THE COLLABORATIVE RESEARCH SUPPORT PROGRAM

ON

FOOD INTAKE AND HUMAN FUNCTION

THE MANAGEMENT ENTITY

FINAL REPORT

USAID GRANT No. DAN-1309-G-SS-1070-00

SUBMITTED BY

D. CALLOWAY C. ROSBERG

H. HORAN S. MURPHY J. BALDERSTON

S. SELVIN G. BEATON

JANUARY 29, 1988

UNIVERSITY OF CALIFORNIA, BERKELEY

BEST AVAILABLE

ERRATA TO ME FINAL REPORT

August 17, 1988

Following are corrections to the Management Entity Final Report (dated January 29, 1988) on The Collaborative Research and Support Program on Food Intake and Human Function (Grant # DAN-1309-G-SS-1070-00). Please attach this page to your copies of the report.

1. p. 57, Table 9. Change the number of target schoolers in the Kenya anthropometry sample from 183 to 129 and the number of toddlers from 129 to 91.

2. p. 96, Table 32. Add "Kenya Toddlers" to the heading for this table.

3. p. 146, Table 67. Birthweights should be corrected as follows:

Project (Sex)	Age	N	Mean	Std Dev	% Median
Egypt (Male)	Birth	63	3.3 kg	.49 kg	101.0
Egypt (Female)	Birth	47	3.1	.45	96.0
Kenya (Male)	Birth	70	3.2	.47	97.9
Kenya (Female)	Birth	56	3.0	.45	94.8
Mexico (Male)	Birth	39	3.1	.37	96.0
Mexico (Female)	Birth	26	2.9	.40	90.6

The footnote SOURCES should be expanded as follows: Derived by ME from Cross Sectional Anthropometry Data Sets. Birthweights from: Egypt - Pregnancy Outcome File; Kenya - Pregnancy Outcome File; Mexico - Anthropometry File, weights within 4 days of birth. Birthlength was assumed to be the first length measure for each infant.

4. p. 150, Table 72. Change Egypt adult weights and heights as follows:

	Weight			Height		
	N	Mean	Std Dev	N	Mean	Std Dev
Egypt (Male)	187	69.8	11.6	124	166.6	12.4
Egypt (Female)	216	64.7	11.7	182	155.4	5.7

5. p. 169, Table 78. Change Kenya Project number of toddlers (NS) with lower respiratory tract illness from 17 to 10. Change footnote "d" to read, "Number of longitudinal subjects with at least one episode within the longitudinal window".

6. p. 182, Table 83. Add to the toddler measures: At 18, 24, 30 mos, Bayley Infant Behavior Record.

7. p. 298. Enter "1 var" for ferritin measure for Mexico Project.

THE COLLABORATIVE RESEARCH SUPPORT PROGRAM

ON

FOOD INTAKE AND HUMAN FUNCTION

THE MANAGEMENT ENTITY

FINAL REPORT

USAID GRANT No. DAN-1309-G-SS-1070-00

SUBMITTED BY

D. CALLOWAY C. ROSBERG

H. HORAN S. MURPHY J. BALDERSTON

S. SELVIN G. BEATON

JANUARY 29, 1988

UNIVERSITY OF CALIFORNIA, BERKELEY

BEST AVAILABLE

DEDICATION

The very existence of the Collaborative Research program on Nutrition and Function is owing to the late Dr. Martin Forman of USAID; without his concern and continuing stewardship the work would not have been done at all, or having been done, it would not have been carried out with such attention to the needs of those policy makers who must translate findings into programs. It is fitting that this final report of our efforts be dedicated in his memory.

ACKNOWLEDGEMENT

Our colleagues who made up the Scientific Coordination Board were "Management" in the truest sense. We wish to recognize Lindsay Allen, Nimrod Bwibo, Adolfo Chavez, Osman Galal, Gail Harrison, Norge Jerome, Avanelle Kirksey, Charlotte Neumann, and Gretel Pelto - and to thank them for their yeoman service as management while simultaneously guiding their research projects, and carrying on their home duties as well.

Samuel Kahn, our USAID project officer not only acted in and on the program's behalf, but also gave wise and generous counsel on administrative and scientific matters. Irwin Hornstein, formerly of USAID and a participant in early pre-CRSP development, assisted us briefly but importantly in this effort.

We appreciate the spirit of helpfulness in which the External Evaluation Panel approached the difficult task of keeping the CRSP to the right path. Sol Chafkin, Marian Yarrow, and Gerald Keusch were our advisors from before the CRSP was funded to the very end; they richly deserve our special thanks. T.N. Srinivasan, Donovan Thompson, Richard Adams, and Philip Payne made unique and valuable contributions throughout the CRSP.

Among the Berkeley family we are particularly grateful to two whose names never appear on our budget. Robert Edwards of the Sponsored Projects Office served as chairperson of the Institutional Council, steering it -- and ourselves -- through times of difficult administrative decision-making and ably looking after our business affairs on campus as well. Karin Beros' gracious and unobtrusive guidance of our day-to-day operations within the Institute of International Studies was a key element in the efficiency of our office staff and their pleasure that accompanied the work.

All of the ME staff were special and all are appreciated. On the technical side, particular credit goes to our senior programmer David Lein, and to Birgit Leibel, Patsy Fosler, Donald Rose, Claudia Waters, and Atnaf Keffelew. Our successive principal administrative assistants handled the varied and difficult tasks that kept our activities glued together: Katherine Condon, Diana Badger, and Katherine Merrill. We could not have managed without this dedicated group, and the array of short-term, part-time and work-study helpers who were such a cheerful presence. Our thanks to all of you.

TABLE OF CONTENTS

Dedication iii

Acknowledgment iv

Table of Contents v

List of Tables vii

List of Figures xi

Summary and Conclusions xiii

I. Background 1

II. Organizational Structure of the CRSP 1

III. Narrative Chronology of Principal Management Activities 4

IV. Fiscal Summary 27

V. Organization and Management Assessment 29

VI. Scientific and Technical Assessment 37

A. Selection of Communities, Target Households, and Subject Classes .39

B. Size of the Cross-Sectional and Longitudinal Samples 51

C. Food Composition Data Bases 63

D. Food Intake 73

E. Resting Metabolic Rate 115

F. Anthropometry 143

G. Morbidity and Mortality 157

H. Psychology 181

I. Reproduction 189

J. Social and Economic Measures. 199

VII. Data Management and Archiving 209

VIII. Testing CRSP Hypotheses: Review of the Project Analyses 227

IX. References Cited 245

Appendix A. Members of CRSP Organizational Bodies: SCB, IC, TAGs . . 247

Appendix B. Members of the External Evaluation Panel 251

Appendix C. Management Entity Staff and Consultants 253

Appendix D. Chronology of CRSP Activities 257

Appendix E. CRSP Basic Files and Master Index of Basic Variables . . 267

LIST OF TABLES

1. Criteria for Selection of Communities 40
2. Composition of Households in CRSP Projects 48
3. Number of Targets in the Cross-Sectional Sample 52
4. Number of Longitudinal Subjects 54
5. Number and Entry Dates of Longitudinal Households 54
6. Number and Entry Dates of Longitudinal Schoolers 55
7. Number and Entry Dates of Longitudinal Toddlers 55
8. Number and Entry of Longitudinal Pregnancies 56
9. Number of Targets in the Longitudinal Sample 57
10. Frequency of Food Intake Measures on Lead Males in Longitudinal Households 57
11. Frequency of Food Intake Measures on Lead Females in Longitudinal Households 58
12. Frequency of Food Intake Measures on Longitudinal Schoolers 58
13. Frequency of Food Intake Measures on Longitudinal Toddlers 59
14. Frequency of Food Intake Measures on Longitudinal Preg/Lact Women . 59
15. Frequency of Anthropometry Measures on Lead Males in Longitudinal Households 60
16. Frequency of Anthropometry Measures on Lead Females in Longitudinal Households 60
17. Frequency of Anthropometry Measures on Longitudinal Schoolers . . . 61
18. Frequency of Anthropometry Measures on Longitudinal Toddlers 61
19. Frequency of Anthropometry Measures on Longitudinal Preg/Lact Women 62
20. Frequency of Anthropometry Measures on Longitudinal Infants 62
21. Summary of Food Composition DataBases for the Three Projects 64
22. Observed Distribution of Energy and Nutrient Intakes ME-defined Longitudinal Samples 75
23. Probability Assessment of Observed Protein Intake 83
24. Estimated Effect of Dietary Methodology Changes, Mexico Project 87
25. Effect of Methodology Changes on Estimated Nutrient Intakes Among 28 Adult Males, Mexico Project 88
26. Range of Mean Energy Intakes by Interviewer 89
27. Mean Toddler Energy Intakes by Egypt Project Interviewers 90
28. Energy Intakes Reported by Different Interviewers for the Same Toddler, Egypt Project 91
29. Estimates of the Partitioning of Variance in Kenya Toddler Intakes . 94
30. Comparison of Intakes Recorded on Adjacent Days, Kenya Toddlers . . 95
31. Comparison of Observed Unexplained Variance in Pooled Two Day Data 96
32. Comparison of the Partitioning of Variance in Two Time Periods, a Period of Food Shortage and a Period of Stability 96
33. Estimates of the Partitioning of Variance in Kenya Schooler Intakes 98
34. Estimates of the Partitioning of Variance in Kenya Adult Intakes 98
35. Contingency Table Examination of Direction of Same Day Deviations in Maternal and Toddler Intakes - Kenya 100

36. Estimates of the PArtitioning of Variance in Mexico Toddler Intakes (Final Methodology) 101
37. Estimates of the Partitioning of Variance in Mexico Adult Female Intakes (Final Methodology) 102
38. Estimates of the Partitioning of Variance in Mexico Adult Female Intakes (Early Methodologies) 102
39. Estimates of the Partitioning of Variance in Egypt Toddler Intakes (4-9-4 Calculations) 104
40. Estimates of the Partitioning of Variance in Egypt Adult Female Intakes (4-9-4 Calculations) 105
41. Examples of the Predicted Attenuation of Bivariate Correlation and Regression Analyses as a Function of Number of Observations Pooled in the Estimate of Intake, Kenya Project Data 106
42. Comparison of Variance Components for Selected Dietary Measures Across Projects 107
43. Energy Intake Ratios for Several Kenya Households 111
44. Energy Intake Ratio, Summary Statistics by Household. Kenya Project 113
45. Distribution of Low Energy Intake Ratios Over Time, Kenya project 114
46. Distribution of Replicated RMR Observations over Entire Study Period 117
47. Comparison of Observed RMR and Predicted RMR (Adult Males in Kenya and Mexico) 117
48. Regression of RMR on Weight and on Predicted RMR of Adult Males . . 118
49. Crude Estimates of the Partitioning of Variance (Adult Male Subjects) 119
50. Comparison of Directions of Change in RMR within Persons across Time (Mexico Data) 120
51. Examination of Measurement Day Effect on RMR - Mexico Data 121
52. Effect of Day of the Week on RMR Estimates - Mexico 122
53. Effect of Sequence of Measurement during the Day on RMR Estimates - Mexico 122
54. Operating Definitions of "Famine" Periods - Kenya 123
55. Frequency of Replicated Measures, Kenya 124
56. Paired Analyses of G-Famine Effects on Body Weight, Kenya 125
57. Impact of "Famine" on RMR and Food Intake: Kenya Adult Males . . . 127
58. Impact of "Famine" on RMR and Food Intake: Kenya Adult Females . . 130
59. Impact of "Famine" on RMR and Food Intake: Kenya Schoolers 133
60. Effect of "Full Correlation" on PArameters of RMR Estimation . . . 141
61. Regression Analyses of the Impact of "Full Correlation" on RMR, Kenya Adult Males 141
62. Variables Common to All Projects 143
63. Reliability of Weight Measures in Toddlers - Egypt 144
64. Reliability of Weight Measures in Toddlers - Kenya 144
65. Indices of Height Measurement Validity, Adult Males 145
66. Indices of Length or Height Measurement Validity in Infants, Toddlers and Schoolers 145
67. Infant Weights and Lengths at Birth, Age 3 to 6 Months, Sexes Combined 146
68. Weights and Lengths of Toddlers by Age - Egypt 147
69. Weights and Lengths of Toddlers by Age - Kenya 148

