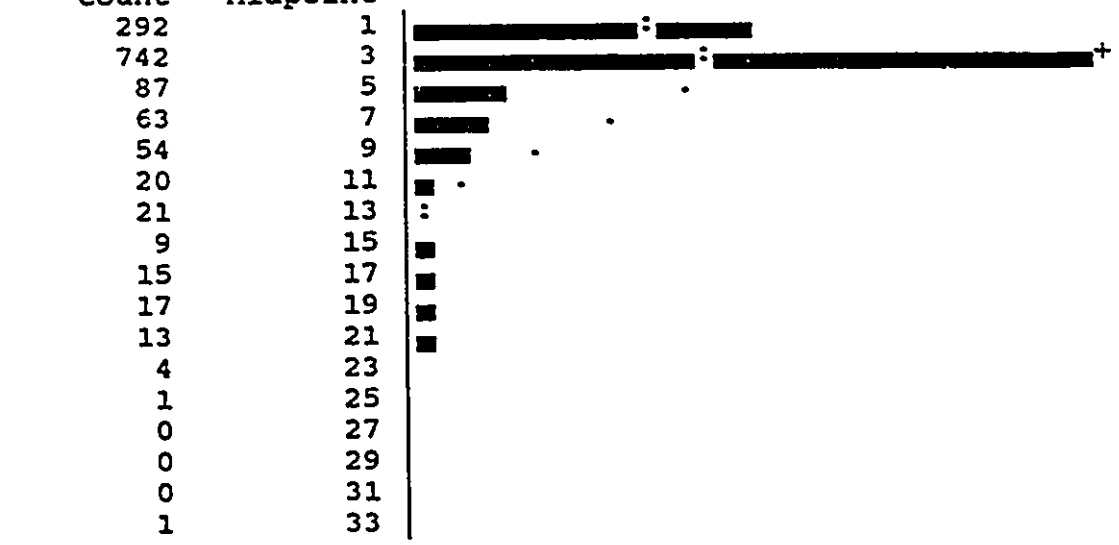
Histogram Frequency

Percentile	Value	Percentile	Value	Percentile	Value
1.00	2.000	10.00	4.500	50.00	7.500

Valid Cases 1251 Missing Cases 91

PREVINDX

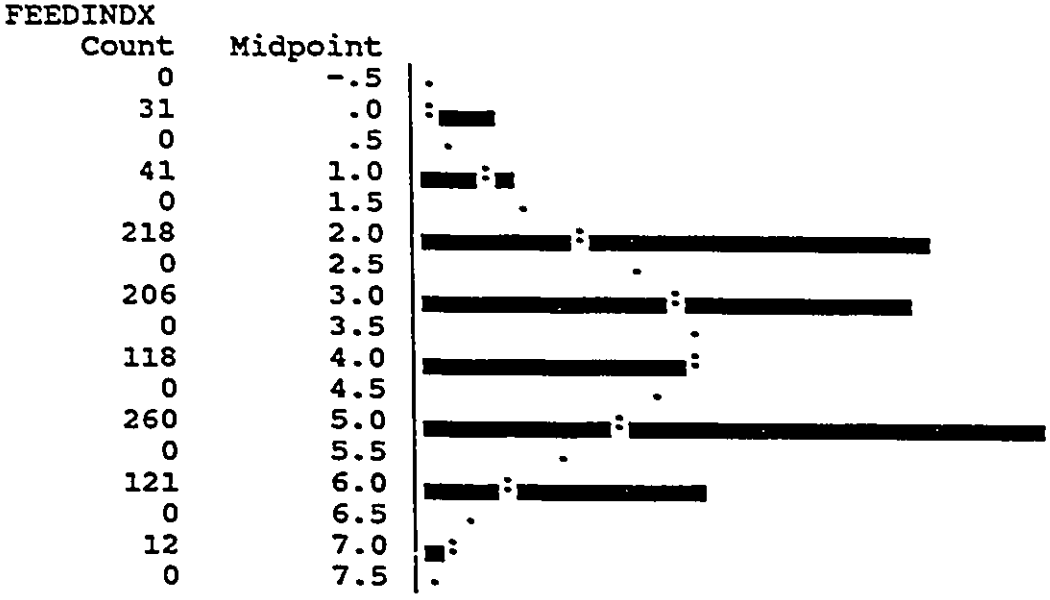
Count	Midpoint
292	1
742	3
87	5
63	7
54	9
20	11
21	13
9	15
15	17
17	19
13	21
4	23
1	25
0	27
0	29
0	31
1	33



Percentile	Value	Percentile	Value	Percentile	Value
1.00	.000	10.00	.000	50.00	2.000

Valid Cases 1339 Missing Cases 3

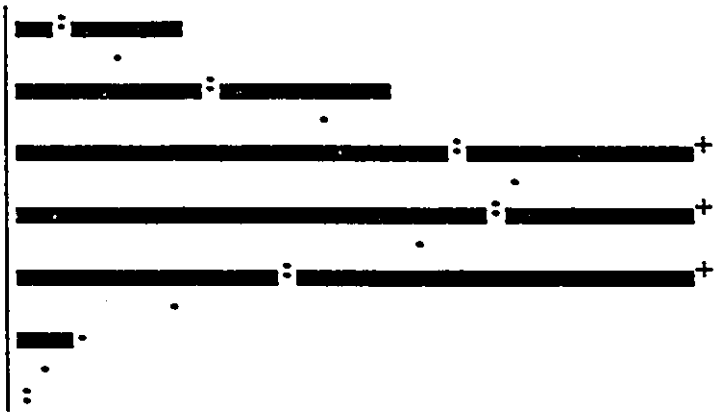
FEEDINDX



Percentile	Value	Percentile	Value	Percentile	Value
1.00	.000	10.00	2.000	50.00	4.000
Valid Cases	1007	Missing Cases	335		

MORBID2

COUNT	VALUE
71	.00
0	1.00
159	2.00
1	3.00
339	4.00
1	5.00
375	6.00
0	7.00
365	8.00
0	9.00
27	10.00
0	11.00
4	12.00



I.....I.....I.....I.....I.....I

0 80 160 240

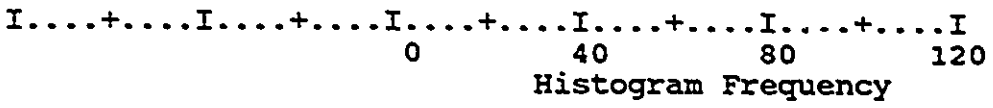
Histogram Frequency

Percentile	Value	Percentile	Value	Percentile	Value
1.00	.000	10.00	2.000	50.00	6.000

Valid Cases 1342 Missing Cases 0

NUTRIDEX

Count	Midpoint
62	3
126	13
119	23
24	33
9	43
22	53
11	63
19	73
30	83
25	93
25	103
43	113
24	123
13	133
7	143
15	153
6	163

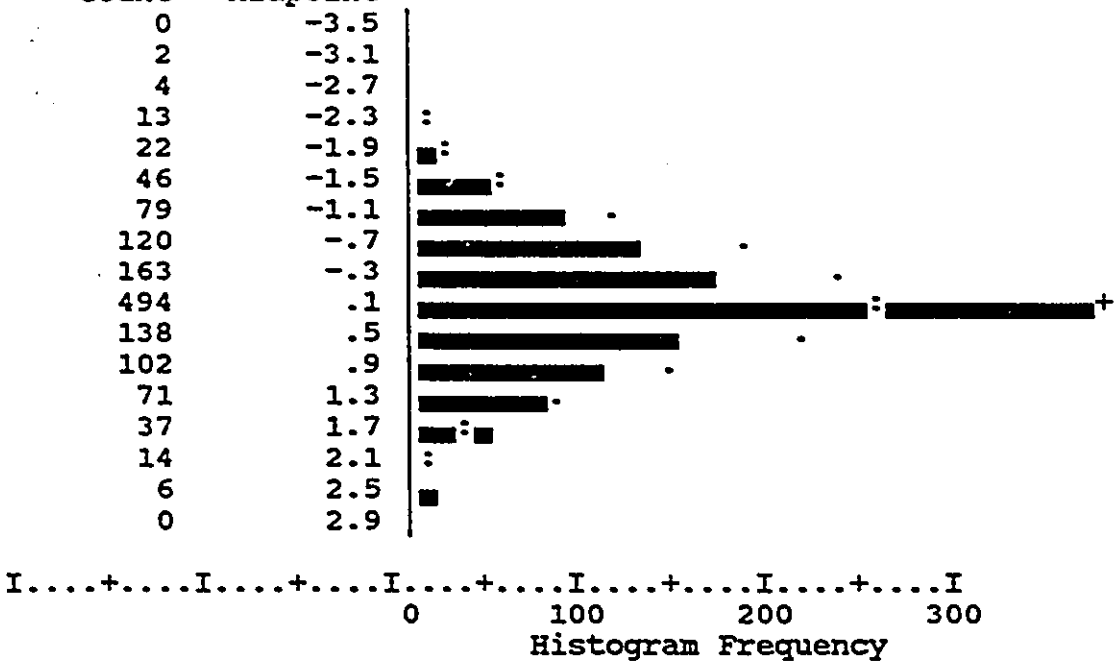


Percentile	Value	Percentile	Value	Percentile	Value
1.00	.000	10.00	6.500	50.00	26.000

Valid Cases 580 Missing Cases 762

APPENDIX D
DISTRIBUTION GRAPHS FOR GROWTH INDICATORS[1]

ZHA4



ZHA4

Mean	-.002	Median	.000	Mode	.000
Std Dev	.834				

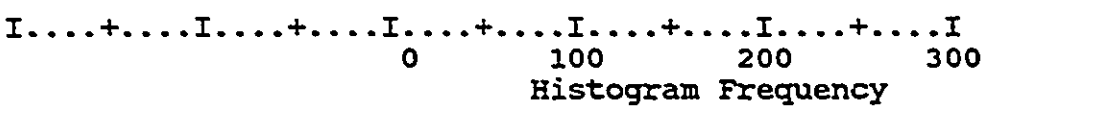
Percentile	Value	Percentile	Value	Percentile	Value
1.00	-2.266	10.00	-1.059	50.00	.000

Valid Cases 1311 Missing Cases 31

[1] Dots on graphs show outline of normal distribution

ZWA4

Count	Midpoint
1	-3.2
1	-2.8
5	-2.4
12	-2.0
56	-1.6
82	-1.2
117	-.8
168	-.4
489	-.0
135	.4
110	.8
87	1.2
50	1.6
18	2.0
7	2.4
3	2.8
1	3.2



ZWA4

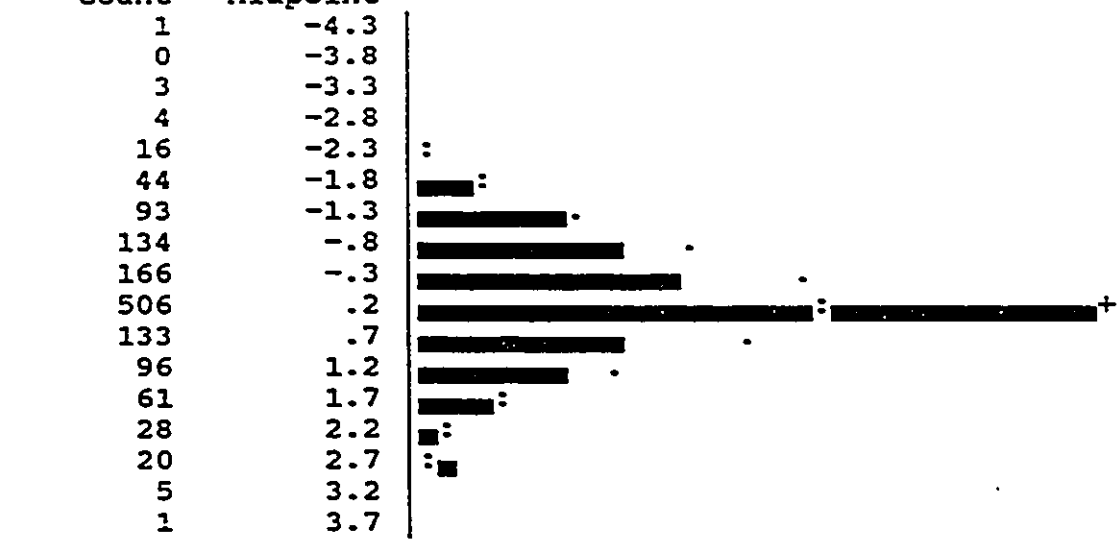
Mean	.000	Median	.000	Mode	.000
Std Dev	.837				

Percentile	Value	Percentile	Value	Percentile	Value
1.00	-2.016	10.00	-1.093	50.00	.000

Valid Cases 1342 Missing Cases 0

ZWH4

Count	Midpoint
1	-4.3
0	-3.8
3	-3.3
4	-2.8
16	-2.3
44	-1.8
93	-1.3
134	-.8
166	-.3
506	.2
133	.7
96	1.2
61	1.7
28	2.2
20	2.7
5	3.2
1	3.7