70. Weights and Lengths of Toddlers by Age - Mexico 149
71. Schooler Weights and Heights (at Entry) 150
72. Adult Weights and Heights 150
73. Planned Morbidity Protocol 159
74. Estimated Number of Days Observed within Longitudinal Windows . . . 163
75. Identification of Episodes of Morbidity: ME Approach 164
76. Illness Symptoms and Frequency of Reporting: Mexico 167
77. Estimated Duration of Specified Classes of Illness, Mexico 168
78. Mean Duration of Illness Episodes in Children, Egypt and Kenya . . 169
79. Average Incidence of Specified Disease Categories
Among Longitudinal Infants and Children: Egypt, Kenya, Mexico . 170
80. Comparison of Frequencies of Application of Severity Codes
by Target Type 172
81. Comparison of Reported Food Intakes of Toddlers on Days
Reported Ill and Days not Reported Ill: Mexico 174
82. Mortality during CRSP Phase II in Study Households 179
83. Summary of Planned Measures of Psychological Function
Basic and Common Data Sets 182
84. Cognitive Scores of Adults and Schoolchildren 187
85. Comparison of the Brazelton Scores in Kenya and Egypt and
Lester Scores in Mexico 188
86. Definition of Reproductive Periods in Project Final Reports 190
87. Maternal Food Intake during Pregnancy and Lactation
as Reported by CRSP Projects 192
88. Coincidence of Reproductive Phase and Food Shortage in Kenya . . . 193
89. Maternal Height and Weight during Pregnancy and Lactation
as Reported by CRSP Projects 195
90. Resting Metabolic Rate of Pregnant and Lactating Women
as Reported by Kenya and Mexico Projects 197
91. Spearman's r's of Selected Kenya Final Report SES Variables
from Form 631 205
92. Spearman's r's of Selected Kenya SES Variables and
Toddler Energy Intake 208
93. Egypt Basic Files 217
94. Kenya Basic Files 219
95. Mexico Basic Files 221
96. CRSP Archive Format 224

LIST OF FIGURES

1. Organization and Management Structure of the CRSP: Program and Projects 2

2. Histogram of Distribution of Energy Intakes of Toddlers, kcal/d 76

3. Histogram of Distribution of Energy Intakes of Toddlers, kcal/kg/d 77

4. Histogram of Distribution of Energy Intakes of Schoolers, kcal/d 78

5. Histogram of Distribution of Energy Intakes of Schoolers, kcal/kg/d 79

6. Histogram of Distribution of Energy Intakes of of Adult, kcal/d 80

7. Histogram of Distribution of Energy Intakes of of Adults, kcal/kg/d 81

8. Body Length of 18-30 Month Olds Plotted by Age and Calender Date, Kenya Project 153

9. Body Weight of 18-30 Month Olds Plotted by Age and Calender Date, Kenya Project 154

SUMMARY AND CONCLUSIONS

The 1977 World Food and Nutrition Study of the US National Academy of Sciences gave its highest priority to determination of the functional consequences of chronic mild-to-moderate undernutrition. In 1981, the US Agency for International Development funded a Collaborative Research Support Program (CRSP) on this topic, under the auspices of the Board for International Food and Agricultural Development. Field research studies were carried out in Egypt, Kenya, and Mexico. Collaborating institutions were: the Egyptian National Institute of Nutrition with the Universities of Arizona and Kansas and Purdue University; the University of Nairobi with the University of California at Los Angeles; and the Instituto Nacional de Nutricion, Mexico with the University of Connecticut. Funding was accomplished by a grant awarded to the University of California, Berkeley, which served as the Management Entity (ME) and subcontracted funds to Purdue University, UCLA, and University of Connecticut for their research and to be passed through them to the overseas institutions for research support in the field. Collaborating US institutions contributed at least 25% of the US costs of this research. Research data and complete documentation, including final reports of the three projects, have been archived with USAID as the terminating activity of ME.

As the CRSP evolved, responsibility and authority assigned to the several components of the management structure evolved as well. A Scientific Coordination Board (SCB), composed of US and host-country principal investigators (PI) and one representative of the ME, had final responsibility for research design and other technical matters. An Institutional Council, whose members were administrative officers from the US project institutions and ME, provided oversight of budgets and management functions. The scientific merit of the projects and program was evaluated periodically by an interdisciplinary group of experts, the External Evaluation Panel, who reported to USAID. Areas of management, organization, and infrastructure which presented serious or recurrent problems are discussed in this report, together with recommendations for future consideration. In retrospect, it is clear that from the beginning of a program, there needs to be an unequivocal definition of roles, responsibilities, and authorities of all major actors, from the most senior level of management to the level of field data collector. There should also be clear agreements on the transfer and archiving of data, and on the sharing of information to promote mutual understanding.

The goal of the Nutrition CRSP was to test the hypothesis that food intake related, in a causative manner, to functional outcomes that are believed to have import for the development of individuals, households, and communities. A key design consideration was that this generic hypothesis, and its many specific hypotheses, should be examined in different biological and cultural ecosystems. Accomplishment of the most basic goal of the CRSP called for interproject analyses of the data collected with care and effort by the three projects.

Complex field research involving human subjects is never perfect and this CRSP is no exception. ME's scientific assessment of the projects' data as archived centrally and the analyses presented in their individual final reports thus appears somewhat critical. Problems have been identified in the completeness

of collection of data on key variables and differences in methodology, sometimes very subtle, between projects (and in one case within a project over time) that limit or prohibit comparison between projects. Criticisms not withstanding, it is ME's considered judgment that these data sets are among the best ever produced in similar settings. ME's concern is with cross-project evaluation; the problems identified impact less severely on analysis of data within a project.

Neither time nor funds were sufficient to complete analysis of the data. Recognizing this reality in the last two years of the CRSP, the SCB established priorities among the many original hypotheses of the CRSP design. Adhering to this agreement, each project focused analytical attention on three areas: the effect of food intake on morbidity (and physical growth) in toddlers, the effect of food intake on psychological development (and social behavior) in toddlers, and the effect of food intake on pregnancy outcome (and lactation performance). In the press of time, analyses were carried out before the basic processes of data validation and refinement were complete. The ME assessment and the projects' final reports should be read with this in mind.

Project reports present a considerable volume of descriptive statistics as a background to analysis (by prior agreement). There is temptation to overinterpretation of these presentations. While pairwise comparisons are useful descriptions of relationships within a data set, their clear limitation is the inability to take into account multiple relationships or effectively to examine direction (causality) of associations, thus making accurate biological interpretation difficult. Furthermore, measurement error can attenuate the relationships seen with the net effect that the observed associations do not portray correctly, or perhaps even detect with statistical significance, the underlying biological relationships. The major advantage of the rich data sets collected in the CRSP is the possibility of defining the role of a number of factors in the relationship between nutrition and growth, morbidity, and cognitive status of individuals. Although application of multivariate techniques (preferred in most situations but not all) requires careful attention, it has the potential to exploit the exceptional richness of the projects' data sets. The point being made is that there is risk inherent in overinterpretation of what, at present, amount to preliminary analyses.

This said, all projects have reported correlations between some dietary variables and some outcome variables. For example, at the level of correlation analyses, all projects found associations between measures of intake and achieved body size; the relationships were not always consistent within or across projects. Nevertheless, there remains an underlying basis for assuming that there is a relationship, expressed in broad terms, between an individual toddler's intake and achieved size and between maternal intake and infant size in the first six months of life. Body size is a consistent predictor of cognitive development, as measured by Bayley scores, at 30 months of age and a less consist predictor of morbidity. In the final analysis, it may well be that size per se is not the issue, but rather an understanding of the processes, including food deprivation, that result in smallness.

The dietary variable identified in the present analyses as significant is sometimes energy intake but more often in Kenya and Mexico, an element of dietary "quality" reflected in protein intake from animal sources or energy from animal foods or fat. Much further work is required to determine if these "quality" factors have a causal relationship to outcomes, if their significance is a statistical anomaly due to variance features, or if they are simply an indicator of some as yet unidentified true effector variable.

A clear and important finding from all projects is that growth falters early in the lives of people living at the margin. The impact is iterative: larger mothers have larger babies. Faltering is demonstrable by about the infant's fourth month, on the average, and is clear by the sixth month. Children who are small at 18 months of age predictably are also small at age 30 months. Examination of the data on children aged seven to nine years and of other nontarget siblings present in the households has yet to be done; it will be important to understanding the significance of the concomitants of early growth and to evaluating the probability, if not possibility, of catch-up growth in these environments and circumstances.

Much remains to be done before the analyses of even the selected priority hypotheses can be said to be complete. The work to date represents but a small fragment of the work that can be done. There has been a major investment of time and effort by a very large number of people, most notably the monumental contribution by subjects in the three communities that were studied. What is now needed, to capitalize upon this investment, is that dedicated teams of skilled investigators who are prepared to invest major effort in the analyses of these data be given the resources needed to support the analysis of data already collected rather than being diverted to the collection of new data. All projects have enriched the scientific communities that were overseas partners in this research. It would be tragic if these partners were omitted, by default of funds, from participating in the full evaluation of their work, which would allow them, who are best positioned to do so, to translate the scientific findings for the formulation of effective national nutrition health policy.

It is the plea of ME that the opportunity, indeed the obligation, to complete the task initiated by the Nutrition CRSP will not be allowed to pass.

I. BACKGROUND

The first formal proposal for research on the topic of food intake and human function was made by a committee convened in connection with the World Food and Nutrition Study of the National Academy of Sciences (NAS) (1,2). The concept was developed further at a workshop convened by the NAS in 1977 (3). The United States Agency for International Development awarded a contract to the University of California, Berkeley (UCB), in April 1978, for development of a research program plan that could be considered by the Board for International Food and Agricultural Development (BIFAD) for funding under the provision of Title XII of the Foreign Assistance Act. Projects to be included in a Collaborative Research Support Program were recommended by an External Steering Group following peer review of research proposals from eligible institutions. The final planning activity was to convene the potential US and host country investigators in December 1979 to flesh out a collaborative research design and to define management functions and structure. (For details of planning, see ref. 4.) The US collaborators selected UCB to be the Management Entity.

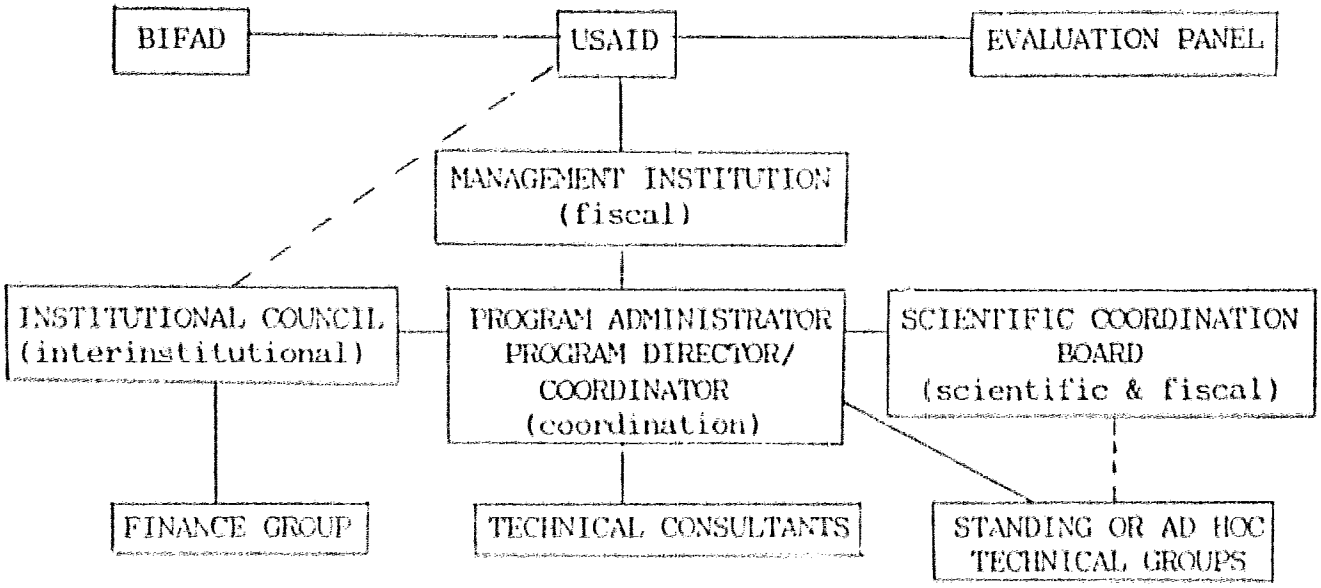
After another, approximately 18 month-period of USAID review and deliberation, the CRSP grant document was signed by USAID (on September 18, 1981); the grant was accepted by the University of California on December 15 of that year. The major tasks necessary for establishing the scientific, management, and financial components of the program were set in motion some two years after plans had been completed. This delay meant that research teams had to be assembled anew as many intended investigators had turned their attention elsewhere. Changes that had occurred in host countries meant a renegotiation of terms of and arrangements for overseas collaboration. The funds awarded were based on budgets prepared in 1979, with no adjustment for inflation. Thus, time and money were constrained from the beginning.

II. ORGANIZATIONAL STRUCTURE OF THE CRSP

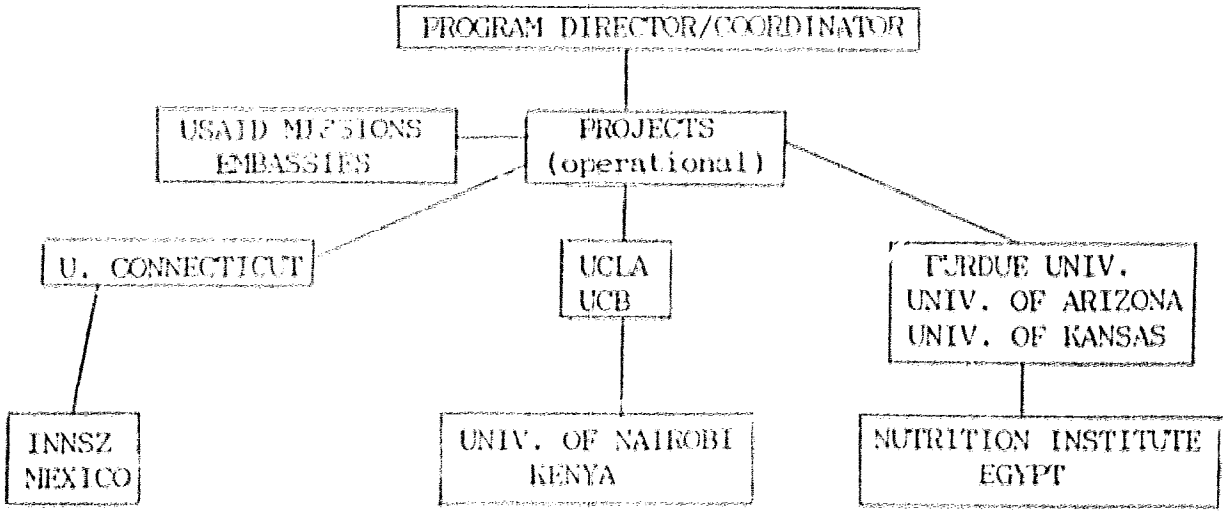
Central coordination and management of the interinstitutional Collaborative Research Support Program on Food Intake and Human Function (the Nutrition CRSP or simply CRSP in this report) was required to sustain the scientific coherence and integrity of the projects as a collective effort, and to arrange fiscal and other administrative matters such as project/program evaluation, reporting, travel authorizations, and the like. As noted above, advance planning for the CRSP included consideration of an appropriate management structure among which these responsibilities were to be divided. Recommended were a Scientific Coordination Board (SCB), an Institutional Council (IC), and a Policy Consultative Group (PCG). Federal requirements were that the CRSP be implemented through a Management Entity (ME), and that an External Advisory (later Evaluation) Panel (EEP) be named to act as a peer review body. An organizational chart showing the relationships of these bodies is displayed as Fig. 1.

Figure 1

a. Program Structure



b. Project Structure



Scientific Coordination Board. Operationally, the SCB served as scientific decision-maker of the CRSP. As such, it had final responsibility for research design, including review and approval of non-core project research, and other technical matters such as publications policy. Voting members were one US and one host-country Principal Investigator (PI) from each project and one person from ME (see Appendix A). The SCB met about twice each year, often with additional PIs in attendance and sometimes with other project and ME staff as auditors or advisors on particular topics. ME was a scientific collaborator in the transactions of the SCB, and attempted to monitor the projects' performance in implementing SCB directives. The SCB was aided in its scientific planning by ad hoc Technical Advisory Groups (TAGs) that included project investigators and outside technical consultants as required in specific research areas.

Institutional Council. The IC's role was to facilitate interinstitutional arrangements including review and approval of project budgets, and general oversight of management functions. In budget matters, the IC acted with the advice of a Finance Group, comprised of the institutional IC members and one PI from each US institution. The IC members were business or research administrative officers selected by the US institutions (see Appendix A). Voting was by project. The IC had general administrative oversight of the CRSP and was empowered to communicate with USAID. In practice, IC recommendations were made directly to ME.

Policy Consultative Group. Because it was foreseen that the results of the CRSP would have important implications for government policy, a PCG was formed by USAID staff, in consultation with ME. The PCG, consisting of policy makers from the host countries and two US policy analysts, met once to assess policy relevance of findings presented by US and host country PIs.

External Evaluation Panel. The interdisciplinary nature of the CRSP required that the EEP be drawn from several disciplines (see Appendix B). It was the EEP's responsibility to evaluate scientific merit of the projects and program, including the research design, and to advise USAID of its findings and recommendations. The EEP carried out its function by meeting as a group or selectively with the SCB, by site visits to the projects, and by review of written reports.

The Management Entity was designated as the institution receiving the USAID grant. It was vested with ultimate responsibility for both the scientific and fiscal management of the project, consistent with direction and advice from the EEP and IC.

Among its various roles, ME was made specifically responsible for dispersing funds to US institutions for country projects, including those funds to be passed through to other US institutions and participating foreign institutions. In addition, ME was to be responsible for both the scientific and fiscal audits of the use of CRSP funds, and for holding PIs of lower tier agreements accountable for their institutions' use of funds to support their specific research program based on decisions taken by the SCB. ME carried out its work through the office of a Program Director/Coordinator with a modest

administrative and clerical staff, an economy made possible by the generous support of UCB's Institute of International Studies, which housed the ME activity. At the early stages of planning data analysis and management, members of the UC statistics department advised ME. Graduate students from that department served in the second year. Later, the ME was augmented by a data management/nutrition expert and a computer programming staff, as well as the retaining of consultants as required, some on a continuing basis (see Appendix C).

Management of project operations and funds was the responsibility of PIs and their institutions. The US PIs were all university faculty with continuing academic responsibilities to their US institutions, which made continuing residence abroad impossible. Each project's US institution appointed a full-time director as manager of field operations who worked under the day-to-day oversight of the host country PI, a burden assumed in addition to their normal and very substantial obligations to their own institutions.

III. NARRATIVE CHRONOLOGY OF PRINCIPAL MANAGEMENT ACTIVITIES

A chronology of meetings, site visits, reports, etc. appears in Appendix D; what follows is a brief description of major ME activities in each year of the CRSP.

A. Year One: 1 September 1981 - 31 August 1982

Overview

The Management Entity office was established at the University of California, Berkeley, under the direction of Principal Investigators Professors Doris Calloway and Carl Romberg. While the search for a Program Director was underway, Professor George Beaton served as Interim Director of ME until April 1982. During the first part of the year, contracts were formulated, and CRSP organizational groups established. Administrative arrangements were made between ME and the US institutions, and between the US institutions and their corresponding overseas institutions. Research activities during this first year were largely planning and organization.

The most significant scientific accomplishment of this period was the further development of the interim plan for the core longitudinal study, following from that proposed to USAID in the final report of the planning activity (4). This interim plan in large part determined the design and implementation of the "Phase I" pilot studies and collection of preliminary community and household data related to site selection. The completion of detailed Technical Advisory Group (TAG) reports provided the basis for selection of independent and dependent variables to be included in the core research design.

1. Management Entity and Administrative Organization

While the position of Program Director had not been filled by the close of Year One, it was envisaged that the person occupying this position would be the liaison between ME, the scientific management of the CRSP as a whole, and the management of the individual projects. It was intended that the appropriate person would be nominated by ME, concurred on by USAID, and approved of by both the SCB and the IC. A national search was undertaken to fill the position, but reached no conclusion.

During the second and third quarters of the year, sub-contract agreements were negotiated, and logistics for field work developed. A meeting of the SCB pro tem was held at Berkeley in October 1981 to plan for the preliminary field studies and establish the mechanisms for the exchange of methodology information. In late November names were proposed for the EEP, the IC, and the PCG. The EEP and the IC were appointed; appointment of the PCG was deferred.

Eight TAGs, which represented the seven primary areas of the study--nutritional intake, nutritional status, activity, cognition, social performance, reproduction, and disease--as well as statistical design and data management, were appointed (Appendix A). Each was instructed to consider a set of key questions related to the purpose of the CRSP, propose methods of data collection, and suggest models for data analysis. The resulting reports included suggestions for the collection of research variables, drafts of survey instruments, and the identification of aspects requiring further testing or modification in the field.

TAG reports were reviewed at an April 1982 SCB meeting in New Orleans. A CRSP policy was adopted stipulating that regardless of funding source, all supplementary studies involving core research subjects must be approved by ME and the SCB. The Egypt project delivered notification of Dr. Ronald Watson's departure as PI at Purdue University, and Dr. Avanelle Kirksey was endorsed as his replacement.

Also at this meeting the recommendation was made that Dr. Judith Balderston be invited to accept a position as Deputy Program Administrator while the difficult search for a Program Director continued. Although a prominent nutritional scientist was being sought, it was agreed that Dr. Balderston's background in economics, her prior experience in other research related to education and nutrition in LDCs, and previous management of an AID-sponsored project made her well-equipped to assume responsibility within the Management Entity. As a condition of the appointment, it was recommended that the Deputy Program Administrator report to Professor Calloway, who would act as Program Administrator and would commit significant time to the program; Professor Beaton, who had left his post as CRSP Interim Director in April, was to continue to provide specific nutritional science expertise as Senior Consultant to ME.

2. Research Planning and Preliminary Field Work

In June 1982, a meeting of the US PIs and an ad hoc Statistics and Data Management Committee was held to discuss the construction of the integrated Interim Design from the separate TAG proposals, for use in testing in the field during the summer. Decisions reached at this meeting were refinements of the original project planning design; discussions focused on the selecting of criteria in relation to the overall design, the selecting of households and subjects, the scheduling of observations in the household, the measurement of specific variables, and the methods and procedures appropriate for data collection.

Throughout this first year, the organization of activities by project between each set of US and LDC institutions was accomplished by extensive collaborative efforts through PI visits, the exchange of scientific staff, and coordination of activities at the field sites by field staff. During this period, each project arranged visits of US PIs to their collaborating overseas institution and to the field site. Overseas PIs and scientists also were brought to planning meetings in the US.

a. Phase I Project Activities. To guide selection of the field sites, several communities were evaluated in relation to agreed-upon criteria, which were: appropriate size, homogeneity, adequate infrastructure, and the presence of mild-to-moderate malnutrition. (These criteria had been developed in 1979 (4) and are fully described in the "Phase I, Preliminary Survey" document, circulated during the first year of the project.[1]) Several communities in each country were considered, and comparisons and choices made among them.

Phase I studies of food intake were also carried out in order to test field methods and to determine the variability of energy intake from day to day (information needed to establish sample size). Preliminary collections of anthropometric and clinical data were made in order to test methods of observation for feasibility and reliability.

Field work during this first year of the CRSP progressed, understandably, at different rates. In the Mexico project, work proceeded smoothly after the April 1982 signing of the subcontract agreement between UCB and the University of Connecticut. The field staff were appointed, and since well-established facilities were located near the field site, it was possible to move workers rapidly into the area of the study in Solis. Community selection began, censuses and ethnographic studies were carried out for several communities. During the summer and fall of 1982, food intake data were collected, and indices for dietary complexity developed.

The Egypt project experienced some delay in the signing of contracts. While the project's planning activities in the host country were underway by May of '82, difficulties in negotiations deferred the start of actual field work until after the contracts were signed in late August. The directorship of the

[1] Documents referred to in this report are archived in microfilm or microfiche with USAID.

collaborating Egyptian Nutrition Institute changed with the retirement of Dr. Hekmat El-Sayed Aly, on June 1, 1982.; Dr. Osman Galal was appointed Director and assumed the role of Project PI.

In the Kenya project, long delays were experienced in the signing of the inter-campus agreement between ME at UCB and UCLA, due to continued negotiations with respect to the UCLA budget. In addition, contract negotiations between UCLA and the University of Nairobi were delayed, partially owing to an attempted coup in the country in August 1982, which resulted in the closing of the University for several months. Before the contracts were signed, however, an exploration of possible village sites was undertaken by Nairobi-based field staff, including the planning of ethnographic surveys, preparation for mapping, and the conducting of demographic surveys.

3. Data Management Activities

The TAG on Statistical Design and Data Management, which had convened in January 1982, proved extremely useful in defining tasks and data flow from field collection operations, and in emphasizing the responsibilities of the data management groups. The Deputy Program Administrator worked with the US data managers to coordinate and simplify the processes of quality control and analysis; it was recommended that all universities would use the SAS (Statistical Analysis System) management system for statistical analysis and data manipulation. It was also agreed that in the US, data would be stored and managed at UCLA and the Universities of Connecticut and Kansas, as well as at the ME central facility at UCB. Following the recommendation of the TAG, it was further agreed that data entry, verification, and the first stage of editing would take place at the oversens country locations, with full precautions being built into the system to ensure that data at the country and corresponding US institutions and the central data bank at Berkeley would be kept consistent with regard to data corrections. Data flow procedures were also outlined by the US data managers, and in the latter part of the year, plans were made for the formation of an ongoing committee of data managers.

B. Year Two: 1 September 1982 - 30 September 1983

Overview

The second year of the CRSP marked the completion of Phase I activities--community selection, preliminary data collection, pilot testing of instruments for feasibility in the field, conducting socioeconomic surveys, ethnographic studies and community census, and the preparation for and approval of a "Phase II" research design (the core longitudinal study) by the EEP. These activities were accomplished through several inter-project meetings during the period, along with an extensive review process.

1. Country Project and Field Activities

In addition to the year's inter-project meetings, each component project continued its individual planning and field activities. These included the development of site facilities, the hiring and training of personnel, the

ordering of equipment, and various other tasks connected with the establishment of the broad-scale field research.