I....+....I....+....I....+....I....+....I....+....I....+....I

0 120 240 360

Histogram Frequency

ZWH4

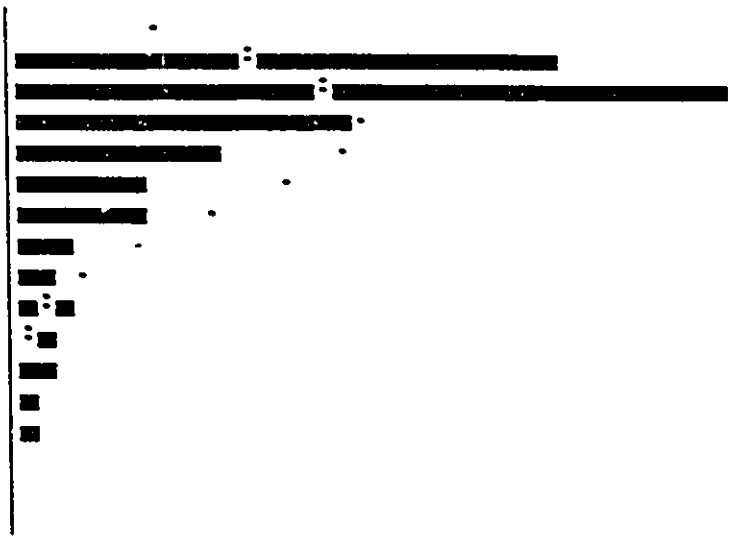
Mean	.046	Median	.000	Mode	.000
Std Dev	.984				

Percentile	Value	Percentile	Value	Percentile	Value
1.00	-2.299	10.00	-1.226	50.00	.000

Valid Cases	1311	Missing Cases	31

HAPCT HEIGHT FOR AGE4 %ILE

Count	Midpoint
0	-10.0
231	-2.5
306	5.0
144	12.5
85	20.0
54	27.5
53	35.0
20	42.5
19	50.0
20	57.5
12	65.0
13	72.5
7	80.0
6	87.5
3	95.0
3	102.5
0	110.0



I....+....I....+....I....+....I....+....I....+....I....+....I

0 80 160 240

Histogram Frequency

HAPCT HEIGHT FOR AGE4 %ILE

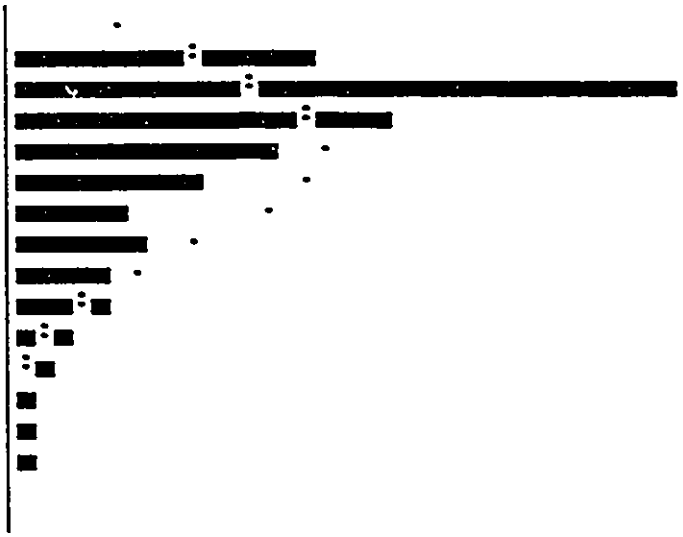
Mean	14.746	Median	6.535	Mode	.000
Std Dev	19.366				

Percentile	Value	Percentile	Value	Percentile	Value
1.00	.000	10.00	.177	50.00	6.535

Valid Cases	976	Missing Cases	366

WAPCT

Count	Midpoint
0	-10.0
130	-2.5
281	5.0
160	12.5
110	20.0
77	27.5
47	35.0
55	42.5
43	50.0
43	57.5
24	65.0
14	72.5
8	80.0
7	87.5
6	95.0
2	102.5
0	110.0



I....+....I....+....I....+....I....+....I....+....I....+....I
0 80 160 240
Histogram Frequency

WAPCT

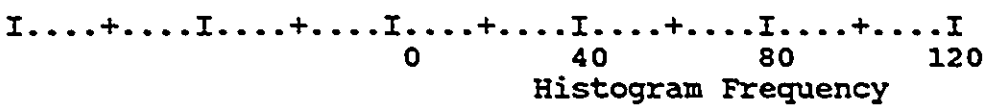
Mean	20.580	Median	12.410	Mode	.630
Std Dev	21.607				

Percentile	Value	Percentile	Value	Percentile	Value
1.00	.040	10.00	.840	50.00	12.410

Valid Cases	1007	Missing Cases	335

WHPCT

Count	Midpoint
0	-10.0
8	-2.5
80	5.0
103	12.5
90	20.0
93	27.5
74	35.0
88	42.5
74	50.0
92	57.5
75	65.0
71	72.5
51	80.0
42	87.5
31	95.0
4	102.5
0	110.0



WHPCT

Mean	43.496	Median	41.700	Mode	14.160
Std Dev	26.179				

Percentile	Value	Percentile	Value	Percentile	Value
1.00	1.904	10.00	9.690	50.00	41.700

Valid Cases	976	Missing Cases	366

APPENDIX E
CORRELATIONS among the VARIABLES USED IN THE REGRESSIONS

Cases used for Dependent: Feedindx (938 cases)

	FEEDINDX	MORBID1	PREVINDX	FAMILIAL	ECONDX
FEEDINDX					
MORBID1	.006				
PREVINDX	.037	.127			
FAMILIAL	-.058	-.009	.130		
ECONDX	-.020	-.022	.051	.287	
NUTRIDEX (414 cases)	-.005	.058	.028	.068	.188

Cases used for Dependent: Morbid2 (938 cases)

	FEEDINDX	MORBID2	MORBID1	PREVINDX	FAMILIAL	ECONDX
FEEDINDX						
MORBID2	.167					
MORBID1	.006	.125				
PREVINDX	.037	-.000	.127			
FAMILIAL	-.058	-.074	-.009	.130		
ECONDX	-.020	-.064	-.022	.051	.287	
NUTRIDEX (414 cases)	-.005	-.189	.058	.028	.068	.188

Cases used for Dependent: Z-score of Height for Age (N=911)

	FEEDINDX	MORBID1	FAMILIAL	ECONDX	PREVINDX	MORBID2	ZHA4
FEEDINDX							
MORBID1	.003						
FAMILIAL	-.055	-.007					
ECONDX	-.023	-.026	.293				
PREVINDX	.042	.127	.131	.054			
MORBID2	.166	.113	-.077	-.074	-.000		
ZHA4	-.046	-.150	.095	.082	.030	-.013	
NUTRIDEX (405 cases)	.010	.060	.059	.192	.029	-.183	.113

Cases used for Dependent: Z-score of Weight for Age (N=938)

	FEEDINDX	MORBID1	FAMILIAL	ECONDX	PREVINDX	MORBID2	ZWA4
FEEDINDX							
MORBID1	.006						
FAMILIAL	-.058	-.009					
ECONDX	-.020	-.022	.287				
PREVINDX	.037	.127	.130	.051			
MORBID2	.167	.125	-.074	-.064	-.000		
ZWA4	-.034	-.189	.066	.086	-.039	-.040	
NUTRIDEX (414 cases)	-.005	.058	.068	.188	.028	-.189	.082

Cases used for Dependent: Z-score of Weight/Height (N=911)

	FEEDINDX	MORBID1	FAMILIAL	ECONDX	PREVINDX	MORBID2	ZWH4
FEEDINDX							
MORBID1	.003						
FAMILIAL	-.055	-.007					
ECONDX	-.023	-.026	.293				
PREVINDX	.042	.127	.131	.054			
MORBID2	.166	.113	-.077	-.074	-.000		
ZWH4	-.042	-.169	.039	.084	-.070	-.045	
NUTRIDEX (405 cases)	.010	.060	.059	.192	.029	-.183	.046

Cases used for Dependent: Weight for Height Percentile at Time 2 years (N = 910)

	FEEDINDX	MORBID1	FAMILIAL	ECONDX	PREVINDX	MORBID2	WHPCT4
FEEDINDX							
MORBID1	.005						
FAMILIAL	-.057	-.005					
ECONDX	-.021	-.027	.296				
PREVINDX	.041	.128	.130	.055			
MORBID2	.167	.112	-.076	-.074	-.000		
WHPCT4	-.025	-.202	.007	.068	-.089	-.021	
NUTRIDEX (404 cases)	.013	.058	.062	.191	.030	-.184	.020

APPENDIX F
GENERAL SAMPLE - REGRESSIONS

Changes in R-square value

Large Sample:

Dependent Variable = Feeding Index

	Original * (N = 938)		Reduced (N = 941)	
1..	FAMILIAL	.00224	PREVINDX	.00709

Large Sample:

Dependent Variable = Current Morbidity Index

	Original (N = 938)		Reduced (N = 941)	
1..	FEEDINDX	.02698	FEEDINDX	.02777
2..	MORBID1	.04126	MORBID1	.03071
3..	FAMILIAL	.04425	ECONDX	.03350

---

* Original = indexes with maximum number of components.
  Reduced = indexes with reduced number of components.

Changes in R-square value

Small Sample:

Dependent Variable = Feeding Index

	Original (N = 414)		Reduced (N = 414)	
1..	PREVINDX	.00879	PREVINDX	.01876

Small Sample:

Dependent Variable = Current Morbidity Index

	Original (N = 414)		Reduced (N = 414)	
1..	FEEDINDX	.03337	FEEDINDX	.03337
2..	NUTRIDEX	.06636	NUTRIDEX	.04146
3..	MORBID1	.07520		

Changes in R-square value

Large Sample:

Dependent Variable = Zscore Height for Age at Time 2 years

	Original (N = 911)		Reduced (N = 914)	
1..	MORBID1	.02140	MORBID1	.02553
2..	FAMILIAL	.02915	ECONDX	.03280

Dependent Variable = Zscore Weight for Age at Time 2 years

	Original (N = 938)		Reduced (N = 941)	
1..	MORBID1	.03468	MORBID1	.04297
2..	ECONDX	.04031	ECONDX	.05195

Dependent Variable = Zscore Weight for Height at Time 2 years

	Original (N = 911)		Reduced (N = 914)	
1..	MORBID1	.02756	MORBID1	.03492
2..	ECONDX	.03281	ECONDX	.04306
3..	PREVINDX	.03467		

Changes in R-square value

Small Sample:

Dependent Variable = Zscore Height for Age at Time 2 years

	Original (N = 405)		Reduced (N = 405)	
1..	ECONDX	.01914	MORBID1	.02242
2..	MORBID1	.03420	ECONDX	.04078
3..	FAMILIAL	.04191	FAMILIAL	.04525
4..	NUTRIDEX	.04885		

Dependent Variable = Zscore Weight for Age at Time 2 years

	Original (N = 414)		Reduced (N = 414)	
1..	MORBID1	.03595	MORBID1	.04811
2..	FAMILIAL	.04943	ECONDX	.05960
3..	NUTRIDEX *	.05447		

Dependent Variable = Zscore Weight for Height at Time 2 years

	Original (N = 405)		Reduced (N = 405)	
1..	FEEDINDX *	.00554	FEEDINDX	.00554
2..	MORBID1	.03764	MORBID1	.04231
3..	FAMILIAL	.04453	FAMILIAL	.04782

---

* Did not reach statistical significance at the 0.1 level when the direct method was used instead of the stepwise.

Regression Coefficients Weight for Height Percentile

(Regressions Weighted By Accuracy of Birth Date)

Dependent Variable = Weight for Height Percentile
at Time 2 years (N = 910)

Variable	B[1]	Beta[2]	p-value
MORBID1	-1.75798	-.18964	.0000
ECONDX	.74831	.06891	.0430
PREVINDX	-.33929	-.06294	.0573

Cumulative Adjusted Variance .04147

Dependent Variable = Weight for Height Percentile
at Time 2 years (N = 404)

Variable	B	Beta	p-value
MORBID1	-1.51664	-.16808	.0009
PREVINDX	-.45739	-.08960	.0771

Cumulative Adjusted Variance .03547

---

[1] Unstandardized Regression Coefficient
[2] Standardized Regression Coefficient

APPENDIX G
MALNOURISHED SUBGROUPS - REGRESSIONS

Stunted sub-group
(N = 86)

Correlation among independent variables used:

	FEEDINDX	MORBID2	MORBID1	PREVINDX	FAMILIAL
MORBID2	.282				
MORBID1	-.046	.073			
PREVINDX	-.003	-.058	.142		
FAMILIAL	.069	.148	.114	.184	
ECONDX	.046	.067	-.051	.083	.320

Regression Results

Feeding practice: No variables entered in equation.

Current Morbidity:

Variable	B[1]	Beta[2]	p-value	Cumulat. Variance
FEEDINDX	.33758	.28164	.0086	.06836

Height-for-Age (ZHA4):

Variable	B	Beta	p-value	Cumulat. Variance
FEEDINDX	.06788	.29528	.0058	.07632

[1] Unstandardized Partial Regression Coefficient
[2] Standardized Partial Regression Coefficient

Underweight sub-group
(N = 77)

Correlation among independent variables used:

	FEEDINDX	MORBID2	MORBID1	PREVINDX	FAMILIAL
MORBID2	.181				
MORBID1	-.117	.085			
PREVINDX	-.018	-.153	.019		
FAMILIAL	.221	.027	.095	.374	
ECONDX	.220	.018	-.069	.268	.331

Regression Results

Feeding practice:

Variable	B[1]	Beta[2]	p-value	Cumulat. Variance
FAMILIAL	.13560	.22065	.0538	.03600

Current Morbidity:

No variables entered.