As part of their completion of Phase I activities, all three projects were intended to collect preliminary dietary intake (converting to energy) and anthropometric information on subsets of the population, for purposes of examining the range of nutritional deficiency to be found in each community, and to assure that the proportion of households experiencing mild-to-moderate malnutrition would be adequate for research design needs. Not all projects fully accomplished this work. Also during this period, the three projects worked together to develop various methodological protocols that were later documented in a manual prepared by ME.

Because of continued delays in Kenya project contract negotiations, USAID representative Dr. Irwin Hornstein accompanied Professor Charlotte Neumann to Kenya to facilitate the procurement of an agreement acceptable to the government of Kenya. Finally, in February 1983, the formal subcontract between the University of Nairobi and UCLA was signed. As noted above, some field activities were undertaken before the official contract signing. In the Embu district--the chosen site--staff recruitment and training began in early September 1982, along with the census and survey of households, and collection of preliminary data on anthropometry, morbidity, and reproduction, using pre-coded forms. These preliminary Phase I data were then corrected and entered on computer diskettes by the Kenyan field team, and sent to UCLA.

With the selection of Kalama as the Egyptian field site, Phase I field work began and preliminary data for this project were collected on anthropometry of pre-school and school-age children, food intake variability, community descriptions, and preliminary morbidity surveys. A field director was appointed and began work at the Nutrition Institute in Cairo in early October.

The Mexico project, which had already established its field work during the Spring of 1982, completed the collection and preliminary analysis of data in a pilot study conducted in San Francisco de Solis, a small community in the Solis Valley.

2. Phase II Research Design and Review

At a November 1982 SCB meeting held in Virginia, all PIs and ME representatives collaborated to create a Phase II research plan that would encompass the complex interrelationships between intake and each of the major functions designated by the NAS (3) for study. The participants discussed the appropriate unit of study (households and individuals), the choice of target subjects, the sampling strategy, and the specific attributes of functions to be included in the design.

The selection of nominees for the EEP was completed at the November 1982 SCB meeting. In December, ME assisted members of the SCB in drafting the research design report, which was forwarded to the EEP for evaluation.

The EEP met in January of 1983 with project representatives and ME to confer on the proposed design. Members of the panel suggested that this Phase II

plan needed simplification, to "assure the most direct and cost-effective way to clarify the central question of intake-function relationships." They strongly recommended that the focus of study be reduced to two high priority, "low risk" functional relationships during the first year of Phase II: (a) the association of intake with child morbidity, and (b) the association of intake with child psychological development. This recommendation had the effect of shifting the focus of interest from the household to specific target individuals within the household, and mandated a reconstruction of the design.

The EEP also advised that, because social performance functions were ill-defined in the current research design, before social performance measures could be included in the study "research must be individually tailored to the various research sites and local experiences." Along these lines, they suggested that each project initiate a series of "case studies" to define better the country-specific research questions and to develop appropriately tailored plans, either for their own project or for all three projects in parallel, to be implemented during the second year of Phase II.

The EEP further recommended that non-core studies (i.e. research not included in the collaborative design) of particular interest to investigators be limited to not more than 15% of each project's budget. In addressing the organizational structure of the program, the EEP recommended that each project employ a statistician/data manager, and that Berkeley recruit a statistician to provide analytical assistance to the three projects. Lastly, the EEP recommended that plans be reformulated, and that a refocused design be submitted for review in June or July of that year.

The CHSP response to the EEP's recommendations on design modification was formulated at the May 1985 meeting of US PIs and ME. It was decided that during the first year of Phase II, efforts would focus on three major outcomes - child morbidity, child psychological development, and reproductive function. It was decided that research would address six main hypotheses, as follows:

1. Maternal intake during pregnancy and lactation influences endowment at birth, and during the infant's first six months of breast feeding. Also, maternal intake affects maternal activities for sanitation and child care. Thus, infant behavior and morbidity are affected by maternal food intake.

2. Food intake of the child between 18 and 30 months of age affects his/her morbidity, body weight, and psychological development, while maternal intake affects child care and sanitary practice. Thus, behavior and morbidity of the young child are affected by the food intake of both child and mother.

3. Food intake of the child between 7 and 9 years affects the child's morbidity and behavior, while parents' intake affects behavior towards the child.

4. Food intake of adults affects their morbidity, responsiveness, and performance of usual responsibilities, and this in turn affects the morbidity and performance of other household members.

5. Adjustment of RMR in adults provides a major path of adaptation for restoring energy equilibrium.

6. Household food intake affects household morbidity.

The May meeting also included a lengthy discussion of sampling and other methodological issues pertaining to these hypotheses. A document entitled "Revised Research Design - Phase II, June 1983" (the so-called "Blue Book") ensued from these discussions. In this document, ME translated SCB research plans into specific variables and schedules. It was circulated to EEP members, whose reactions were in turn forwarded to the USAID Nutrition Office Director, Dr. Martin Forman. Panel members agreed that the refocused design was now acceptable, and that the CRSP should proceed to Phase II. A meeting of the EEP Chairman, the USAID Program Officer, and members of ME was held in Berkeley in mid-August, at which time it was officially decided that the CRSP should proceed to Phase II data collection by October 1983.

3. Management and Administration

Given the complexity of the CRSP structure, the establishment and monitoring of its organizational relationships and reporting mechanisms proved to be quite intricate, and at times cumbersome, during this second year. Because of the involvement of so many investigators from such a variety of academic disciplines, the organizational process required considerable effort to maintain collaboration among the projects and management. Despite the difficulties inherent in such a complex set of institutional arrangements, management of the CRSP proceeded relatively smoothly, due to the commitment of participants at all levels.

A report on "Policy Implications and the Nutrition CRSP" was written by Dr. Judith Geist in consultation with Professor Carl Rosberg, Co-Program Administrator. The report, which dealt with the issues of public policy that are closely related to the CRSP research areas, was submitted to AID for circulation as background for a yet-to-be named PCG.

A number of administrative and organizational details were successfully addressed at the major meetings of the year. At the November 1982 SCB meeting the SCB voted in a new US Co-PI for the Kenya project. This project was originally intended to be a joint effort between UCB and UCLA, but the intended UCB PIs (Professor Calloway, who had become UCB Provost of the Professional Schools and Colleges in 1981, and Professor Ostwald, who was terminally ill) were no longer available. Professor Janet King and Assistant Professor Robin Gorsky served, at different times and in that order, as Co-PIs.

a. International Travel. A number of difficulties arose involving travel billing, and the certification of the use of non-US airlines. International travel authorization and reporting procedures were reviewed to improve coordination with the USAID Program Officer. The services of Thomas Cook, Inc., the travel agency with whom ME had established a contract in March of

198[illegible] were continued, an arrangement which provided the advantage of exempting the bulk of travel costs from an indirect cost assessment.

b. Human Subjects. The UCB Committee for the Protection of Human Subjects received and reviewed all approvals and materials relating to research work involving human subjects conducted by the collaborating CRSP US and overseas institutions. On the Committee's advice, ME developed instructions and an umbrella protocol for use by collaborating PIs in establishing protocols for their own campus committees' review.[2]

4. Data Management Activities

All three projects were involved in preliminary data analysis during this period, and worked throughout to coordinate data management plans in conjunction with ME. At a December 1982 meeting convened by Dr. Balderston and ME Data Manager Patsy Fosler, the US data managers and overseas representatives of the data management teams further discussed plans for parallel data management activities across projects. Other accomplishments at this meeting included the discussion of a method of organizing the material for form development, the initiation of plans to develop the specific lists of hypotheses and variables, and the addressing of important data editing issues.

ME completed its planning for centrally archiving project data and offered advice to projects on further refinements of data management functions, outlined as follows:

a) translate protocols into the required data capture instruments;

b) establish the pattern and timing of data flow from host country to US institution to Berkeley, and establish alternatives;

c) prepare data capture instruments for data collected outside the main system (such as centralized laboratory determinations), and establish data flow processes;

d) set up quality control procedures and data cleaning operations;

e) establish a data management computer system and methods of interactions between host country, US institution, and ME;

---

[2] Initial approval was given 10/31/83, conditioned upon submission of an umbrella protocol to be developed for the Program as a whole. The umbrella protocol was submitted on 11/14/83 by ME, and the Human Subjects Review Committee provided no further information or action on it. In 1984 and again in 1985 the approval came up for renewal in mid-year, information was requested of the ME office, and the approval duly extended. This process stopped in 1986, since the field phase was, in every case, reported to us as having been terminated in the first half of that year, and the CRSP was therefore no longer dealing directly with human subjects.

f) work closely with statistical consultants to determine analytical strategies for examining research questions;

g) establish a plan for sharing techniques and other data management information between projects;

h) develop data book scheme for circulation to and interactions with PIs, to determine PI data interests and forms for request of computer analyses;

i) develop techniques and processes which would result in the development of a final multi-country data set for parallel inter-country analyses and eventual public use.

In September 1983, the Berkeley data management team prepared an official document, "Manual on Data Collection & Data Flow," which incorporated these protocols. This manual was based on design decisions taken by the SCB at its May, 1983 meeting. Variable sets, organized by function and target subject, were to follow timing and methods as determined by SCB decisions.

C. Year Three: 1 October 1983 - 30 September 1984

Overview

The major events of this third year of the CRSP were marked by the initiation of Phase II research data collection and the negotiation of a new administrative organization for the program, including the reorganization of the role of ME, the reformulation of the research design, and the decision to abandon the search for a Program Director, and instead to create and fill the position of Program Coordinator. A major, unexpected development for the Kenya project was the onset of the drought in that region of Africa. It is notable that having selected three "stable" countries, each one had now presented major "disruptions" that acted to affect implementation of the program or would influence interpretation of the research results.

1. Country Project and Field Activities

With the new research design for Phase II having been approved by the EEP in August 1983 and subsequently accepted by USAID, the projects proceeded to Phase II in the fall of 1983 and winter of 1984. Phase II data collection had begun in all projects by January 1984. Before the data collection on core study households and subjects could commence, however, each project had to complete the organization of its field office operations, select and train additional staff, refine all instruments and methods of collection, and choose target subjects. Site visits were made by Drs. Calloway and Balderston to review the projects' progress with Phase II implementation and to identify areas requiring PI attention; observations and recommendations were made to senior project staff at an exit interview in each case, summarized in travel reports submitted to USAID, and copied to PIs. Because of delays encountered in staffing and establishing procedures, data flow had not been initiated by the time of the SCB meeting in May 1984, although some case study information was made available in preliminary form.

2. Research Design Changes

The projects were obligated to follow the EEP-approved core design until specific approval for modification was given by the SCB, and ratified by ME. In the fall of 1983, ME assembled and circulated to all members of the SCB the proposals for changes that had been received from individual projects, or which had been initiated by ME after detection of design description errors. A response to the first set of proposals was sent out. Some had been unanimously approved by the projects and hence become design modifications. Other design change requests were deferred for discussion at the May 1984 SCB meeting in Berkeley.

At this May meeting, the following matters arose with reference to the primary independent variables:

a) The Egypt project reported difficulty with its food intake methodology. The SCB agreed that the project should attempt a new approach, on which the project was to report at the next (September 1984) SCB meeting.

b) The Mexico project presented information to suggest that the approved method of collecting morbidity data was unreliable for the Solis setting. A new method was proposed and approved by the SCB. ME objected, however, on the grounds of scientific design, and exercised a reserved veto power, notifying the SCB that ME would seek advice from outside epidemiologic consultants. The advice was sought, and confirmed the ME concerns. The Mexico project was advised to continue with the morbidity protocol that had been submitted to the EEP. (In point of fact, this directive was not followed.)

c) Project psychologists had sent to ME a series of proposed changes in the approved protocols for cognitive studies. These involved what appeared to be differences in instruments and tests across projects. It was agreed that projects would decide which subtests and additional measures would be performed, and that projects would advise ME of their decisions.

d) In order to provide an opportunity to draw back into the study those areas that the EEP had recommended be deferred from the first year of Phase II, project-designed case studies were to be made at each field site to provide information for the design of studies for inclusion in the second year. The major proposal for add-on study was related to activity of the lead female, using time allocation as the basic test instrument. It was agreed that projects should refine both the constructs and instrumentation of this area of investigation on a project-specific basis, and, where possible, contribute information to the activity categories designated by the SCB.

In 1983, projects submitted brief descriptions of non-core research projects that they wished to initiate. The preliminary statements were returned to projects together with comment by ME on each proposal, and in view of clear pressures on the CRSP budget, projects were urged to seek outside funding for meritorious proposals. In 1984, it was decided that projects could submit

proposals for non-core research to the SCB for review and approval, to be accompanied by specific budget plans showing that core commitments and the additional protocols could be funded from existing sources.

In September 1984, ME completed and circulated to projects a revised manual delineating a Basic Data Set, reformulated according to design changes proposed in May, 1984 and adopted in August, 1984, constituting the essential minimum variables to be collected by all projects (the so-called "Red Book"). Projects were to declare what additional variables were included in project-specific designs, but without omitting any items from the Basic Data set.