Weight-for-Age (ZWA4):

Variable	B	Beta	p-value	Cumulat. Variance
MORBID1	-.02225	-.19498	.0874	.05035
FEEDINDX	.03353	.17169	.1314	

[1] Unstandardized Partial Regression Coefficient
[2] Standardized Partial Regression Coefficient

Wasted sub-group
(N = 114)

Correlation among independent variables used:

	FEEDINDX	MCRBID2	MORBID1	PREVINDX	FAMILIAL
MORBID2	.228				
MORBID1	-.115	.176			
PREVINDX	.223	.089	.191		
FAMILIAL	.071	.031	.229	.312	
ECONDX	.035	-.095	.018	.141	.304

Regression Results

Feeding practice:

Variable	B[1]	Beta[2]	p-value	Cumulat. Variance
PREVINDX	.07101	.25423	.0073	.05892
MORBID1	-.09398	-.16404	.0804	

Current Morbidity:

Variable	B	Beta	p-value	Cumulat. Variance
FEEDINDX	.40831	.25126	.0067	.076884
MORBID1	.19062	.20475	.0264	

Weight-for-Height (ZWH4):

Variable	B	Beta	p-value	Cumulat. Variance
MORBID2	.04325	.24742	.0085	.06542
MORBID1	-.03155	-.19392	.0381	

[1] Unstandardized Partial Regression Coefficient
[2] Standardized Partial Regression Coefficient

APPENDIX H
AGE GROUPS - REGRESSIONS

Correlations
Less than 6 Months of Age
(N = 252)

	FEEDINDX	MORBID2	MORBID1	PREVINDX	FAMILIAL	ECONDX
MORBID2	-.060					
MORBID1	-.077	.024				
PREVINDX	.080	-.109	.107			
FAMILIAL	-.007	-.062	-.048	.176		
ECONDX	.064	-.030	-.044	.119	.270	
NUTRIDEX (N = 121)	.022	-.220	-.107	-.033	.016	.228

	FEEDINDX	MORBID1	FAMILIAL	ECONDX	PREVINDX	MORBID2	NUTRIDEX
ZHA4	-.028	-.145	.081	.039	.109	-.008	.132
ZWA4	-.006	-.178	-.009	.075	-.013	-.074	.083
ZWH4	.005	-.183	-.048	.135	-.074	-.048	.029

More than 6 Months Old
(N = 686)

	FEEDINDX	MORBID2	MORBID1	PREVINDX	FAMILIAL	ECONDX
MORBID2	.166					
MORBID1	.113	.197				
PREVINDX	-.002	.039	.150			
FAMILIAL	-.057	-.071	.000	.112		
ECONDX	-.035	-.071	-.018	.020	.292	
NUTRIDEX (N = 293)	-.020	-.182	.149	.063	091	.170

	FEEDINDX	MORBID1	FAMILIAL	ECONDX	PREVINDX	MORBID2	NUTRIDEX
ZHA4	-.048	-.157	.099	.097	-.008	-.012	.104
ZWA4	-.039	-.199	.092	.089	-.051	-.026	.083
ZWH4	-.042	-.173	.066	.066	-.068	-.038	.054

Regressions
Less than 6 Months of Age
(Large Sample; N = 252)

FEEDINDX:

No variable entered

MORBID2:

Variable	B[1]	Beta[2]	p-value	Cumulat. Variance
PREVINDX	-.04124	-.10851	.0856	.00732

ZHA4:

Variable	B	Beta	p-value	Cumulat. Variance
MORBID1	-.04880	-.15951	.0121	.02917
PREVINDX	.02074	.12706	.0452	

ZWA4:

Variable	B	Beta	p-value	Cumulat. Variance
MORBID1	-.05463	-.17788	.0046	.02777

ZWH4:

Variable	B	Beta	p-value	Cumulat. Variance
MORBID1	-.05858	-.17662	.0050	.04170
ECONDX	.05473	.12693	.0429	

[1] Unstandardized Regression Coefficient
[2] Standardized Regression Coefficient

Less than 6 Months of Age
(Small Sample; N = 121)

FEEDINDX:

Variable	B[1]	Beta[2]	p-value	Cumulat. Variance
PREVINDX	.03979	.16388	.0725	.01868

MORBID2:

Variable	B	Beta	p-value	Cumulat. Variance
NUTRIDEX	-.00984	-.21996	.0153	.04038

ZHA4:

Variable	B	Beta	p-value	Cumulat. Variance
MORBID1	-.05072	-.17469	.0574	.02223

ZWA4:

Variable	B	Beta	p-value	Cumulat. Variance
MORBID1	-.05753	-.21064	.0204	.03634

ZWH4:

Variable	B	Beta	p-value	Cumulat. Variance
MORBID1	-.06350	-.20870	.0212	.05935
ECONDX	.06921	.17173	.0570	

[1] Unstandardized Regression Coefficient
[2] Standardized Regression Coefficient

Regressions
More than 6 Months Old
(Large sample; N = 686)

FEEDINDX:

Variable	B[1]	Beta[2]	p-value	Cumulat. Variance
MORBID1	.06518	.11265	.0031	.01125

MORBID2:

Variable	B	Beta	p-value	Cumulat. Variance
MORBID1	.16242	.17929	.0000	.05959
FEEDINDX	.22539	.14396	.0001	
ECONDX	-.06339	-.06306	.0895	

ZHA4:

Variable	B	Beta	p-value	Cumulat. Variance
MORBID1	-.05627	-.15593	.0001	.03481
FAMILIAL	.02614	.07891	.0490	
ECONDX	.02815	.07043	.0789	

ZWA4:

Variable	B	Beta	p-value	Cumulat. Variance
MORBID1	-.07246	-.19905	.0000	.04539
FAMILIAL	.03070	.09255	.0134	

ZWH4:

Variable	B	Beta	p-value	Cumulat. Variance
MORBID1	-.07590	-.17344	.0000	.03155
FAMILIAL	.02693	.06704	.0801	

[1] Unstandardized Regression Coefficient
[2] Standardized Regression Coefficient

More than 6 Months Old
(Small sample; N = 293)

FEEDINDX:

Variable	B[1]	Beta[2]	p-value	Cumulat. Variance
MORBID1	.05463	.09925	.0899	.00645

MORBID2:

Variable	B	Beta	p-value	Cumulat. Variance
FEEDINDX	.29594	.18658	.0008	.11737
MORBID1	.19719	.22586	.0001	
NUTRIDEX	-.01097	-.21211	.0002	

ZHA4:

Variable	B	Beta	p-value	Cumulat. Variance
ECONDX	.05127	.13348	.0343	.04571
MORBID1	-.03854	-.11050	.0573	
FAMILIAL	.03608	.11304	.0726	

ZWA4:

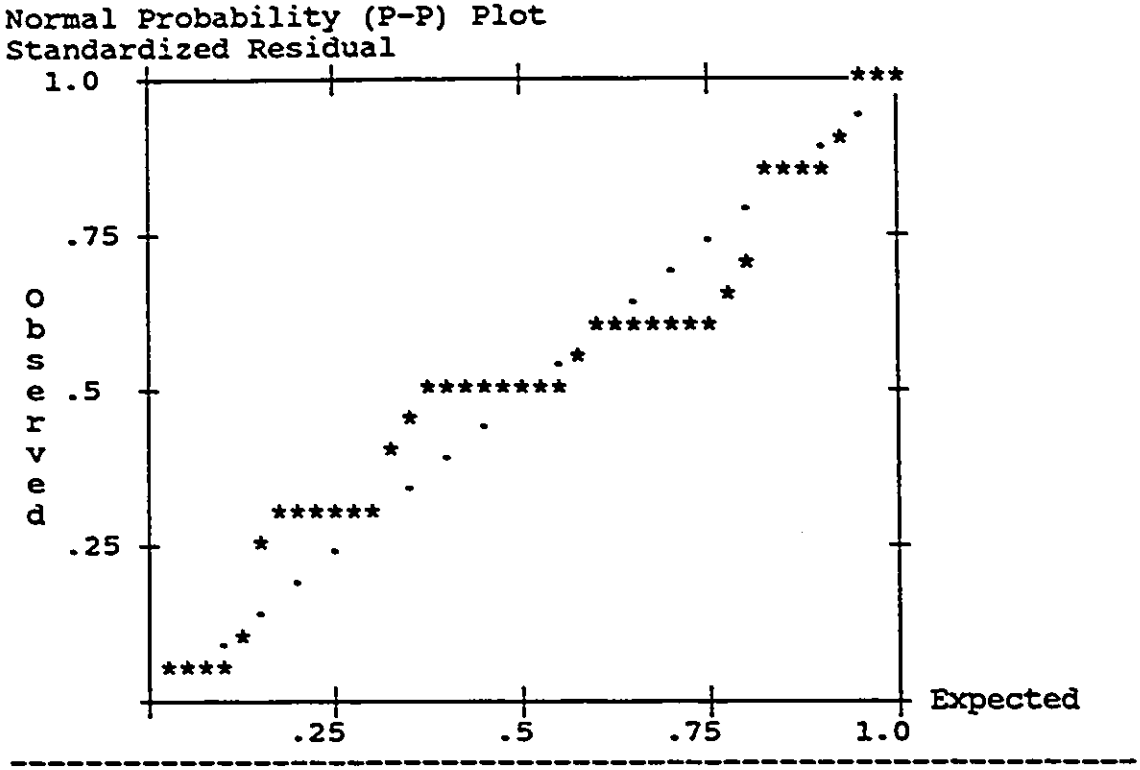
Variable	B	Beta	p-value	Cumulat. Variance
MORBID1	-.07238	-.20956	.0003	.07603
FAMILIAL	.05862	.18379	.0013	
NUTRIDEX	.00199	.09744	.0892	

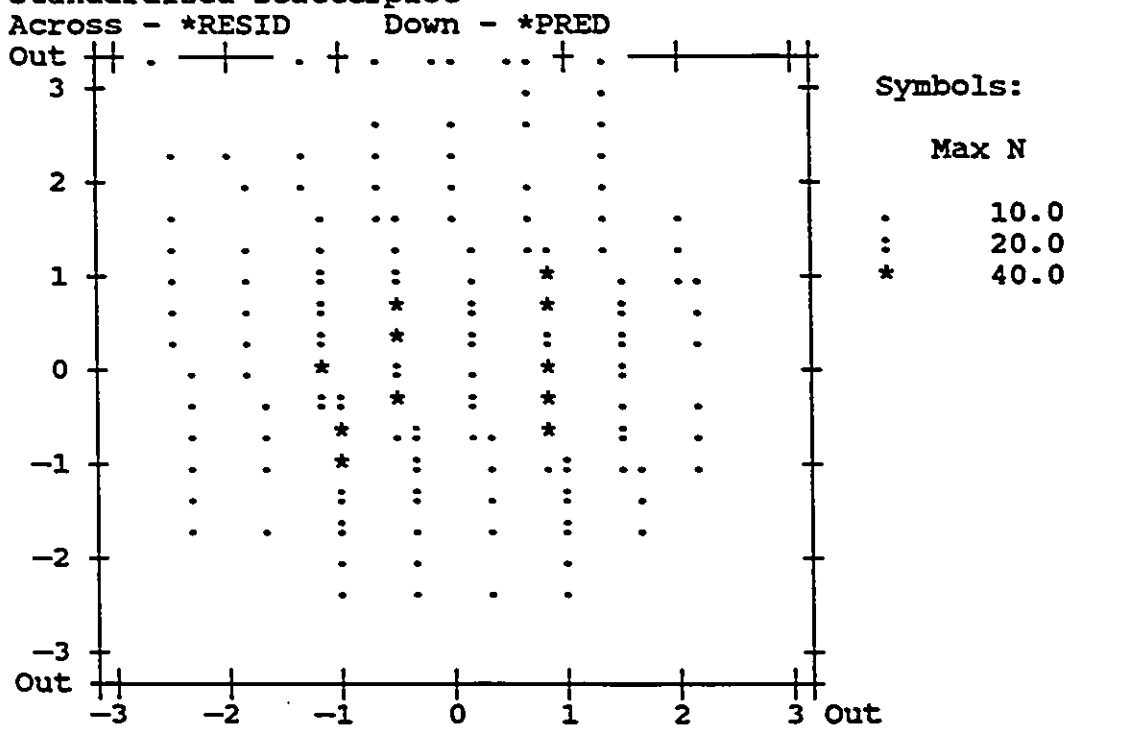
ZWH4:

Variable	B	Beta	p-value	Cumulat. Variance
MORBID1	-.07853	-.18516	.0015	.05321
FAMILIAL	.06098	.15716	.0068	