3. Management and Administration

During this third year of the program, ME facilitated the implementation of the newly approved design through visits to the country project sites and the organization of SCB and other meetings, and fulfilled various monitoring functions from the UCB office. These included assisting in the resolution of personnel problems within projects, executing certain USAID mandates, and recruiting consultants to assist in scientific decisions. Pressures that had developed during the initial implementation of the CRSP relaxed with the formal commitment by US investigators to meet all core data collection and analysis requirements within the planned total budgets, and with the ensuing negotiation by ME and the SCB of a new administrative organization and partitioning of responsibilities.

a. CRSP Reorganization. One of the principal changes in the reorganization that took place during this year was the SCB's assuming a greater responsibility, via the creation of an SCB Executive Committee, for implementing the agreed-upon research design within individual projects, and for providing a forum in which collaborating projects could present and discuss mutual problems and experiences. The objectives of this shift were several:

a) to give increased autonomy to projects, as opposed to ME, in the design of country-specific research;

b) to assign primary responsibility for decisions to the level at which they were made;

c) to reinforce appropriate peer review procedures for the assessment of the merit of the decisions;

d) to differentiate the role of ME from the projects, the SCB, and the IC.

Within this scheme, the Management Entity was to consult and advise upon but not be responsible for scientific decision-making, with the proviso that the basic core design was fixed (i.e. not open to revision). The "Red Book" represented the definition of the core requirements. ME continued its responsibility for the budgetary and contractual aspects of administering the CRSP funds, as the primary USAID Grantee, and was vested as a formal voting member of the IC. ME also retained the following principal responsibilities: to provide a structure for review and approval of research proposals, assist

projects in the implementation of approved designs, monitor the projects' scientific performance and design effectiveness on behalf of USAID, determine and implement interproject design and data analyses, and collect and archive the research data.

Under this new arrangement, the authority and responsibility for elements of scientific design outside the basic core were vested in the SCB, with ME advising the Board of any concerns it had about the decisions. An important implication of this arrangement was that projects, individually and collectively, would accept responsibility for the scientific merit and financial feasibility of their decisions.

Key to the altering of the ME role was the creation of a new position - that of Program Coordinator. The renewed search for Program Director in 1983 had closed without success, and, during the May 1984 SCB meeting, the PIs requested that a Program Coordinator be sought rather than a Program Director, emphasizing the perceived need for help to the projects in coordinating, rather than providing direction for, their work. Following approval by USAID, an active search was begun in July 1984. Dr. Hubert (Hugh) Horan was selected to fill the position, and began active duty in September of that year. It was agreed that his responsibilities would include assisting the SCB in its task of further defining and achieving consensus on research questions, working with the projects to facilitate and implement the SCB-approved design, and consulting with the rest of ME to maintain and monitor the progress of the collaborative data sets.

b. Fiscal Management and Planning. Target total expenditures for each of the three projects and for ME were established at the time of the CRSP initiation. But as the start-up of Phase II had been delayed, concern arose as to whether these sums would be adequate. ME was repeatedly and firmly advised, however, that USAID would not be willing to provide additional funds. An implication of the "fixed price contract" was that projects would be required to complete the required work (design, data collection, data analysis, and reporting) within the allotted budget. As a precautionary step, ME directed that research chargeable to the USAID funds be limited to core activities (as noted above), and in October called for revised project budget projections through the completion of the CRSP (at that point, anticipated in 1986). ME reviewed these budgets, which called for about $1.2 million more than USAID had committed, and offered suggestions on changes and savings that might be effected. At the May 1984 meeting of the IC, the project representatives gave formal assurance that all required performance elements of the CRSP would be completed within the overall target budgets.

With this assurance, ME felt it appropriate to relax its close scrutiny of expenditures. It was decided that henceforth, Management would monitor expenditures not in terms of specific items (except with regard to USAID policies and similar contractual requirements), but rather in terms of the categories of expenditures in the approved budgets submitted by the project institutions. If expenditures exceeded those in the budget, the projects would be asked to review them in relation to their commitment to complete all work within the target budget. It was intended that this process would allow

PIs the flexibility to exercise their judgment independently, and to respond more directly to their own institutions in justifying expenditures.

4. Data Management

The data management staff at ME organized schedules and associated variables lists based on both the new, EEP-approved design and on instruments provided by the projects. While it had been hoped that identical forms would be utilized for data collection across projects, this was not feasible; projects found that different modes of questioning and observing were necessary, and that each would have to develop its own instruments tailored to the cultural conditions specific to its site. Moreover, since projects were attaching different weights to the various aspects of the core design, questions were formulated with differing degrees of detail. Nevertheless, because of the need to examine research questions across the three projects, ME required that the basic data set be collected in each project. For this purpose, a list of the variables that presented the required core data were incorporated into the Data Management Manual.

The Manual presented a set of instructions on the use of SAS, a list of SAS files and variables, and the methods of data flow from the field to the central university archive. As UCLA data managers objected to the use of SAS, ME volunteered to do the SAS entry, when necessary, provided that the Kenya project clearly identify the data file format on its tapes. Methods of quality control--which involved detailed plans for data flow, error correction, documentation, and final archiving of the official files--were also included in the Data Management Manual.

During this year, data management systems were established in each field location as follows: In Kenya, a commercial establishment in Nairobi was employed for data entry, with verification to be undertaken in Nairobi, and a second set entered at UCLA. After several months, it became apparent that there were too many errors in the Nairobi data entry, so the job was shifted to another firm during the summer of 1984. In spite of a major effort on the part of UCLA, it was not possible to arrange for data management within the collaborating institution, the University of Nairobi.

In Mexico, an agreement originally had been made for data entry and processing to be undertaken by IBM of Mexico and the Mexican Ministry of Public Health. Complications developed in these arrangements, however, and in May 1984 the project PIs decided they could not be completely dependent on Mexican facilities other than their own for the capture of project data. Arrangements were then made to carry out data entry on micro-computers at the field site, and to have data transferred to the University of Connecticut on diskettes for entering into mainframe data files.

In Egypt, several alternatives for data entry were investigated, and arrangements were made for this activity at Al-Ahram by July, 1984. Prior to this date, all data had been sent on paper to the University of Kansas for entry into computer files.

Although each project's data management system was distinct, in keeping with CRSP policy each had made arrangements to maintain a complete set of computerized data in the host country for use by investigators in that country.

As part of the June 1984 CRSP reorganization, ME's role in the CRSP data management and analysis were revised. It was agreed that ME would be responsible for cross-project analysis of the collaborative data sets (but not project-specific data sets), with advice and direction being taken from the projects through the SCB. A part of this process was to involve testing analytical approaches during the course of the CRSP, provided, of course, that data flowed to Berkeley on schedule.

5. Project Performance Monitoring

As part of its role in monitoring the performance of the projects during this phase of the CRSP, it was ME's responsibility to ensure that data were flowing from the projects to Berkeley, and to keep the SCB apprised of research progress. It was apparent, even at this stage of the project, that projects saw themselves as being unable to adhere to revised, previously planned data flow procedures and schedules. At the May 1984 meeting, PIs agreed on a schedule of data flow from the projects to Berkeley, understanding that compliance with the schedule would be a measure of performance.

Another element of project monitoring was the review of site deficiencies as detected in the course of field visits made by members of both ME and the EEP. The EEP visits, arranged by ME, were as follows: Professor Philip Payne and Dr. Gerald Keusch to the Egypt site in July 1984, and to the Kenya site in September 1984; and Drs. Marion Radke-Yarrow and Donovan Thompson to the Mexico site in August 1984.

D. Year Four: 1 October 1984 - 30 September 1985

Overview

Notable progress was made during this period of the CRSP, which was a watershed year, being the final of two years of field research, and the first of several years of data analysis. With field demands still at their peak, it was necessary to turn attention simultaneously to considerable efforts in data management and analysis.

After the CRSP reorganization, the major role of ME's Program Coordinator was one of fostering a supportive, decentralized style of management, and of increasing the amount of personal contact with and between the projects. To further these ends, three major meetings were conducted during this period, two of them official SCB meetings, and one an informal meeting of PIs in Cairo, Egypt.

1. SCB and Other Meetings

The year began with the end of the September/October 1984 SCB meeting in Berkeley, at which the EEP reviewed the CRSP's progress. The meeting was

dominated by demands for and discussion of analytical strategies for the data being collected; the EEP felt keenly that analysis on a preliminary basis should have already begun, and that the absence of clear strategies for future analysis needed remedying as a first priority. The members expressed strong feelings that the initiative for project analysis should be at the project level. The panel delivered a strong mandate that the projects provide proof of analytical capability within approximately six months. The SCB agreed, and some strategizing for accelerating the data management and analysis process was initiated at that time.

ME then hired a program-wide statistical consultant, Professor Stephen Fienberg of Carnegie-Mellon University, and an ME statistical consultant, Professor Steve Selvin of UCB. Dr. Balderston (former Deputy Program Administrator) moved into the Data Management component of ME, where she continued the development of data analysis and management under the title of Management Coordinator for Data Systems and Analysis, a position which she retained until leaving the project in May 1986.

In January 1985 the PIs organized a workshop in Cairo, Egypt on data management and analysis, which was attended by US and host-country PIs, Drs. Horan and Beaton from ME, and many of the Egyptian senior investigators involved in the CRSP. This meeting marked the first definitive turn of attention from field problems to analytical strategies. In addition, Program Coordinator Horan took the opportunity after the meetings to visit the Egyptian and Kenyan field sites; Professor Beaton revisited the Kenya site a few months later as part of a trip sponsored by another agency.

In February/March 1985, an SCB meeting was held in Berkeley, concurrently with both IC/FG meetings and formative meetings of the Data Analysis Group, a task force with the four senior statisticians as core members, and data managers and investigators participating as needed. The SCB discussions dealt almost entirely with questions of data analysis. At this meeting, a formal proposal was put forward by the data analysts and PIs, and was approved by SCB, that data analysis be conducted at the project level. ME should serve as a catalyst and coordinator of an expanded statistical group but should not take the initiative in primary analysis. ME would retain the responsibility for interproject interpretation of parallel analyses.

The IC/FG meetings focused at length on the projected cost overruns in the Egypt and Kenya projects. At this group's meeting in May 1984, each of the three projects and ME had committed to living within its projected five-year budget, which posed no special problems for ME or for the Mexico project, but did pose very considerable problems for Egypt and Kenya. These projects' revised planning budgets still projected a total cost overrun of some $460,000 combined, in spite of the considerable economies instituted. The IC recommended that the Kenya project find another $98,000 in budgetary economies, and that Egypt find another $102,000. Mexico, which projected a very slight surplus, was asked to find $100,000 in potential budget cuts, while ME volunteered $103,000 in cuts. These funds were then made available to projects based on need and demonstrated ability to use them effectively. In order to make its share of additional funds available to the projects, ME substantially reduced and reorganized its own budget, at the same time

offering to bear the costs (including airfare, per diems, and meeting arrangements) of all SCB, statistical, and data analysis meetings. The IC authorized ME to reassume greater fiscal responsibility on budget planning and execution.

In late April, a meeting was held in Queretaro, Mexico, involving the full EEP, PIs from each project, and representation from ME, including its principal consultants. At that meeting, the EEP reviewed both progress and the specific process of coordinated project-level analyses that had been set in motion and gave enthusiastic approval. Even in the face of demanding and competing field research obligations, the PIs were able to present sample analyses of data together with the beginnings of a planned, coordinated analytical strategy. This formalized the transfer of responsibility to the projects.

2. Project Evaluation

a. Egypt Project. Serious problems were encountered during this initial period of data collection. An investigation of the food intake data and methodology was undertaken by project staff in conjunction with Professor Beaton because of conflicting interpretations of validation trials that had been conducted, and because it appeared that in the research data sets energy intakes per kilogram were unusually low in adults, particularly males, and unusually high in toddlers. The problem was addressed with a complete review of the validation trials, a statistical analysis of the research data sets, and a visit to the field site for discussions with Egyptian investigators. In addition, the project completed a detailed examination of the records and computation algorithms. By April of 1985, specific errors had been detected and corrected, and it was concluded that there was no specific reason to fault the food intake methodology of the Egypt project, although concern about the unusual distribution of toddler food intakes continued, and the EEP expressed serious reservations.

A second problem was the amount of missing data. The issue was initially addressed at the Cairo meeting in January 1985, and reviewed at length during the subsequent February/March SCB meeting.

b. Mexico Project. In April, the PIs informed ME that a revised food intake methodology had been instituted in the field as of January 1985. The change, to more professional interviewers, was perceived by Mexican personnel as necessary to maintain project-household relationships. ME expressed concern about this midstream change and requested information on validation studies. In June 1985, ME commissioned a group of consultants to advise both the University of Connecticut and ME on appropriate measures to investigate and suggest remedies for problems ME perceived to exist, i.e.,

i. Data flow from this project seemed particularly slow and unsatisfactory.

ii. There was at least one field site management problem judged likely to have created a validity and reliability problem, in ME's view. (This

view was not shared by the UConn PIs, but ME judged the point in question important enough to be worth a second opinion.)

iii. Revised estimates of missing data from this project were also considerably delayed, and, in the view of ME, critically needed for the analytical planning process.