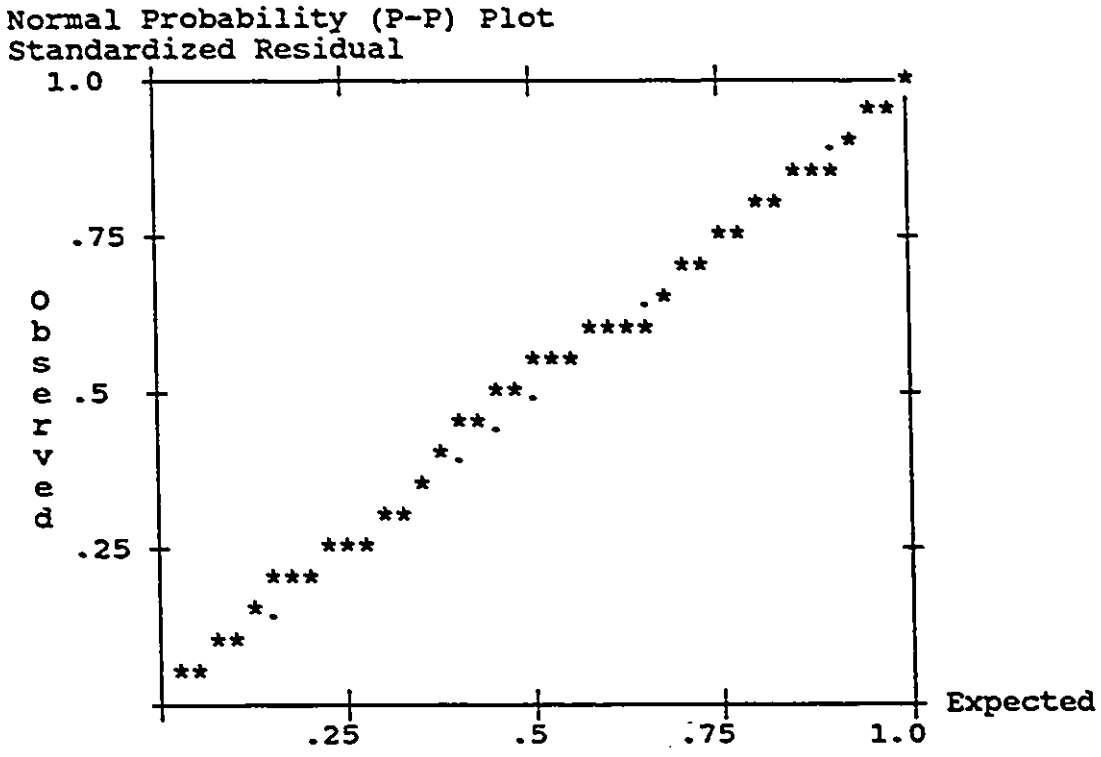
[1] Unstandardized Regression Coefficient
[2] Standardized Regression Coefficient

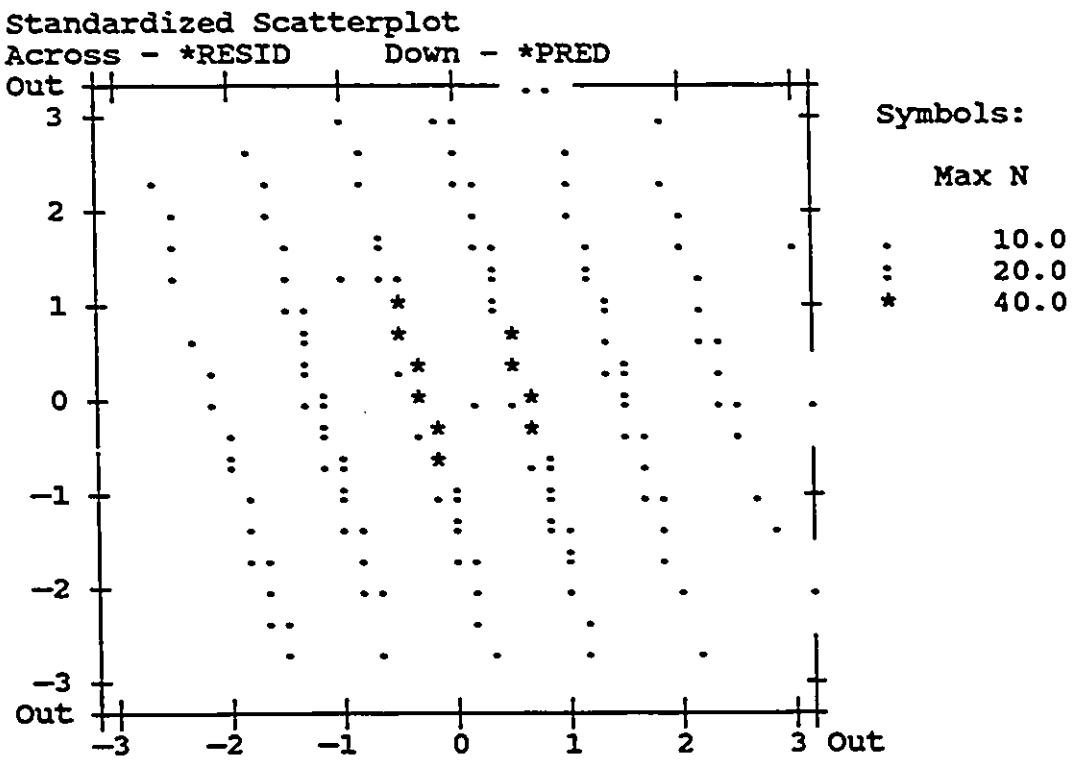
APPENDIX I
Residual Plots of General Regressions
Outcome = Feedindx



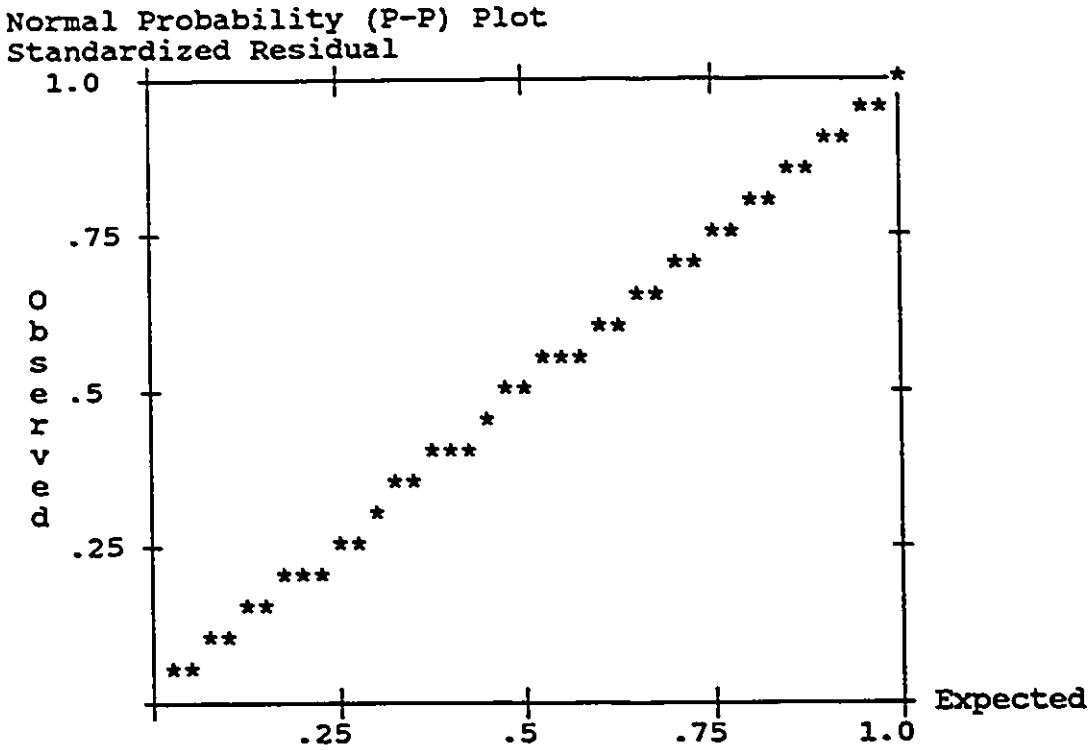


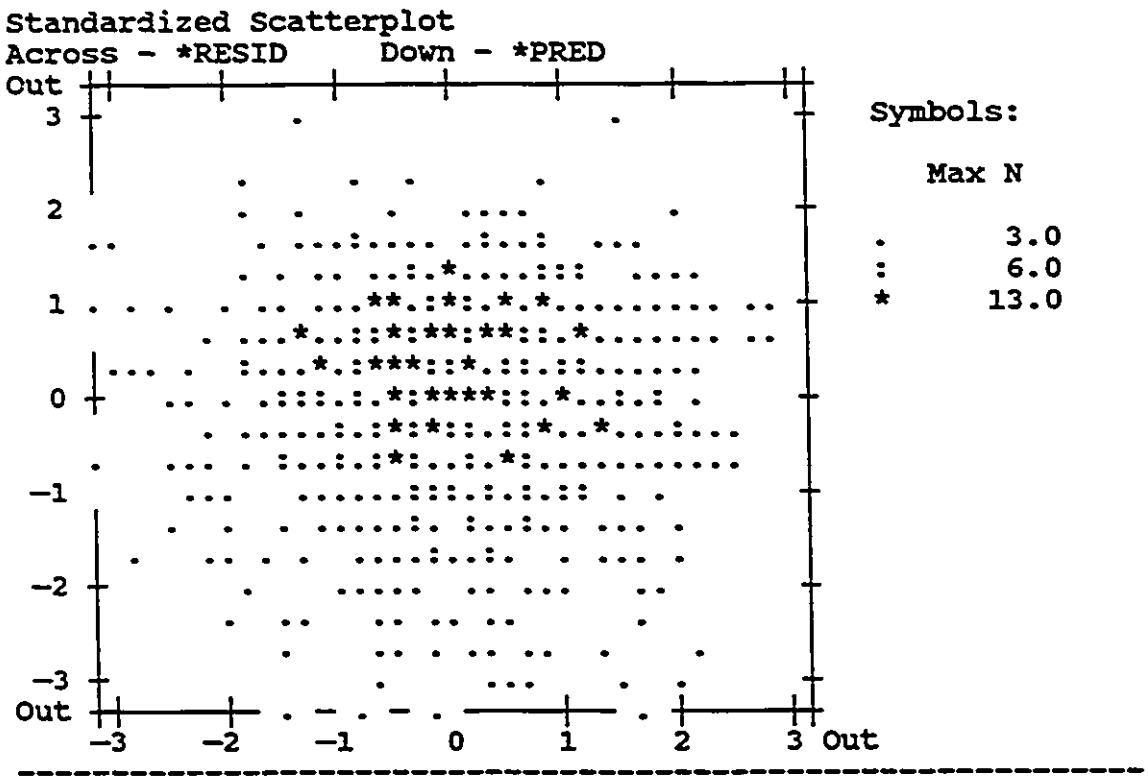
Outcome = Morbid2



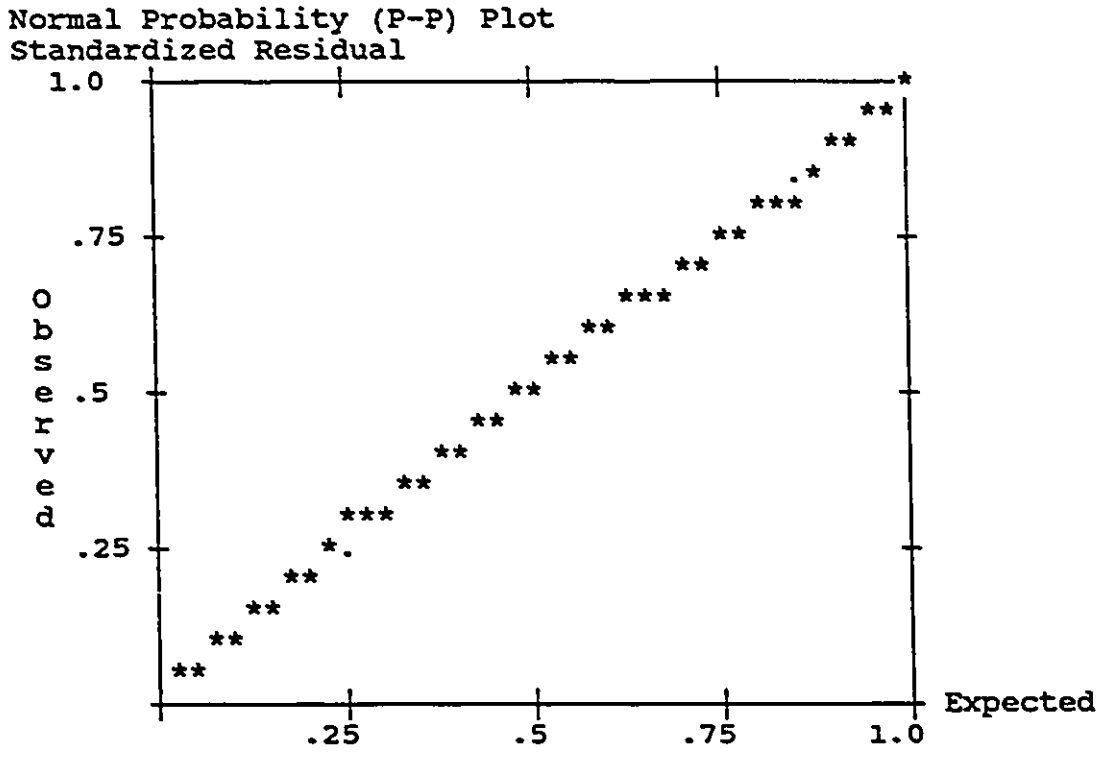


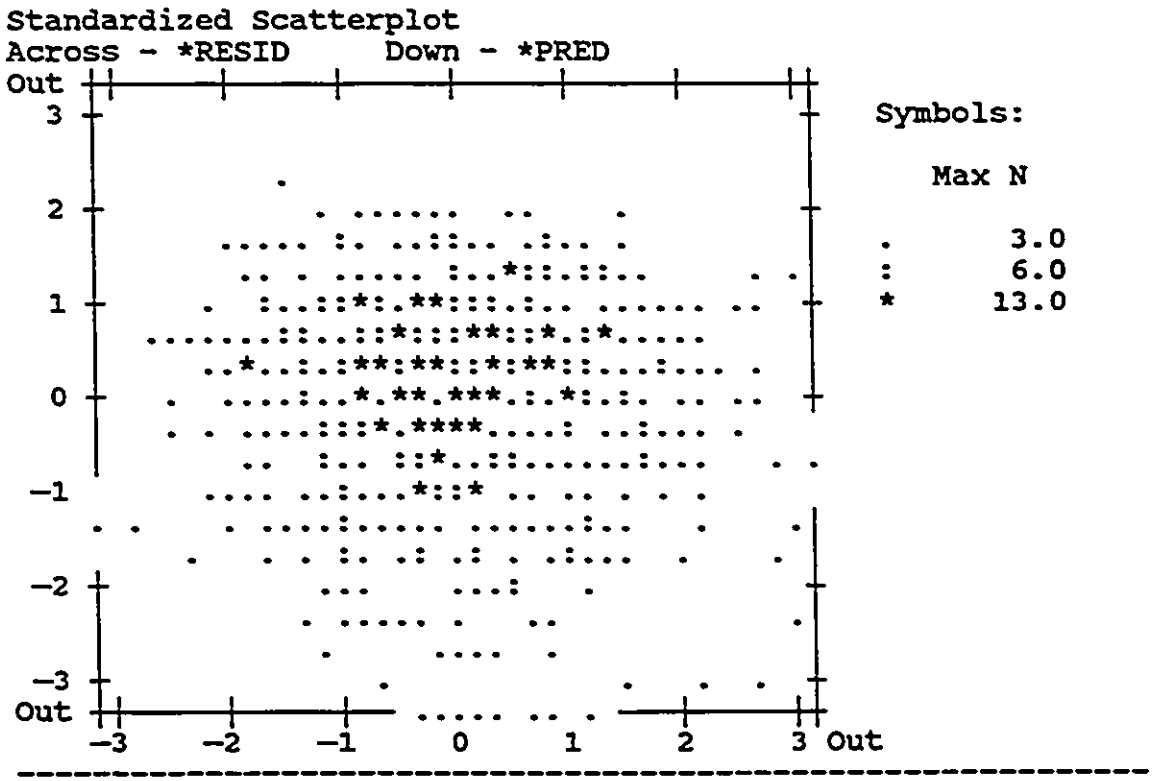
Outcome = ZHA4



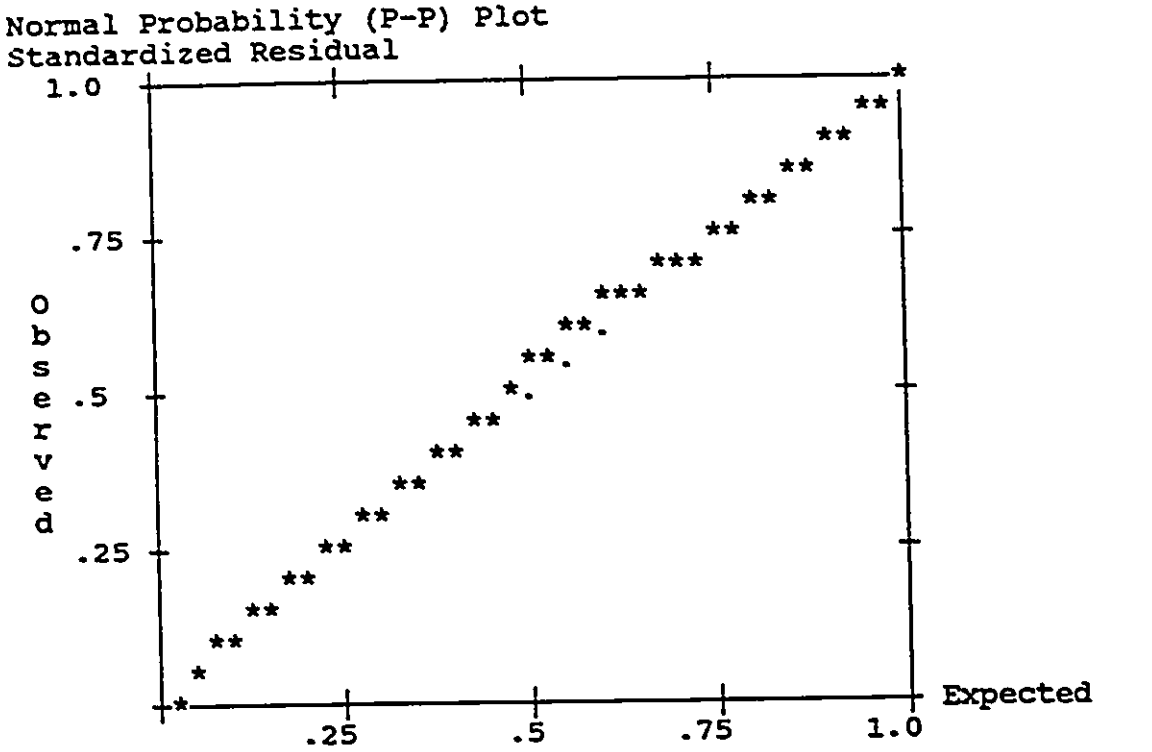


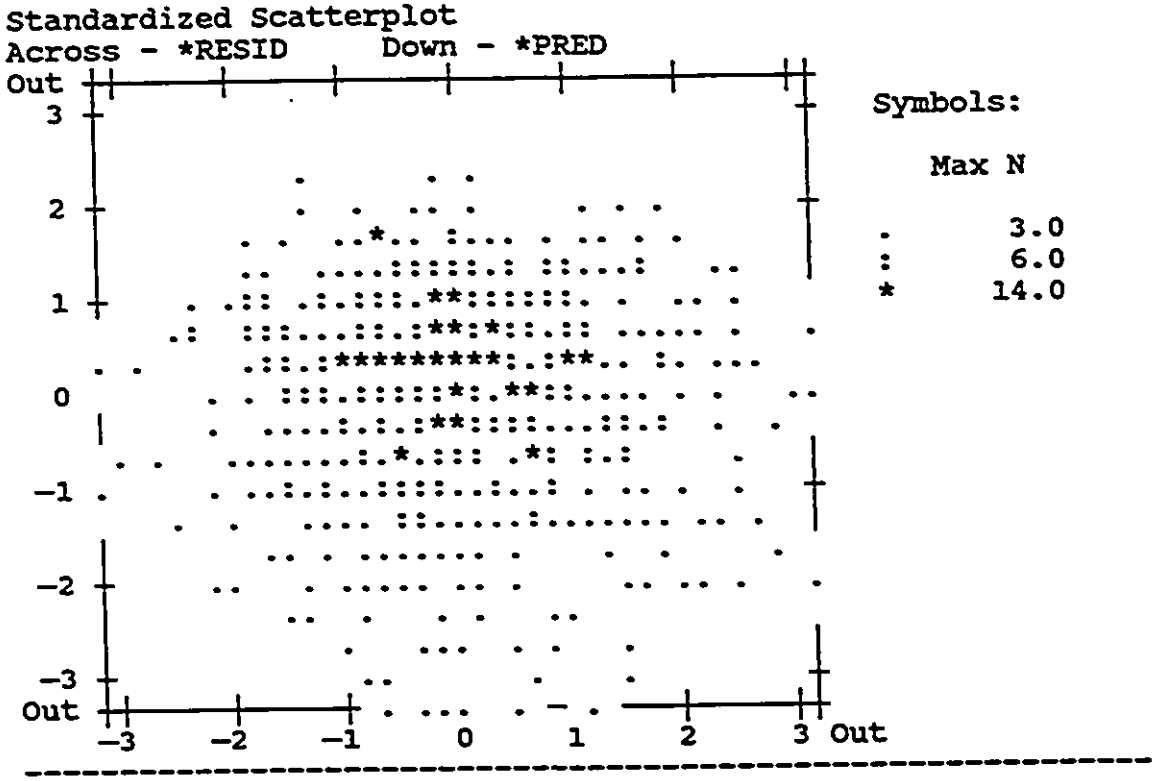
Outcome = ZWA4



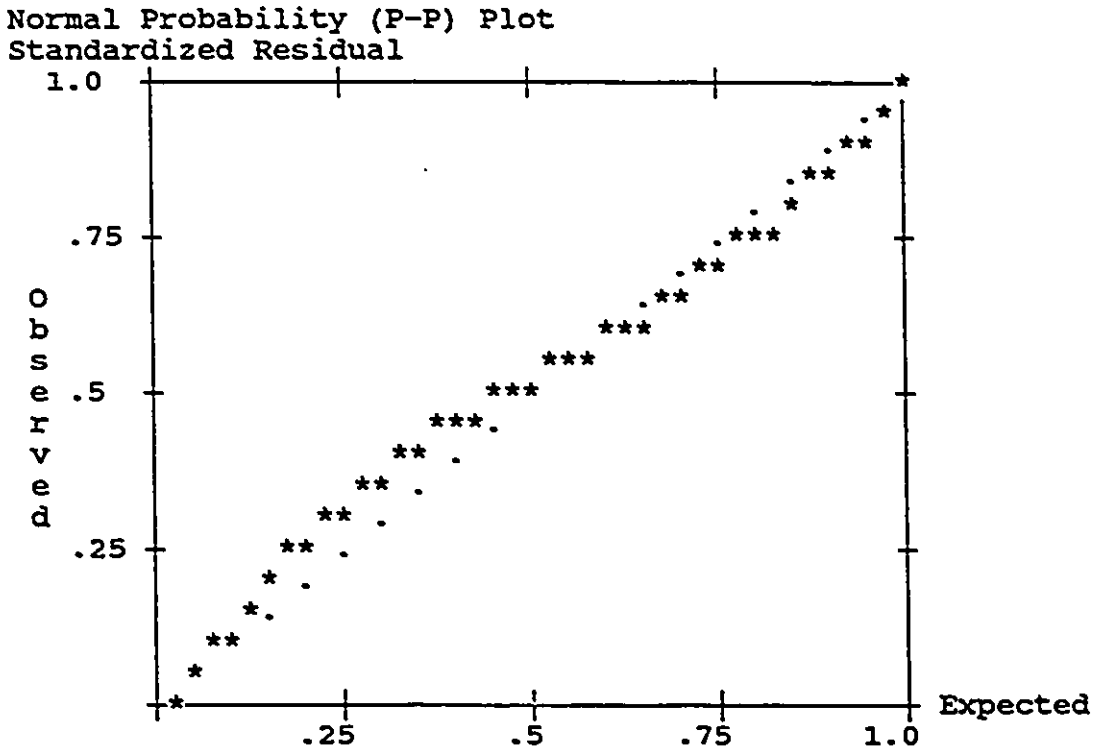


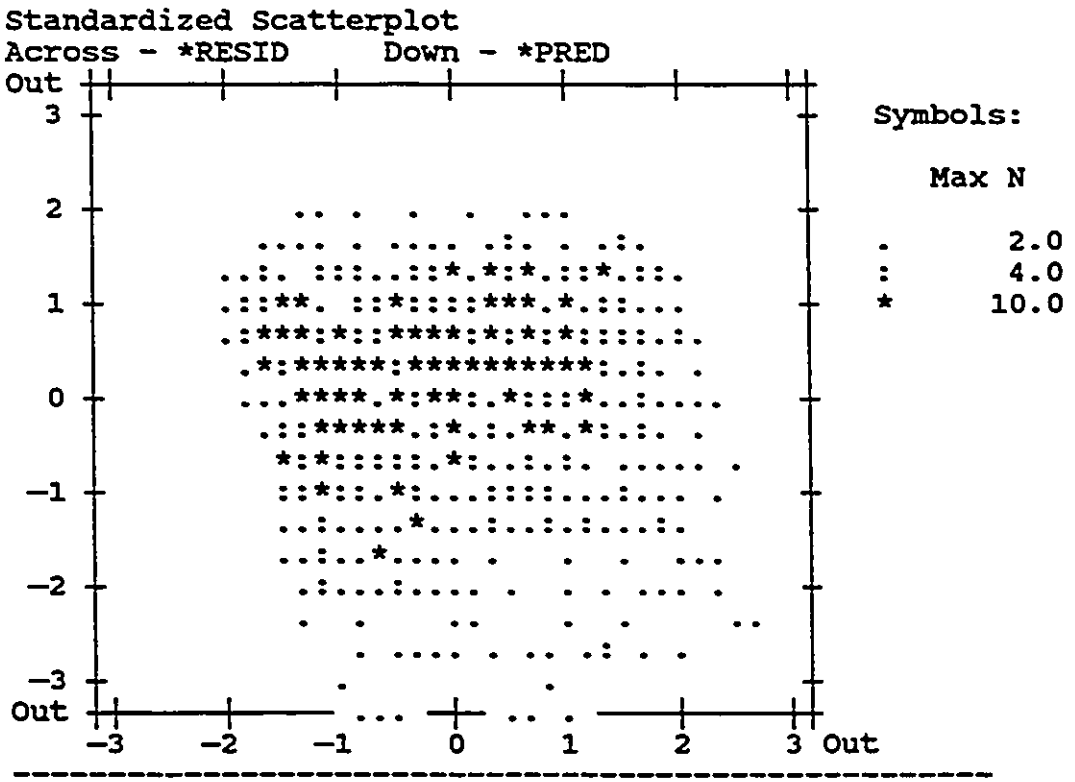
Outcome = ZWH4





Outcome = WHPCT





APPENDIX J

Comparison of Cases Used and Cases Not Included in the Analysis because of Missing Values, on Variables comprising any of the Dependent (growth) or Independent (indexes) Variables.