Professor Donovan Thompson, biostatistician and EEP member, was asked to chair the consulting panel. Dr. Raymond Neff, Assistant Vice Chancellor for Information Systems and Technology at UCB, served as the ME representative, and Dr. Edmundo Alvarez, Medical Officer at the Panamerican Health Organization (PAHO), was chosen by the UConn/INNSZ PIs. The consultants reported that while the data management system in operation at the time was adequate for the present, it had some unusual vulnerabilities; and that ME's concern about certain aspects of field science (i.e., missing data, adherence to agreed protocols) was indeed valid, and another field science concern (relating to psychology testing) had arisen during the investigation. Although UConn sent a report on missing data to UCB during the course of the consultancy, consultants agreed that more precise information on the exact nature of the missing data would be needed for the analytical process. ME strongly recommended that the project's PIs investigate the circumstances and persons involved in the field science questions detailed in the report, as well as field site management generally, and that steps be taken to locate and remedy any problems identified.

c. Kenya Project. It became apparent early in this period that the drought experienced in Kenya would severely affect food availability in the field location. Project staff naturally assumed an ethical responsibility to assist the study villages in obtaining food relief (and later, seed supplies), which unavoidably had some cost to the project. The effect of this food crisis on the ability to test research hypotheses could not be predicted, and project staff paid particular attention to documentation of the food situation and responses to it.

3. Data Management and Analysis

Data management and analysis activities picked up considerably during this year. Data files, as they became available in each project, were passed along to Berkeley for archiving and conversion into SAS format when necessary, with files being combined within projects across time and topics. The Berkeley data management group also scanned the data and performed simple summary statistics computations. Investigations of several relationships were undertaken to examine the effects on food intake produced by a variety of factors, including variance of individual food intake, entry date for toddlers, the sequence of food intake measurements, and the Ramadan fast.

In June 1985 plans were made for a trial of parallel analyses of target toddlers across projects. ME prepared files, using Egypt project data, for two three-month periods (beginning at 18 months of age, and extending for six months), and issued memoranda listing a recommended set of variables needed to carry out contingency and initial regression analyses. These included toddler food intake, anthropometry and morbidity, and maternal intake. The variable

definitions were refined based on feedback from the projects, and from ME consultants. In July, sample contingency tables were circulated and proposed as a model that individual projects could alter and refine to suit their own data analysis needs. ME then ran prototype correlation and regression analyses to serve as guides for all three projects.

Dr. Suzanne Murphy was hired as ME Data Manager and Assistant Research Nutritionist in September 1985.

E. Year Five: 1 October 1985 - 30 September 1986

Overview

The important events of this year involved two large meetings, renewed attention to the PCG, serious fiscal difficulties for one project (see section IV), and an audit of the CRSP's scientific performance by USAID.

In effect, Phase II, the field phase of the project, was substantially completed for all three projects during this year, although both the Mexico and Egypt projects continued in the field in order to complete pregnancy outcome studies. Data flowed from all three countries to their corresponding US institutions on an increasing scale, while significant attention was given to scientific and technical problem-solving. The shift of focus from field research to data analysis was completed during this time.

1. SCB and Other Major Meetings

a. Data Analysis Group. In October 1985 an important Data Analysis Group (DAG) meeting was held in Storrs, CT, in which several (but not all) project data managers participated. Although the DAG, designed to provide a forum for discussion and decision-making on analytical issues, was originally composed principally of the project statisticians and an appropriate representation of PIs, it became apparent early on that the data managers were definitely needed as a part of the process, for several reasons. They had been dealing more closely with data flow, cleaning, and archiving than any other group in the CRSP, and some decisions would have been impracticable or too costly if made without their active participation. The CRSP data managers were also a very senior group of collaborators, all with considerable experience, frequently with doctorates in very relevant disciplines. Their general advice proved to be quite consistently worthwhile, as they were technically knowledgeable as well as experienced in the practical realities and difficulties of data management. The more critical aspect of the dialogue was an increasing sense of frustration among project statisticians, due to the delays in data flow and "data cleaning".

b. Policy Consultative Group. In November 1985 Dr. Forman of USAID passed through Berkeley, and in consultation with Program Co-Administrators Calloway and Rosberg, worked on bringing the PCG to active status, to foster the application in developing countries of the research findings becoming available at the end of the CRSP. Between November and February 1986, a broad process of consultation resulted in final nominations for the PCG, as well as

a working understanding of the many basic organizational and administrative details affecting it.

At the February 1986 SCB meeting, discussions were held on how to manage the policy aspects of the CRSP. Projects reported being under some pressure, particularly in Egypt and Kenya, to reveal the results of the CRSP at once, before the data had been adequately analyzed and refined by the investigators. It was agreed that the CRSP has an obligation to try to provide an organized process of policy dialogue on its results. It was also recognized, however, that the projects should avoid creating an expectation that the CRSP mechanism could address the problems that the results may identify. Further agreement was made as to the PCG's function as a mechanism for continuing guidance and stimulation of the dynamic process of policy analysis and formation.

c. SCB. The February 1986 SCB meeting at UCLA continued the work of the DAG as well as of the SCB and other CRSP groups. The meeting was dominated by scientific and analytical discussions focusing on the clarification of priorities, detailing of the process for testing the basic hypotheses, and several individual scientific questions.

2. CRSP Performance Audit by USAID

In March 1986, two USAID representatives conducted a performance and productivity audit of three CRSPs, including the Nutrition CRSP. The object of the audit was to determine what the original purpose and intent of the research had been - as witnessed by the grant documents - and to form an opinion as to whether or not what had been promised by the grantee was being (or was likely to be within the allowable time period) delivered or not. Both reviewers emphasized that having a properly archived and adequately documented data set was an expectation of critical import to USAID.

3. Financial and Contractual Arrangements

Because of renewed financial planning efforts on the parts of all the projects and ME, the CRSP found itself prepared to enter its final Period VI with significantly more funds than projected the previous year. Two factors were behind this unexpected surfeit: one was that because data delivery was behind schedule in two of the three projects, costs associated with data delivery were therefore also lagging by up to three months. More significantly, the projects had voluntarily increased their cost-sharing loads and cut costs through a number of intelligent fiscal choices. The Program Coordinator's office had reduced its staff considerably, as had the ME data management team.

4. Data Management and Analysis

Data management activities at ME focused primarily on archiving during this period, as well as on continuing the data flow procedures established previously. As data tapes were received from the projects, programmers routinely copied them, converted selected variables to SAS as necessary, merged individual data files as required, made selected checks of the data quality, and generated preliminary reports on the contents of each. By the end of September 1986, the majority of the field data had been received from

all three projects, and the status of the files tabulated. As noted later, problems began to emerge with regard to discrepancies between field data archived in Berkeley and "cleaned" data sets being developed at the project level. This affected ME's ability to generate summaries and other monitoring procedures that agreed with project analyses.

In addition, documentation related to the data archive was assembled, including manuals, survey instruments, variable lists, and detailed mappings of variables from survey instruments. This documentation, subject to any revisions, became part of the permanent record of the CRSP, for all researchers' use.

Attention was turned during the latter half of the year to the specific format that the final CRSP data archive would take. It was recognized that the archived files would necessarily encompass two aspects of the final data sets: the similarities between the original design and the actual data of the three field projects, and the dissimilarities imposed by the realities of the actual field work. Users of the CRSP data would need to clearly understand these similarities and differences in each variable and, in some cases, changes in the collection methodology within a variable over time. Thus, it was critical that the archived files be fully documented.

Three levels of file structure were proposed for the archive: the field data sets, which would reflect the data as originally entered onto tapes (usually in the host country); data sets archived by the US institutions, with necessary changes made by ME staff for structuring files into the Basic Variable format; and analytic data sets, created either by the projects or by ME for the purpose of performing specific analyses.

An important aspect of ME's archiving task was the description and documentation of the contents of each data set. To this end, the ME data management group began working on descriptive statistics of some of the more important basic variables during this period, including food intake, anthropometry, and morbidity.

Finally, in addition to its work on the final archive, the ME data group also performed selected analyses on research hypotheses suggested by project PIs, and investigated several food intake issues for the Egypt and Kenya projects.

F. Year Six: 1 October 1986 - 30 September 1987, extended to 31 January 1988

Overview

During the first half of the CRSP's final year, ME devised plans for the generation and delivery of Final Reports to AID. Arrangements were made for the archiving of essential CRSP files, and for printing and microfilming the Final Reports. Project subcontracts were extended to complete submission of data files, documents and reports for ME final processing. Program Coordinator Horan resigned effective November 6, 1987 and by December 1, ME staff was reduced to three full-time equivalents plus the Program Administrator and key consultants. CRSP officially closed its operations on

January 31, 1988, 10 years after the NAS workshop that recommended the research be done.

1. SCB and PCG Meetings

There was a major and final joint meeting of the SCB and other component groups in November 1986. Several different agendas were addressed at this series of meetings (which included EEP representatives, statisticians, psychologists, data managers, IC members, ME, and the SCB): the normal scientific and business discussions of the SCB; establishment of basic approaches to analysis of the psychology data by project psychologists, statisticians, and PIs; an informal EEP evaluation of the CRSP's progress since the May 1985 meeting in Queretaro; review and adjustment of the analysis of the morbidity procedures for toddlers; and discussion of the timing and composition of the Final Reports to USAID.

The evaluations made by the EEP members at this meeting were positive and encouraging. Professor Thompson, one of the two EEP members present, expressed his belief that at least partial answers would be found to the CRSP's ambitious research questions. He also commented that it was urgent that the CRSP find more funds to continue data analysis as a group, but under a different administrative structure.

The PCG met in Berkeley in July 1987[3] to discuss the process of identifying policy implications that might emerge from the CRSP's research findings. EEP member Sol Chafkin served as Chair of the meeting, which included overseas and one US PI from each project, as well as policy-oriented analysts from the Egypt and Kenya projects, and Co-Program Administrator Carl Rosberg. These discussions represented the first organized invitation to all PIs to focus on policy implications as a means of clarifying for policy decision-makers the significance of the research findings and the choices they face with respect to keeping or modifying existing policies, or formulating new ones. The PCG noted that the CRSP, in relating energy intake to specific functional consequences, was making a major conceptual departure from earlier research dependence on anthropometry, and that the research and analysis had produced a rich mine of data that would be of use to assistance agencies in defining policies and programs for local areas as well as for whole countries.

2. Budget and Contract

In November of 1986, ME submitted a formal request to AID for another extension of the CRSP, through January 31, 1988. The request was granted, and once revised, an extended grant contract with some new provisions was received from AID in early March 1987. It then became clear that careful research was needed to ensure that both new and old requirements would be met. ME turned its attention to the area of equipment acquisitions, as changes in the contract requirements rendered it necessary to reconstruct ME records of the host countries' authorized purchases, and to begin planning ahead for the

---

[3]Funds for this meeting were not supplied from the CRSP budget, but from a contract placed with the Western Consortium for the Health Professions.

disposition of equipment. Project subcontracts were extended to November 15, 1987, with reallocation of funds from the ME budget, to assure completion of the data archive and its documentation.

3. Project Evaluation

Project staff turned their attention to completion of three principal tasks: a) completion, including cleaning and documentation, of the data archive and its transmittal to Berkeley; b) computation of descriptive statistics for all principal variables and analysis of findings relating food-energy intake to reproductive outcome and to toddler growth, morbidity, and cognitive development; and c) preparation and submission of the final written report. First drafts of project reports were received for review by ME during July and August 1987; drafts were returned with comment to projects within two weeks of receipt. Revised final reports were submitted by the Kenya project in early October, the Mexico project in mid-October, and the Egypt project in late October. Each project arranged for formal presentation of its findings to host country officials so that the governments would be informed of the relevance of the CRSP data to country programs and policies. The Kenya project meeting was held in September 1987. The Egypt project presentation will take place in March, 1988; the Mexico presentation remains to be scheduled.

4. ME Activities

ME undertook many scientific and evaluation tasks during this period to verify the completeness and quality of projects' data files and to chart a plan for future analyses of the data. The outcomes of these efforts are described in Part IV, below. Tasks of archiving, microfilming, editing, and reproducing reports required much time and attention.

Copies of all data archived with ME were distributed to US and overseas PIs and to Professor Beaton at Toronto in December, 1987. ME submitted its own final report in draft form for review by project PIs in December, 1987. With submission of that report and the archived data to USAID, the work of the CRSP was completed.

G. Future Plans

Much work remains to be done with these rich data sets. PIs from the three projects have submitted a joint proposal to NIH for analysis of data on functional anthropometric indicators. Each US PI also has taken initiatives to secure funding for continued collaboration with the project host-country personnel and institutions. Professor Beaton put project PIs in touch with representatives of the World Bank as a potential source of funding certain analytical work that the projects could not undertake within the CRSP.

A few ME principals and staff plan to devote additional time and effort, to prepare a CRSP overall program report, summarizing project findings and exploring these in a broader, cross-project context. That report will be available in June, 1988.

The complete archive will be in the public domain and available to other scientists by request to USAID after January, 1989.

IV. FISCAL SUMMARY

CRSP funding was authorized by USAID in 1981 and there were no subsequent revisions to the total allocation. Shortfalls in project needs were met, as far as possible, by internal reallocation between projects and ME, as follows:

9/81 Planned Allocation	12/87 Actual Allocation	Total Adjustments
ME:		
$2,348,976	2,173,788	<175,188>
Program Total:		
$9,451,024	9,626,212	
$11,800,000	$11,800,000	
Program Detail:		
Egypt:		
$3,353,101	3,390,106	37,005
Kenya:		
$3,084,306	3,245,989	161,683
Mexico:		
$3,013,617	2,990,117	<23,500>
$9,451,024	9,626,212	-0-

Problems with budgeting and expenditures have been mentioned in Section III, above. Modest rebudgeting from the Mexico project was made possible because of temporary advantages from dollar-denominated operations in a severe inflationary situation in Mexico. The Egypt and Kenya projects were faced with substantially higher travel costs than Mexico because of distance from the field sites. Egypt supported a more complex management with three US institutions involved. Kenya had higher field costs, due to much greater expatriate presence than in the other projects. ME cut as many corners as possible to release funds to the projects, including substantial savings from subcontracting for travel.

Significant economies were effected late in Phase II due to delay in data delivery in some cases, and the loss of analytical staff in other cases. The end result was that fewer analyses were done in this period than had been planned. It was a savings no one would have chosen, but it did leave the CRSP

with more money available at the end of FY86 than had been anticipated. Thanks to the availability of these funds, extensions of time were secured from USAID, eventually extending the CRSP through January of 1988 -- sixteen months beyond its original termination date, a very considerable achievement for a project so tightly budgeted and through so many inflationary years.

A requirement of funding under Title XII as administered through BIFAD/USAID is that collaborating US institutions contribute at least 25% of the research costs from their own resources. Exempt from this matching requirement were funds passed through to collaborating host country institutions (US institutions did not collect overhead on funds so transferred), and the Management Entity budget. All collaborating US institutions met or exceeded the mandated matching requirement. The University of California, Berkeley provided significant voluntary cost-sharing as time given to CRSP administration by Professors D.H. Calloway and C. Rosberg, and by maintaining indirect cost assessment at the low level in place when the CRSP began. The University of Toronto contributed the full-time services of Professor George Beaton during a sabbatical leave year, as did the University of California, Berkeley during Professor Calloway's leave periods in 1984 and 1987.

Projects provided some services and support to the communities that so generously participated as subjects of the CRSP research. Outdoor school facilities were improved in Kalama; the Kenya project provided some food assistance during the drought in Embu; the Mexico project upgraded outreach health facilities in the Solis Valley. These modest contributions are the only concrete, immediate rewards to the most important participants in the projects. Some research equipment was left with the collaborating host-country institutions to enable them to continue this line of endeavor.

ME arranged for audits of the overseas components by contracting with the firm Deloit, Haskins, and Sells. There are no unresolved audit issues.

V. ORGANIZATION AND MANAGEMENT ASSESSMENT

The purpose of this chapter is to examine problems associated with the administration of the CRSP: the problems that might have been avoided or significantly reduced by better planning or by a different organizational design. Although hindsight is usually easier than foresight, hindsight, if appropriately interpreted, can be an important basis for foresight on another occasion.

Much of what can be said in this regard will necessarily contain a great deal of subjective judgment. To minimize this risk, four people who had played central roles in the central administration of the CRSP met together to discuss the matter. They drew up a list of points on which there was some measure of agreement, conceding that it was unlikely that there be many points of total consensus On one point, however, there was no difference of opinion: before the initiation of the CRSP, inadequate attention was given to questions of management, organization, and infrastructure. This is not to deny that considerable thought and planning went into the development of the original CRSP structure, nor that existing CRSP federal guidelines at that time offered some guidance. But ME maintains that inadequate attention was paid to organizational and structural issues in setting up this CRSP.

Very early on in the history of the CRSP, ME proposed to USAID that there be a study of the proposed CRSP organizational structure. Professor C. West Churchman, an eminent authority on systems and organizations was available and interested in conducting a study; the sum involved was relatively small. USAID declined to fund this additional effort, and ME believes that this CRSP--and possibly other CRSPs--have suffered from this lack.

A. The Formation and Selection of CRSP Component Projects

The character and composition of this CRSP is a legacy of the planning strategy (4). In retrospect, the established process of forming a CRSP was unnecessarily cumbersome and expensive, and unsuited to the organization of a study to be replicated in diverse locations.

The first organizational step was to investigate independently both overseas and the US institutions that would be eligible to participate in the CRSP. The second was to invite proposals against a specified research topic. This almost total canvass was costly, time-consuming, and produced few solid results. The strongest US institutional links were overwhelmingly with one geographic region, Latin America and The Caribbean. As it was undesirable (even scientifically unsound) to implement the CRSP in only one cultural/agricultural area, some reaggregation of investigators and sites was required. (Purdue in Egypt with University of Arizona and University of Kansas, and UCLA in Kenya with UCB). As in life, forced unions imposed an element of stress.

B. The Selection of ME

All things considered, it is the belief of those who lived through the early years of the CRSP that it would have been better not to have had the ME at UCB. There are several reasons for this. We grant that it is never possible to know what might have been, and therefore our reasoning here has necessarily something of the speculative about it. Nevertheless, the feeling is that the double bind of Professor Calloway's early close involvement in the CRSP and her demanding office as Provost made for a very difficult combination of circumstances.

An uninvolved institution would have been able to begin with more independence, and with the ability to supply a Program Director. Despite two nation-wide searches for a Program Director, UCB was unable to attract an appropriate person with nutritional science background. Some have felt that this was a serious impediment, particularly in the early stages of the field research effort and in the arbitration of differing scientific points of view.

C. Collaborative Design Versus Collaborative Theme

In most CRSPs, the "collaborative" refers to a looser set of relationships than was thought desirable in the Nutrition CRSP. The Nutrition CRSP had only three projects, all of them very large, with a critical importance attached to the comparability of results and hence, comparability of methods and design. Since many of the problems that might be cited relate in one way or another to the issues of a coordinated or replicated design, one must ask whether, in retrospect, such a design is feasible or even desirable, given the unique features of each country site (and the unique interests of each cluster of investigators.)

The obvious strength of the adopted design was the potential for cross-project interpretation and hence for development of both setting-specific and generalizable inferences that could be expected to have greater policy relevance.

The perceived weakness of such a design is that it inhibits opportunistic, country-specific/investigator-specific ventures; such ventures could exist, for example, in a centrally-reviewed and scientifically-monitored set of "independent" projects working toward a predefined theme of interest. (This might be seen as akin to the other CRSPs). The trade-offs between the two approaches seem fairly clear: each has its advantages and disadvantages and each serves a different goal.

In retrospect, the Nutrition CRSP had elements of project independence - by intent and by default. The intentional component is to be found in the Phase I work, in which projects, by design, were free to explore their own settings with a common general goal. The unintentional components were deviations from the conceptual design that actually occurred in the operation of field research programs in Phase II.

One might suggest that in part these deviations arose because, in fact, a common base of theory and design had not been established before research data collection was initiated. Each PI had thought about the questions this research was to address, as phrased by those involved in planning, and each group had arrived at a research strategy that addressed the questions from their own perspective in a manner that best would employ their diverse skills. Total agreement about even the main hypothesis to be tested was lacking. To forge a collaborative design, each was asked to forego some special interest or study some aspect in which they had no interest.

A collaborative theme is probably relatively easy to maintain over time. A collaborative design, given the differences in site and direction, proved extremely difficult to enforce over time, especially with limited budget for meetings and site visits.

Some would feel that there should have been significantly more freedom in the conduct of research given to the individual PIs; others believe that there should have been more of a visible central authority. Most seem to agree that things would have gone better had there been either more letting go or more constant central authority.

It seems clear that the Agency agreed with the view, and adhered to it in the face of increasing difficulty, that the three-way comparability was critical; therefore the centralized authority model was the one to maintain and foster. It is the clear opinion of ME that the means to maintain and foster such a model of collaborative design and execution were not given to ME by the Grant terms, a point developed below. (See section E., below)

D. Alternative Research Programming Strategies

Two alternative approaches to the development and implementation of a collaborative design could have been considered but were not. The decision to undertake the principle CRSP research (Phase II) simultaneously in three countries after a two-year period of organization and pilot study in each (Phase I) was driven by constraints of time and money.

Early on it was suggested that a CRSP research project be launched at first in one country, to provide the experience on what could be undertaken effectively in the field and to redefine and narrow the research question. This alternative was never considered realistic, given the process that had been followed in identifying countries and institutions and selecting among competing research proposals. It was difficult to maintain the integrity of research teams even through the period of delayed funding. It would have been impossible to sustain two teams while the third began its work, and cumbersome and costly to have loaded all principles onto one site. More importantly, as the outcome of the CRSP will show, the problems and possibilities are quite different in different sites; despite their previous field experience elsewhere, the US investigators found limited carry-over from it in new circumstances.

The other, more feasible, strategy would have been to allow a research design to emerge after a longer period of development in each place. Phase I might

have been expanded to be a series of self-contained studies of communities, leading to both hypothesis-building and design considerations for a collaborative Phase II. Desirable as this may seem, now or even in the early phases of the Nutrition CRSP, this was never a real possibility for this CRSP. But in the future, a CRSP could be established as a research Program extending over, perhaps, a 10-year period rather than as a multi-site research Project with a finite (short) lifespan and predetermined budget allocation. If there is/continues to be recognition of the validity and importance of the questions addressed by the Nutrition CRSP, USAID should seek out ways of establishing an ongoing, managed, research program rather than sponsoring single projects.

E. Management Structures

Given the decision that the Nutrition CRSP was to be a centrally-coordinated, design-replicated study of three population groups, the overall management structure (Berkeley, US institutions, Country institutions, Field Projects) must be challenged retrospectively in several ways.

A fundamental weakness, originally seen as a strength, was the concept and role assigned to the SCB and embodied in the primary grant and subcontracts. The SCB was given an advisory role that was, de facto and ultimately de jure, a decision-making role. However, the PI members of the SCB had a clear conflict of interest that few, if any, could truly overcome. They were torn between the perspective and opportunities of their own country projects and the perspective and needs of the collaborative program. As a result, they were unable or unwilling to make important decisions, affecting one project or another, or affecting all projects together, that would have been necessary for full implementation of the conceptual structural model of the CRSP.

ME was responsible and accountable to USAID for the conduct of the projects and the delivery of a valid research product, yet it had no tools to assure compliance other than the Draconian measure of discontinuing a project. Other CRSPs with multiple small projects could drop a single project without injuring the program. But with a three-project CRSP designed for representativeness by cultural areas, dietary staples, and crop production patterns, loss of a project would harm the program.

Intermediate steps were not available to ME. Budgetary control, once the CRSP was launched, was relatively absent. Allowable costs had to be reimbursed by regulation, even in the absence of promised results. Line-item veto power was not an option open to management, for all practical purposes. ME could, of course, raise issues and objections, but the PIs were always able to override them if they wished or to ignore deadlines and agreements.

Prospectively, this might imply that there should be strong, central, scientific management with central decisions made by that management and with binding effects on project implementation. The merit of such decisions should be monitored by an informed panel of peers - not acting as policemen or referees, but rather acting as constructive critics, a role the EEP was intended to serve. For such a management scheme to work, there would have to be much closer contact between the central management and the field-level projects than existed in the CRSP. This implies much more travel of central

management to the field and much more control, by central management, of what is being done in the field. If and when field-level problems developed, problems could be resolved with the help of additional resources.

Other lessons were learned in the managerial relationships between the project-specific US institutions, their counterpart Country institutions, and the field project teams. Major conflicts between the goals of targeted research institution-building led to quite different perceptions, across projects, of the desirable degree of country-institution autonomy.

Whatever management plan is adopted for future CRSPs, there are two critical areas where ME is convinced a CRSP would be better served by a more central authority. These are:

a. A voice in certain key hires in the field, especially in the hiring of Field Directors. The ability to direct the dismissal of certain key persons for cause.

b. The ability to intervene promptly and appropriately if hard deliverables are not received at specified stages of the CRSP. Field data on tape would be a possible example of such deliverables, as would reports and manuals.

F. Data management and Analysis

A clear lesson of the present Nutrition CRSP, in keeping with the lesson from many other projects both large and small, is that active researchers interested in data gathering generally accord too low a priority to data management and almost inevitably underestimate the time and funds needed for data analyses.

In future programs, it may be that the central management must take a more active role at the field, country institution, and US institution levels, in establishing and monitoring the components of the data management process. It may imply also that data collection must stop much earlier in the life of the program/project (with clear implications for the definition of research questions and research design) to allow both time and budget for data analyses.

The EEP recommended disaggregation of these duties, the opposite of what is implied above. From the standpoint of institution building, which appealed to the EEP, this would seem to have been sound advice. From the standpoint of operation of a coordinated project, the EEP advice should be taken only if there are substantial resources to permit major duplication of effort.