T-test on Continuous Variables

Variable Used/missing	Number of Cases	Mean[1]	t	DF	prob
Z-score H/A	903	-.0024	-.20	1159	.842
	258	.0100			
Z-score W/A	903	.0179	1.17	1185	.242
	284	-.0526			
Z-score W/H	903	.0637	.91	1159	.364
	258	-.0032			
W/H Pctile	903	43.4617	.54	965	.589
	64	41.6325			
Age at Weaning	903	1.5967	-3.84	976	.000
	192	2.1257			
Age at Time 1	903	22.1245	-2.58	1145	.010
	244	25.5000			

X^2 test on Categorical Variables

Sex	Male	Female	$p(X^2)$
Cases used	51.9%	48.1%	.8956
Cases missing	51.3%	48.7%	
Measles	Had	None	
Cases used	65.0%	35.0%	.7421
Cases missing	65.8%	34.2%	

Prematurity	Singleton	Premature	Twin	
Cases used	95.6%	1.4%	3.0%	.7223
Cases missing	94.7%	2.1%	3.2%	

Mother's Drinking	A lot	A little	Never	
Cases used (903)	7.6%	14.1%	78.3%	.5395
Cases missing (284)	6.4%	16.3%	78.3%	

Clinical Malnutrition	Present	Absent	
Cases used	10.1%	89.9%	.3547
Cases missing	8.0%	92.0%	
Mother's Parity	Low	High	
Cases used	67.7%	32.3%	.3229
Cases missing	71.0%	29.0%	
Mother's Age	Low Risk	High Risk	
Cases used	70.8%	29.2%	.2024
Cases missing	66.5%	33.5%	
Mother's Health	Sick	Well	
Cases used	3.5%	96.5%	.1480
Cases missing	1.8%	98.2%	
Kind of Town	Poor	Better	
Cases used	49.6%	50.4%	.1402
Cases missing	44.4%	55.6%	
Difficult Lactation	Present	Absent	
Cases used	14.7%	85.3%	.0071
Cases missing	10.6%	89.4%	

X^2 test on Categorical Variables

Mother's Education	High	Low	None	$p(X^2)$
Cases used	2.1%	46.5%	51.4%	.0004
Cases missing	3.2%	53.1%	43.7%	
Mother's Income	Some	None		
Cases used	23.9%	76.1%		.0002
Cases missing	35.6%	64.4%		
Diarrhea	Present	Absent		
Cases used	20.8%	79.2%		.0000
Cases missing	8.8%	91.2%		
Mother's other Income	None	Some		
Cases used	88.3%	11.7%		.0000
Cases missing	70.6%	29.4%		

Marital status	Single	Unstable	Stable	
Cases used	6.0%	16.3%	77.7%	.0000
Cases missing	28.8%	16.0%	55.2%	
Father's Drinking	A lot	A little	Never	
Cases used	57.9%	10.2%	31.9%	.0000
Cases missing	37.6%	39.0%	23.4%	
Father's Income	High		Low	
Cases used	71.2%		28.8%	.0000
Cases missing	80.3%		19.7%	
Mother's Allowance	High		Low	
Cases used	79.8%		20.2%	.0000
Cases missing	59.9%		40.1%	

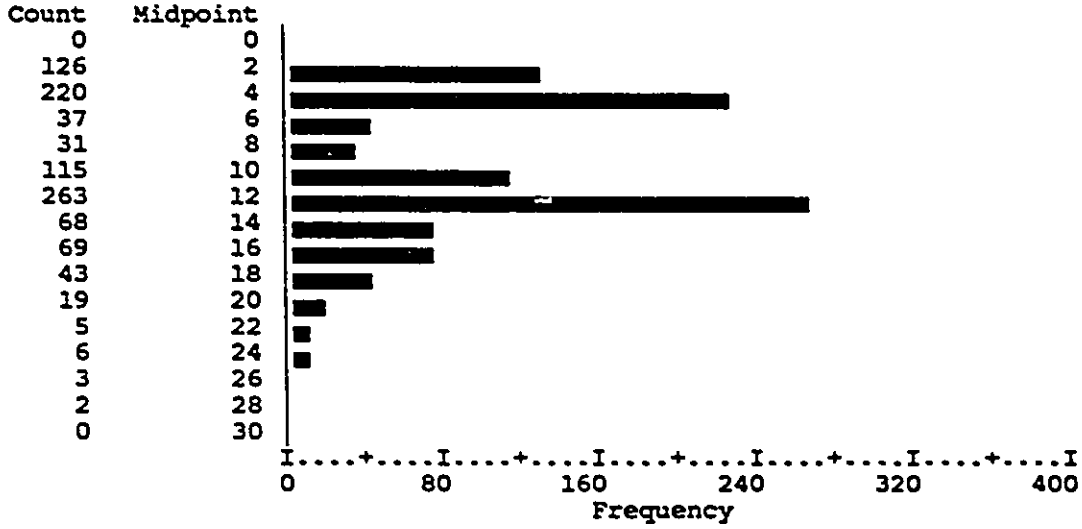
APPENDIX K

Frequencies of Missing Values in Variables of Cases Used in Analyses

(Cased Used = 1007 of the 1342 in original dataset)

VALUE	FREQ	PCT	CUM PCT	VALUE	FREQ	PCT	CUM PCT	VALUE	FREQ	PCT	CUM PCT
1.00	30	3	3	10.00	88	9	53	19.00	9	1	97
2.00	96	10	13	11.00	139	14	66	20.00	10	1	98
3.00	145	14	27	12.00	124	12	79	21.00	5	0	99
4.00	75	7	34	13.00	45	4	83	23.00	3	0	99
5.00	22	2	37	14.00	23	2	85	24.00	3	0	100
6.00	15	1	38	15.00	29	3	88	26.00	3	0	100
7.00	14	1	39	16.00	40	4	92	27.00	1	0	100
8.00	17	2	41	17.00	29	3	95	28.00	1	0	100
9.00	27	3	44	18.00	14	1	97				

Histogram:



Mean	8.945	Std Dev	5.366	Minimum	1.000
Maximum	28.000				

Frequencies of Missing Value in Variables for Cased used in Analyses

Variable Label	Missing (out of total of 1007 cases used)
KIND OF TOWN	0
# IN HOUSE >5 Y.O.	0
# IN HOUSE <5 Y.O.	0
FATHER'S HEALTH	41
MOTHER'S HEALTH	2
MOTHER'S AGE	315
MOTHER'S PARITY	1
DEAD SIB BORN <1966	0
MARITAL STATUS	39
FATHER'S EDUCATION	81
MOTHER'S EDUCATION	15
FATHER'S EMPLOYMENT	59
FATHER'S INCOME	87
MOTHER'S ALLOWANCE	13
GARDEN	7
MOTHER'S EMPLOYMENT	5
MOTHER'S OTH.SUPPORT	5
MOTHER'S OTH.INCOME	70
FATHER'S OH-INTAKE	89
MOTHER'S OH-INTAKE	29
NO. DIARIES DONE	559
MEAT DAYS	566
FISH DAYS	566
BEANS DAYS	566
STAPLE DAYS	566
MILK DAYS	566
FOOD EXPENSE	573
SOURCE OF WATER	569
RANK OF CHILD 1	7
SEX	0
PREMATURITY	0
DATE OF BIRTH	0
ACCURACY OF D.O.B.	0
OCCURENCE OF DEATH	0
NO. YOUNGER SIBS	0
BIRTH PLACE	1
WHO LIVING WITH	0
DIFF. LACTATION	123
AGE AT WEANING	136
AGE BOTTLE STARTED	117
AGE STARTED PORRIDGE	188
AGE STARTED NSIMA	221
HOSPITALIZATIONS	7

Variable Label	Missing
MALNUTRITION	14
CNS-PROBLEMS	14
ENT-PROBLEMS	3
SKIN-PROBLEMS	2
RESP-PROBLEMS	3
EYES-PROBLEMS	3
BLOOD-PROBLEMS	3
GI-PARISITES	3
GI-ENTERITIS	3
SURGERY	3
RX-WITCHDOCTOR	18
RX-HOME	15
MOTHER PREGNANT? 1	5
BREAST FEEDING? 1	14
BOTTLE FEEDING? 1	14
TYPE OF FOOD 1	26
SYMPTOMS OF DIARRHEA 1	1
OTHER SYMPTOMS 1	3
MOTHER PREGNANT? 4	9
BREAST FEEDING? 4	9
BOTTLE FEEDING? 4	11
TYPE OF FOOD 4	12
SYMPTOMS OF DIARRHEA 4	0
OTHER SYMPTOMS 4	0
MEASLES OCCURENCE	18
<5 CLINIC VISITS	5
MEDICAL CHART?	3
SMALLPOX VACCINE	13
DPT VACCINE	4
POLIO VACCINE	4
BCG VACCINE	2
AGE AT 1ST MEASURE	0
AGE AT 4TH MEASURE	0

APPENDIX L

Comparison of Cases with 5 or less missing variables with the rest of cases [1]

T-test on Continuous Variables

Variable (5</>5)	Number of Cases	Mean	t	df	prob.
Mother's age	276 416	27.26 27.51	-0.50	690	.620
# older sibs	368 639	4.20 4.46	-1.99	1005	.047
# younger sibs	368 639	2.46 2.52	-1.02	1005	.308
Mother's Parity	368 638	5.40 5.27	0.74	1004	.462
Father's Income	365 555	55.25 56.40	-0.75	918	.452
Food Expense	365 69	69.20 62.64	1.94	432	.053
Rank of Child	368 632	4.66 4.52	0.80	998	.426
Age at weaning	367 504	1.59 1.58	0.28	869	.778
Age start bottle	366 524	.20 .21	-0.77	888	.439
Age start porridge	357 462	.46 .43	1.40	817	.162
Age start nsima	348 438	.73 .75	-0.97	784	.333
Age at Time 1	368 639	20.45 22.65	-2.02	1005	.043
Weight /Kg at Time 4	368 639	11.89 12.15	-1.21	1005	.228
Length /cm at Time 4	363 613	85.22 86.93	-2.01	974	.044
Age at Time 4	368 639	32.87 33.92	-0.944	1005	.349

---

[1] Because of the multiple comparisons made, only a p-value of about 0.001 or less guarantees a true probability of 0.05.

NB: p-value of 0.000 means <0.001

Comparison of Cases with 5 or less missing variables with the rest of cases
X^2 test on Categorical Variables

Variable (5</>5)	Categories			$p(X_2)$
Town	Poor	Better		
5< missing	59.2%	40.8%		.000 [1]
>5 missing	43.0%	57.0%		
Father's health	Sick	Well		
5< missing	0.8%	99.2%		.024
>5 missing	2.8%	97.2%		
Mother's health	Sick	Well		
5< missing	4.1%	95.9%		.268
>5 missing	3.0%	97.0%		
Marital status	Single	Unstable	Stable	
5< missing	5.2%	19.1%	75.7%	.000
>5 missing	13.8%	14.2%	72.0%	
Father's education	High	Low	None	
5< missing	10.3%	72.0%	17.7%	.005
>5 missing	11.5%	62.3%	26.2%	
Mother's education	High	Low	None	
5< missing	1.6%	47.2%	50.4%	.005
>5 missing	2.4%	47.2%	50.4%	
Father's employment[2]				.220
Mother's employment[2]				.000
Father's OH intake	A lot	A little	Never	
5< missing	60.1%	9.5%	30.4%	.146
>5 missing	58.4%	7.6%	34.0%	
Mother's OH intake	A lot	A little	Never	
5< missing	14.0%	6.8%	79.2%	.052
>5 missing	13.4%	9.4%	77.2%	
# Meat days[2]				.000
# Milk days[2]				.000
Source of water	Tap	Lorry	Well	
5< missing	49.6%	3.0%	47.4%	.331
>5 missing	43.7%	1.4%	54.9%	

Sex	Male	Female		
5< missing	52.7%	47.3%		.971
>5 missing	50.8%	49.2%		
Prematurity	Singleton	Premature	Twin	
5< missing	93.5%	2.7%	3.8%	.118
>5 missing	93.4%	2.3%	4.3%	
Birthplace[2]				.604
Difficult lactation	Yes	No		
5< missing	18.5%	81.5%		.293
>5 missing	16.2%	83.8%		
Malnutrition	Yes	Never		
5< missing	12.8%	87.2%		.063
>5 missing	8.7%	91.3%		
Respir. problems	Yes	No		
5< missing	14.1%	85.9%		.110
>5 missing	7.9%	92.1%		
Gastro-enteritis	Yes	No		
5< missing	28.3%	71.7%		.000
>5 missing	12.8%	87.2%		
Rx-Witchdoctor	Yes	No		
5< missing	8.8%	91.2%		.035
>5 missing	5.8%	94.2%		
Rx-Home	Yes	No		
5< missing	20.8%	79.2%		.014
>5 missing	13.9%	86.1%		
Breastfeeding	Yes	No		
5< missing	45.1%	54.9%		.079
>5 missing	40.8%	59.2%		
Bottlefeeding	Yes	No		
5> missing	15.2%	84.8%		.002
>5 missing	9.3%	90.7%		
Type of food[2]				.052
Diarrhea	Yes	No	With Measles	
5< missing	80.3%	15.3%	4.4%	.216
>5 missing	85.4%	11.7%	2.9%	
Measles				.048

Clinic Visits	None	Some			
5< missing	68.2%	31.8%			.067
>5 missing	74.1%	25.9%			
DPT vaccination	None	One	Two	Three	
5< missing	80.7%	12.0%	1.9%	5.4%.	.024
>5 missing	88.4%	5.6%	2.5%	3.5%	
Polio vaccination	None	One	Two	Three	
5< missing	77.4%	12.0%	4.9%	5.7%	.274
>5 missing	83.3%	9.3%	3.7%	3.7%	
Breastfeeding T4	Yes	No			
5< missing	20.7%	79.3%			.363
>5 missing	18.1%	81.9%			
Bottlefeeding T4	Yes	No			
5< missing	6.3%	93.7%			.497
>5 missing	7.6%	92.4%			
Type of food T4					.747
Diarrhea T4	Yes	No	With Measles		
5< missing	13.0%	85.1%	1.9%		.326
>5 missing	10.0%	88.3%	1.7%		

---

[1] p-value of 0.000 means 0.001
[2] Too many categories to list

APPENDIX M
Actual computer codes for Creation of Some Indexes [1]

```
COUNT MORBIDEX = HOSPEVER CNSEVER CNSEVER ENT SKIN RESP RESP
EYES BLOOD BLOOD  GIA GIA GIB GIB SURGEVER SURGEVER
NGANGA RXHOME MEASLES DIFFLACT (1) PREM (1,2).

COMPUTE PREVINDX = (U5C/2) + (SPX * 2) + (DPT * 2)
+ (POLIO * 2) + (BCG * 2).

COMPUTE NUTRIDEX = ((MEAT * 2) + (FISH * 2) + (BEANS * 1.5)
+ STAPLES + (MILK * 2)) + GARDEN + COST.

COMPUTE FAMILIAL = (TOWN + NOLD + NYOUNG + (HEALTHPA * 2)
+ (HEALTHMA * 2) + AGEMA + PARITY + DEATHS + (MARRIAGE * 2)
+ (EDPA * .3) + (EDMA * .3) + DRINKPA + DRINKMA + WATER
+ HANDICAP + (YOUNGSIB * .5) + BORN + WHOWITH).

COMPUTE ECONDX = ((JOBPA/2) + (JOBMA/2) + INCOME + ALLOWS
+ MONEY + MOTHINC).
```

---

[1] Other non-relevant codes have been omitted for clarification.

APPENDIX N

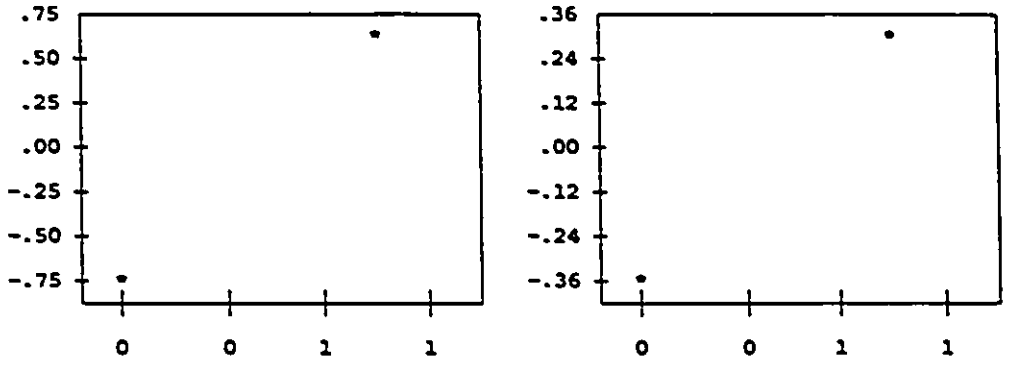
Distributions of the index variables used in discriminant analysis, by group (alive, dead).

SEX

Valid cases:	274.0	Missing cases:	.0	Percent missing:	.0		
Mean	.4854	Std Err	.0302	Min	.0000	Skewness	.0587
Median	.0000	Variance	.2507	Max	1.0000	S E Skew	.1472
5% Trim	.4838	Std Dev	.5007	Range	1.0000	Kurtosis	-2.0113
				IQR	1.0000	S E Kurt	.2933

Frequency	Bin Center	
141.00	1	**********
133.00	2	**********

Bin width : 1
Each star: 3 case(s)



Normal Plot Detrended Normal Plot

K-S (Lilliefors)

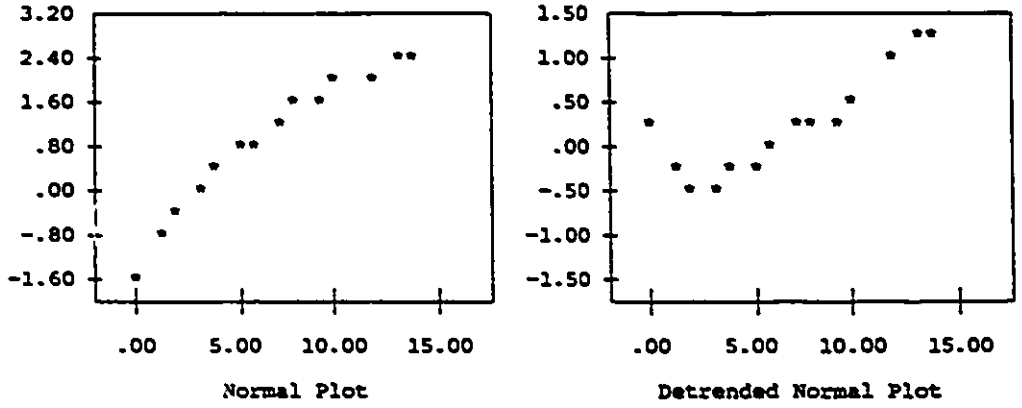
Group	Statistic	df	Significance
Whole sample	.3484	274	.0000
Alive	.3480	136	.0000
Dead	.3479	138	.0000

MORBIDEX

Valid cases:	274.0	Missing cases:	.0	Percent missing:	.0		
Mean	2.8139	Std Err	.1631	Min	.0000	Skewness	1.4089
Median	2.0000	Variance	7.2876	Max	14.0000	S E Skew	.1472
5% Trim	2.5685	Std Dev	2.6996	Range	14.0000	Kurtosis	2.0237
				IQR	3.0000	S E Kurt	.2933

Frequency	Bin Center	
42.00	.50	**********
73.00	1.50	**********
43.00	2.50	**********
36.00	3.50	*********
22.00	4.50	******
14.00	5.50	*******
9.00	6.50	****
16.00	7.50	********
10.00	8.50	*****
9.00	Extremes	****

Bin width : 1.00
Each star: 2 case(s)



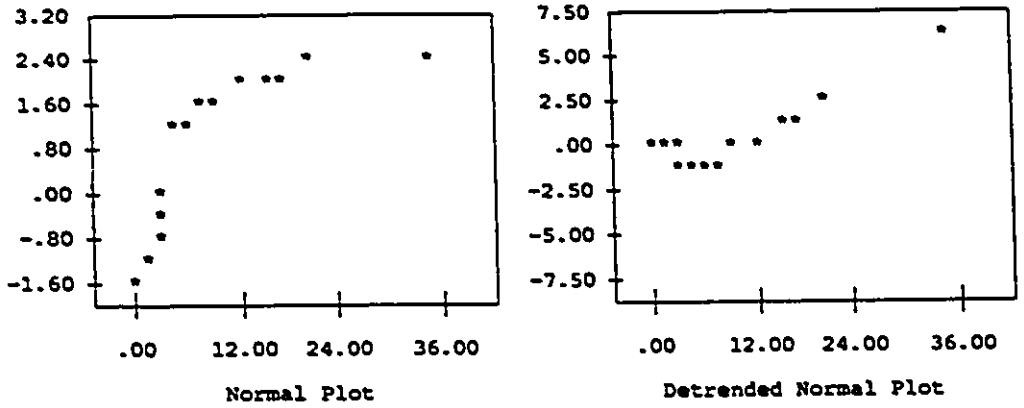
	K-S (Lilliefors)		
Group	Statistic	df	Significance
Whole sample	.1916	274	.0000
Alive	.1481	136	.0000
Dead	.2087	138	.0000

PREVINDX

Valid cases:	274.0	Missing cases:	.0	Percent missing:	.0		
Mean	3.0052	Std Err	.1894	Min	.0000	Skewness	5.2479
Median	3.1685	Variance	9.8311	Max	34.0000	S E Skew	.1472
5% Trim	2.6030	Std Dev	3.1355	Range	34.0000	Kurtosis	41.0839
				IQR	1.1685	S E Kurt	.2933

Frequency	Bin Center	
40.00	Extremes	**********
2.00	.50	*
2.00	1.50	*
72.00	2.50	**********
125.00	3.50	**********
13.00	4.50	****
20.00	Extremes	*******

Bin width : 1.00
Each star: 3 case(s)



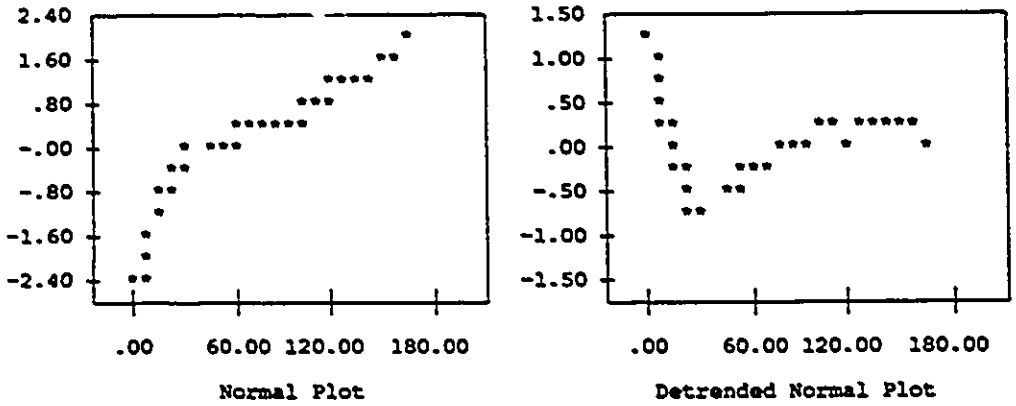
	K-S (Lilliefors)		
Group	Statistic	df	Significance
Whole sample	.3588	274	.0000
Alive	.2957	136	.0000
Dead	.4436	138	.0000

NUTRIDEX

Valid cases:	117.0	Missing cases:	157.0	Percent missing:	57.3		
Mean	51.4274	Std Err	4.2482	Min	1.0000	Skewness	.8967
Median	26.0000	Variance	2111.510	Max	162.5000	S E Skew	.2236
5% Trim	48.2754	Std Dev	45.9512	Range	161.5000	Kurtosis	-.5477
				IQR	73.5000	S E Kurt	.4437

Frequency	Bin Center	
91.00	50.00	**********
26.00	150.00	**********

Bin width : 100.00
Each star: 2 case(s)



	K-S (Lilliefors)		
Group	Statistic	df	Significance
Whole sample	.2450	117	.0000
Alive	.2059	59	.0000
Dead	.3004	58	.0000

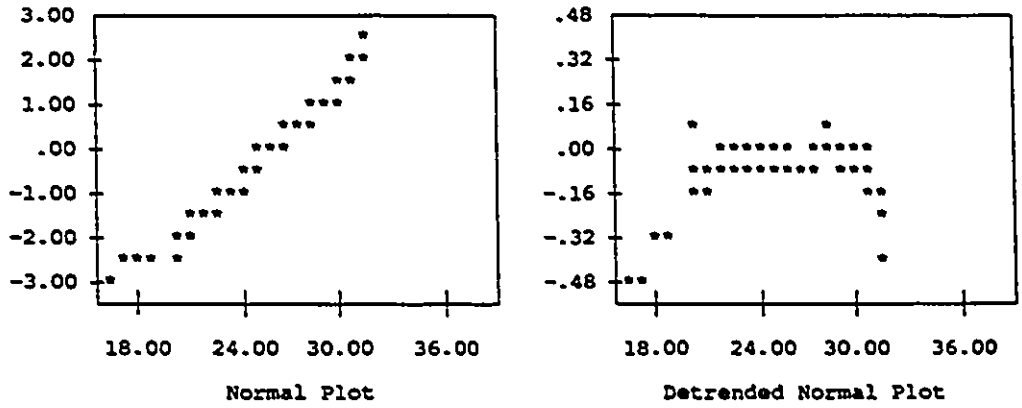
SOCSANDX

Valid cases:	274.0	Missing cases:	.0	Percent missing:	.0		
Mean	25.1307	Std Err	.1681	Min	16.5000	Skewness	-.1460
Median	25.1000	Variance	7.7465	Max	31.7000	S E Skew	.1472
5% Trim	25.1611	Std Dev	2.7833	Range	15.2000	Kurtosis	-.0387
				IQR	3.6250	S E Kurt	.2933

```
Frequency   Bin Center

    2.00  Extremes
    4.00     15.00   *
  258.00     25.00   **********************************************
   10.00     35.00   **

Bin width :     10.00
Each star:       6 case(s)
```



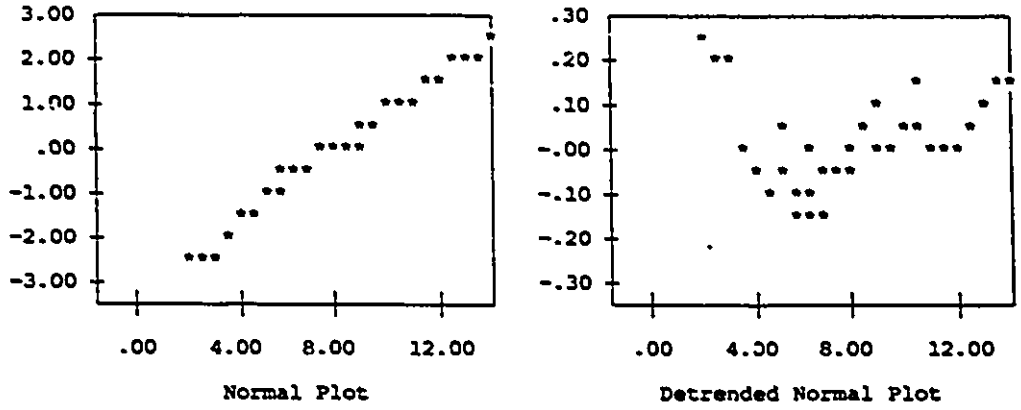
	K-S (Lilliefors)		
Group	Statistic	df	Significance
Whole sample	.0306	274	> .2000
Alive	.0469	136	> .2000
Dead	.0410	138	> .2000

ECONDX

Valid cases:	274.0	Missing cases:	.0	Percent missing:	.0		
Mean	7.3597	Std Err	.1488	Min	2.0000	Skewness	.1719
Median	7.5000	Variance	6.0663	Max	14.0000	S E Skew	.1472
5% Trim	7.3218	Std Dev	2.4630	Range	12.0000	Kurtosis	-.4655
				IQR	3.5000	S E Kurt	.2933

Frequency	Bin Center	
228.00	5.00	**********
46.00	15.00	*********

Bin width : 10.00
Each star: 5 case(s)



	K-S (Lilliefors)		
Group	Statistic	df	Significance
Whole sample	.0851	274	.0001
Alive	.1082	136	.0005
Dead	.0743	138	.0593

ABSTRACT

DETERMINANTS OF GROWTH FAILURE
IN A GROUP OF ZAMBIAN CHILDREN

by Denis G. Allard

Introduction: Many predictors of growth failure and of mortality in children have been identified. Research has not provided so far an accurate perception of the relative importance of each determinant and of the intricate interaction of all these factors. The present study attempts to accomplish this, using a set of data obtained from African children.

The Main Objectives of the study were:

1) To develop adequate indicators of growth, using anthropometric data.
2) To specify a theoretical model of the process whereby predictive factors influence growth.
3) To develop composite indexes of the social, economic, and environmental variables.
4) To calculate path coefficients for factors specified in predictive growth models and thereby to test the models.
5) To refine the models and thereby help to plan interventions that would prevent growth failure and reduce premature mortality.

Methods:

STUDY DESIGN: The analysis used data collected in 1971-72 for a longitudinal study of factors affecting the growth of a cohort of 1342 Zambian preschoolers, aged 1 month to 5 years. Extensive baseline information was collected on each household and on each participating child during the initial interview.

At each of four bi-annual follow-up visits, the children were weighed and measured; dietary and morbidity histories were taken.

DEPENDENT VARIABLES: Three indicator values of weight and length were calculated for each child. These are **'Height-for-Age'**, **'Weight-for-Age'**, and **'Weight-for-Height'**. These were expressed both as percentiles of the U.S. NCHS standard growth curves by sex and as **Z-scores** within the sample. Only the Z-scores were eventually used, approximating more closely a normal distribution than the U.S.-derived measures. Mortality during the study also was examined and was coded as a dichotomous outcome indicator.

PREDICTOR VARIABLES: Seven indexes were formed to represent: feeding practices, current morbidity, prior morbidity, food availability, prevention practice, familial circumstances, and economic status. Feeding practices and current morbidity were used also as intermediate outcomes in testing the predictive models.

ANALYSES: **Multiple linear regression** and **path analysis** were used to test the proposed growth models and to study interactions between the components. **Discriminant analysis** was used for the model of predictors of mortality, seeking to differentiate between the survivors and the children who died. This was supplemented by univariate methods such as comparison of means and comparisons of proportions. The goodness of fit of the growth models was tested on various subgroups of the general sample, including the malnourished, and on different age groups.

Results:

a) PREDICTING FEEDING PRACTICES: There was no consistent determinant of feeding practices across all of the various groups studied.

b) PREDICTING CURRENT MORBIDITY: In general, feeding practices and prior morbidity correlated positively with current morbidity, except in children less than 6 months old and in the underweight sub-group. Familial circumstances and food availability correlated negatively with current morbidity in all groups except the malnourished.

c) PREDICTING GROWTH INDICATORS: Prior morbidity was consistently a negative predictor except for the height in the stunted group, where feeding practices were the only retained predictive factor. Familial circumstances and/or economic status were also important positive determinants of growth except among children aged less than 6 months and in the wasted sub-group. Food availability was a positive predictor of weight-for-age in those more than 6 months of age and in the general sample in which it also predicted height-for-age. As a rule, the percentage of variance explained by the chosen predictors was low but improved (reached 12%) when subgroups of the general sample were analysed. Contrary to expectations and as demonstrated by path analysis most of the effects of prior morbidity, familial circumstances, economic status, prevention, and food availability on the growth indicators were direct and did not appear to operate through feeding practices or current morbidity, except in the malnourished sub-groups.

d) PREDICTING MORTALITY: The degree of prediction of mortality achieved by the derived discriminant equation was not impressive, correctly classifying only 75% of those dying and 60% of the survivors. Prior morbidity was the most

important predictor of likelihood of death among the children, but was inversely correlated. This was followed by the economic status, also negatively correlated. Prevention was in third place, with a positive correlation and Familial Circumstances in fourth place, with a positive correlation. These findings are, for the most part, difficult to interpret, and cast uncertainty on the usefulness of the discriminant ability of the equation. Univariate analysis results confirmed the important association with early death, documented by previous research, of some of the components of morbidity, namely prematurity and twin birth.

Discussion: Specific comments are made on the results of each of the groups analysed as they are presented. A more general technical discussion follows at the end. The potential bias caused by missing cases on internal and external validity is explored and the effect of collinearity of independent variables on their risk of exclusion from the model is discussed. Methodological and conceptual reasons for the low proportion of explained variance achieved are proposed.

Conclusion: Notwithstanding the many caveats expressed in the discussion, the measures that emerge as probably the most useful to improve the growth of most Zambian children (healthy and malnourished, very young and older ones alike) are, in order of importance: 1) to reduce morbidity by treating and preventing childhood diseases; 2) to implement social measures that will improve the education and the health of the parents; 3) to create new economic incentives, attempting to improve indirectly environmental conditions of the household. 4) to provide the financial means to buy an adequate quantity and quality of food, and to access needed competent medical care.

EXTRAIT

DETERMINANTS DE DEFAUT DE CROISSANCE CHEZ UN GROUPE D'ENFANTS ZAMBIENS

par Denis G. Allard

Introduction: Plusieurs prédicteurs de défaut de croissance et de mortalité ont été identifiés chez les enfants. La recherche n'a pas fourni jusqu'à présent une perception juste de l'importance relative de chaque déterminant et de l'interaction compliquée de tous ces facteurs. La présente étude essaie d'accomplir ceci, utilisant pour y arriver un groupe de données provenant d'enfants africains.

Les objectifs principaux de l'étude sont:

1) Développer des indicateurs de croissance convenables, utilisant des données anthropométriques.
2) Spécifier un modèle théorique du processus par lequel les facteurs déterminants influence la croissance.
3) Développer des indexs composés des variables sociales, économiques, et environmentales.
4) Calculer les coefficients de sentier pour les facteurs spécifiés dans les modèles prédicteurs de croissance et vérifier ainsi l'exactitude de ces modèles.
5) Raffiner les modèles et ainsi contribuer à planifier des interventions qui pourraient aider à prévenir le défaut de croissance et à réduire la mortalité prématurée.

Méthodes:

DEVIS DE L'ETUDE: L'analyse a utilisé des données recueillies en 1971-72 pour une étude longitudinale des facteurs affectant la croissance d'une cohorte de 1342 Zambiens pré-scolaires, âgés de 1 mois à 5 ans. Beaucoup d'information de base fut

recueillie sur chaque ménage et sur chaque enfant participant lors de l'entrevue initiale. A chacune de quatre visites de suivi semi-annuelles, les enfants furent pesés et mesurés; leeurs anamnèses nutritionelles et morbides furent notées.

VARIABLES DEPENDANTES: Trois valeurs indicatrices du poids et de la taille furent calculées pour chaque enfant. Elles sont: **'Taille-pour-l'Age'**, **'Poids-pour-l'Age'**, et **'Poids-pour-Taille'**. Celles-ci furent exprimées et comme pourcentiles des courbes de croissance du standard américain NCHS par gendre, et par des **pointages-Z** provenant de l'échantillon-même. Seulement les **pointages-Z** furent utilisés en fin de compte, se rapprochant ainsi plus près d'une distribution normale que le faisaient les mesures dérivées des standards américains. La mortalité survenant durant l'étude fut aussi examinée et fut codée comme un indicateur dichotome.

VARIABLES DETERMINANTES: Sept indexs furent formés pour représenter: les pratiques alimentaires, la morbidité courante, la morbidité passée, la disponibilité de la nourriture, les pratiques préventives, les circonstances familiales, et le statut économique. Les pratiques alimentaires et la morbidité courante furent utilisées aussi comme bornes intermédiaires lors de la vérification des modèles prédicteurs.

ANALYSES: La **régression linéaire multiple** et l'**analyse de sentier** sont les tests qui furent utilisés pour vérifier les modèles de croissance proposés et étudier les interactions de leurs composants. L'**analyse discriminatoire** servit pour le modèle des prédicteurs de mortalité, essayant ainsi de différencier entre les survivants et les enfants qui

moururent. Ceci fut complété par les méthodes d'analyse univariées comme la comparaison des moyennes et la comparaison des proportions. L'exactitude des modèles de croissance fut vérifié sur des sous-groupes variés de l'échantillon général (total), incluant les sous-alimentés, et sur des groupes d'âges différents.

Résultats:

a) PREDIRE LES PRATIQUES ALIMENTAIRES: On n'a pas trouvé de déterminant constant des pratiques alimentaires parmi tous les groupes étudiés.

b) PREDIRE LA MORBIDITE COURANTE: En général, les pratiques alimentaires et de morbidité passée étaient reliées positivement avec la morbidité courante à l'exception des enfants de moins de 6 mois et du sous-groupe dont le poids était sous la normale. Les circonstances familiales et la disponibilité de la nourriture eux se sont avérées négativement reliées à la morbidité courante dans tous les groupes exceptés les sous-alimentés.

c) PREDIRE LES INDICATEURS DE CROISSANCE: La morbidité passée fut un prédicteur négatif dans tous les cas sauf pour la taille du groupe des rabougris, où les pratiques alimentaires furent le seul facteur déterminant retenu. Les circonstances familiales et/ou le statut économique étaient aussi des prédicteurs important à l'exception des enfants âgés de moins de 6 mois et dans le groupe maigre. La disponibilité de la nourriture fut un prédicteur positif du poids-pour-âge chez les âgés de plus de 6 mois et chez l'échantillon général dans lequel il a aussi prédit la taille-pour-âge. De règle générale, le pourcentage de variance expliquée par les

prédicteurs choisis était bas mais s'améliora (atteignit 12%) lorsque des sous-groupes de l'échantillon général furent analysés. Contrairement aux attentes et tel que démontré par l'analyse de sentier la plupart des effects de morbidité passée, des circonstances familiales, du statut économique, de la prévention, et de la disponibilité de la nourriture sur les indicateurs de croissance furent directs et ne semblaient pas opérer par l'entremise des pratiques alimentaires ou de la morbidité courante, excepté chez les sous-groupes des sous-alimentés.

d) PREDIRE LA MORTALITE: Le degré de prédiction de la mortalité atteint par l'équation discriminatoire ne fut pas impressionnante, classifiant correctement seulement 75% des décédés et 60% des survivants. La morbidité passée fut le facteur déterminant le plus important de la probabilité de mourir parmi les enfants, mais y était associée de façon négative. Suivent le statut économique, aussi relié négativement. La prévention était en trcisième place, reliée positivement et les circonstances familiales en quatrième place, avec une correspondance positive. Ces résultats sont, pour la plupart, difficile à interpréter et jettent le doute sur l'utilité de l'habileté discriminatoire de la formule. Les résultats des analyses univariées confirment l'association importante avec une mort précoce de certains composants de la mortalité, surtout la prématurité et la naissance jumellaire, telle que démontrée par la recherche antérieure.

Discussion: Des commentaires spécifiques sont émis sur les résultats de chacun des groupes analysés lorsqu'ils sont introduits. Une discussion technique plus générale suit à la fin. Le biais potentiel causé par les cas manquants sur la

validité interne et externe est discuté et l'effet de la colinéarité des variables indépendantes sur leurs risques d'être exclues du modèle est exploré. Des causes méthodologiques et conceptuelles sont proposées pour expliquer la faible proportion de variance expliquée qui fut obtenue dans cette étude.

Conclusion: Malgré les nombreuses réserves exprimées lors de la discussion, les mesures qui en resortent sont probablement les plus utiles pour améliorer la croissance de la plupart des enfants zambiens (les sains et les sous-alimentés, les plus jeunes et les plus vieux aussi) sont, par ordre d'importance: 1) réduire la morbidité en traitant et prévenant les maladies de l'enfance; mettre en place des mesures sociales qui visent à améliorer l'éducation et la santé des parents; 3) créer de nouvelles incitations économiques, essayant ainsi d'améliorer indirectement les conditions environmentales de chaque logis; 4) fournir les moyens financiers pour l'achat d'une quantité suffisante et d'une qualité adéquate de nourriture, et 5) donner accès à des soins médicaux compétents au besoin.