G. Conflicting Goals in Institution-Building Versus Research Results

A basic premise of Title XII's foreign assistance through BIFAD is that the CRSPs will significantly strengthen the collaborating host-country institutions. We fully endorse this objective, but recognize a conflict of demands. Institution-building will increase the capacity, long-term, of an institution to function well and to conduct research. It would use resources

to train and equip the LDC institution in such a way as to sustain a higher level of effort in the future. It is almost necessarily a future-oriented process, and it very definitely takes time, effort, and money. It implies new personnel and a renewed level of training for existing personnel. Its product would normally be a superior research institution or infrastructure.

Research results are almost necessarily very present-oriented, and are concerned with getting the best short-term performance out of present personnel, resources, and equipment, at the present level of training and readiness. Its product would normally be a superior research paper or other report. There will, no doubt, be a training effect on the institution, but there will also be a heavy diversion of resources from its own organizational needs to the need to produce immediate, or at any rate, very near-term, results.

It is probable that both ends can be adequately served in the same project, but at a considerable cost: it will cost very much more time, and more money.

Within the constraints of time and money afforded the CRSP, it was not possible to serve both masters well. The CRSP had to offer results to the policy planners and to the nutrition community, and that within a relatively brief span of time. Projects and ME did what could be done, en route, in terms of institution-building, but were unable to do more during the present study. Some feel that the CRSP did significant institution-building, occasionally to the detriment of immediate productivity; others feel that the CRSP should have been less concerned with immediate productivity, and substantially more focused on institution-building.

H. Communication

The most difficult aspect of any research is the specification of the research question, in logistically feasible and scientifically valid terms. In most research projects, the second most difficult aspect is the translation of this question into operational activities (design and conduct of the experiment and/or data collection.)

In a program such as the Nutrition CRSP, one might expect that, at least in general terms, the first of these would have been accomplished before the project was initiated. This was the case. However, the translation of the generally-stated research question to specification of actual data needed and then to design of data collection was a much greater task than had been envisaged.

A serious limitation of the present CRSP, and perhaps a very important error in strategy, was an inadequate provision for frequent meetings of investigators--both between US and host-country investigators and between project teams--so that all could evolve together. Obviously, this implies that there should have been a much higher allocation for travel and associated costs - or there should have been a geographically more compact selection of investigators and projects. There was an imbalance between the needs for travel and the funds allocated for the support of travel.

In the Nutrition CRSP, some of these issues may have been exacerbated by the fact that there were major changes in style of management in key staff, and in the interface between ME and projects. There can be no question but that these changes must have exacerbated such communication problems as may have existed.

Conclusion. As a recommendation for the future, it is perhaps important that major effort be placed in the development of a management strategy that offers scientific leadership with continuing individual presence and a budget structure that encourages travel so that PIs and ME can meet both formally and informally at sites appropriate to the topic at hand. Implicit in this recommendation is the associated recommendation that from the beginning of the project there should be a clear definition of roles, responsibilities, and authorities of all major actors, from the most senior level management to the level of the field data collector. There should also be clear agreements on the transfer of data and mechanisms for the sharing of information, and finally, there should be no inhibitions on management's visits to the field sites if a collaborative project is to work. There must be complete sharing of information and understanding.

VI. SCIENTIFIC RESEARCH AND TECHNICAL ASSESSMENT

Field research is never perfect. It is always limited by what proves to be "doable" in a given place, at a given time, with the resources at hand. Because this was a collaborative program with a common design, investigators were largely constrained from substituting new studies for those that presented difficult problems in a specific project setting. The outcome thus is some inadequacy in collection of core data, the degree of inadequacy differing from one variable and one project to another, and some unmeasured loss of what individual initiative might have produced had the investigators been unfettered by the collaborative design.

What this section of the report attempts to do is to provide an evaluation of data archived by the three projects for each of the core independent and dependent variables, as a guide to interpretation of the projects' own reports and in considering future utilization of the archived data. A further purpose in the present examination of the variables on a project-by-project basis is to begin the process of understanding interproject comparisons of analyses. That is, by understanding similarities and differences in the variables, we may better understand differences in findings. Some of what follows will be critical of the completeness and quality of the data. We wish it clearly to be understood, however, that it is ME's judgment that these data sets are among the very best ever produced in similar settings, and that they will prove a rich resource for the broader scientific community.

A. Selection of Communities, Target Households, and Subject Classes

1. Selection of Study Communities

A provisional research design, including criteria for selection of the study community and households, was developed by the US and host-country principal investigators at a meeting in 1979 (4). This was reviewed and endorsed by the SCB pro-tem in October, 1981, and the TAG on Statistical Design and Data Management (January, 1982) (5).

a. Criteria and Procedures. The plan for community-selection was as follows:

After obtaining informed consent of the community, there was to be a systematic collection of information, involving both examination of existing records and collection of information de novo, about communities tentatively selected for study on the basis of criteria summarized in Table 1.

The purposes of this data collection were to confirm that communities were appropriate for study, and to provide a sample frame for the selection of families to be included in the study and a baseline description of the community and its situation in relation to the proposed research variables.

The specific content and approach to data collection were to be determined in the context of individual projects, but the initial screen was to provide evidence from existing data sources or other rapid assessment that data vital to the purpose of associating intake with outcomes could be gathered. Ultimately, data were to be obtained on features relating to the dependent variables and explanatory factors (such as birth rate, family and household size and composition, acute disease incidence, occupational class distribution, the environmental conditions, etc.) These data were intended to characterize the community in such a manner that the criteria of selection could be confirmed, including a general assessment of nutritional conditions.

Existence of this information about the communities gathered in a systematic manner, would facilitate comparison across projects. It was intended also to serve as a determinant of the need for specific recurring assessments, within the research design of specific projects, for the statistical control of potential confounding variables in data analysis, which are not necessarily the same from one location to another.

Projects followed a two-step procedure in selecting the study community. Host-country investigators identified a number of potential sites based on their own knowledge and experience and information from previous surveys, and in relation to national priorities. On-site examination and review of readily obtainable background information led to selection of study districts in Kenya and Mexico and a single village in Egypt. Adequate size and infrastructure weighed heavily as selection criteria at this juncture, as did reasonable proximity to the participating host-country institution. Projects then carried out demographic/ethnographic surveys to confirm the initial judgment and to aid in later identification of study households. Detailed information relating to certain of the criteria was gathered during Phase I (along with development and pilot testing of research methods and instruments), past the

Table 1*

Criteria for Selection of Communities

Characteristics Sought	Indicators of Characteristic
Stability To minimize attrition and loss of repeated data collection capabilities and need for follow-up in different settings.	Low migration rates, minimal labor translocations. Reasonable assurance of absence of major disruptions of the socio-political milieu of the community.
Homogeneity Particularly if more than one village needs to be utilized. To reduce variability of biological variables, divergence of food patterns and family structure and interactions.	Variables not essential to the hypothesis, but likely to confound, are homogeneously distributed across the selected community, (e.g. water supplies and health facilities). Little genetic, ethnic, and cultural diversity in the selected community.
Appropriate Size	Provides either an adequate sampling frame for sample size ultimately determined or, if smaller communities need to be utilized, large enough for a whole population sample (preferably) or: Each of several communities selected must fulfill all criteria and in addition be comparable to each other in vital, ethnographic, demographic, and environmental characteristics.
Appropriate Range of Nutrient Intake	Evidence such as observable status or previous studies provides a reasonable judgment that a suitable range of nutrient intake levels exists in the population.
Adequate Infrastructure	Sufficient organization to provide continuing communication with the population, to assist in obtaining informed consent, and to assist in maintaining population cooperation to minimize attrition.
Exclusions Communities which do not satisfy all the criteria above must not be considered for the project.	Other conditions requiring exclusion are high prevalence of a complicating variable (e.g. diabetes, endemic malaria, voluntary abortion rate), a very low birth rate, atypically high mortality rate among young adults.

* from Ref. 4

time it would have been feasible to relocate the projects had that been desirable.

Homogeneity proved difficult to achieve. Multiple villages were studied in Kenya and Mexico. This was unavoidable as none was of sufficient size to provide the sample population, but these subsets had different water supplies, socioeconomic status and the like. The Egyptian sample was drawn from one village, but these factors were not noticeably more uniform. It is unlikely that the criterion could have been met better than it was, and projects did gather information needed to assess the significance of many potentially confounding variables.

The selection criterion least well evaluated, in ME's view, was the existence of an appropriate range of nutrient intake. That is not to say that the range observed in the study populations was, in the end, not sufficient to test the CRSP hypotheses (that topic is addressed in section VI.D. below), but simply that assurance was, to variable degrees, lacking at the inception of the research Phase II. No project carried out a complete dietary or nutrition survey in advance of selection of the field site or the actual sample study population. Projects generally relied on the existence of a range of nutritional states, as measured by body size, as a proxy for nutritional intake, and appear to have been convinced by the information in hand that the necessary range of food intake would be produced automatically if social, economic, and/or geographic features were taken into account in a household selection process. Once a first cohort of households was selected for study and data on food intake were obtained, there would have been an opportunity to use these data to guide selection of other households; there is no evidence that this was done.

Egypt project:

The Egyptian Nutritional Institute showed commendable attention to the ultimate policy application of CRSP research in selecting a study community. It reported its preliminary selection process[1] as follows:

> ...senior scientists of the Nutrition Institute reviewed the criteria of selection and a task force started [by] re-evaluating the National Nutrition Survey 1978 and the MIT report to identify possible areas of marginal malnutrition.
>
> The NI view was that the selection of the village from the Nile Valley proper according to the criteria may act as a model to be later simulated in any further intervention programs to be set after completion of data collection, that is, to ensure dissemination of activities.
>
> The time factor being a decisive one it was suggested that, in the interest of saving time and effort, selection should be made from the village of the governorate whose information is

[1] Excerpt from NI annual report, p. 3-5.

already available. The selected village [is] to be a model in all respects, agricultural, health and social. After deliberation and discussions, the working group set the following general basis to be taken into consideration:

> 1) To ensure the proper representation of the Egyptian village and hence the possibility of future implementation of nutrition intervention program, the population should be around 10,000 thereby representing the average.
>
> 2) It is preferable that agricultural activity should be conventional, i.e. the production of fruits should not be predominant.
>
> 3) The village selected should obtain the following services: education, medical and social, as well as an agricultural cooperative...
>
> 4) The village should be socially active and be able to contribute, with the efforts of the research teams, towards the success of the project, since execution will depend mainly on personal contact between teams and ...[villagers].
>
> 5) Communication channels (railways, roads) should be available for easy transport.
>
> 6) It is preferable that the village selected should not have been subjected to previous experiments that might have affected its original standard aspects.
>
> 7) It should not be a village of a known specific nutritional deficiency.

According to the criteria set and because of the numerousness of specifications to be taken into consideration, selection was effected in three stages after exclusion of urban, coastal and desert governorates.

The three stages were:

Stage 1: Selection on population basis
Stage 2: Selection based on availability of services
Stage 3: Selection based on field visits to update information

It was then decided to choose one of five villages which fulfilled the requirements. Most of the information was derived from ORDEV (Organization for the Reconstruction and Development of the Egyptian Villages). Kalama, in the Kalubia governorate, was then chosen for the purpose of the project.

After the village was selected, a census was conducted, followed by a sociodemographic survey. Food intake information was obtained on 60 households (33 in 1982 and 27 in 1983), subsequently classified as 16 of high socioeconomic status (SES), 28 of intermediate SES, and 16 of low SES. The geographic distribution of SES was determined and mapped.

Kenya Project:

Site selection was guided by the Kenya Ministry of Health and the Department of Community Health at the University of Nairobi. Reasons for the choice of Embu district were reported[2] as follows:

> 1. No other large scale health or nutrition study was going on in that area.
>
> 2. A desire on the part of the Ministry of Health to increase nutrition awareness in Eastern Province.
>
> 3. Convenient distance of Embu District from Nairobi. Embu is about 100 miles northeast of Nairobi on a good highway. Embu is far enough away so that a large percent of the heads of households or the males are not apt to be working in Nairobi. Logistically it is close enough so that laboratory support is possible and University of Nairobi staff can come to the field site easily.
>
> 4. There is no severe malnutrition on a large scale but rather mild [to] moderate malnutrition as disclosed by the CBS Nutrition Studies of 1979.
>
> 5. Embu District is an area of mixed agriculture and in addition some cash crops, such as coffee and tea. The area varies from a semi-arid, spersely settled area to the highlands that are more densely populated. Three sublocations were picked in Kyeni location because of the relatively dense population (although the households are somewhat scattered). This is not an exclusively semi-arid area nor is it a pure cash crop area raising only coffee and tea, but rather an area of mixed agriculture raising maize, sorghum, beans, and millet.
>
> 6. The tribal makeup of the area is fairly uniform, the people being mainly Kiembu and the language being spoken mainly Kiembu, although Swahili and Kikuyu are widely understood as is English.
>
> 7. There is a cooperative, enlightened, principal provincial medical officer with an interest in nutrition as well as a

---

[2] Excerpt abridged from Kenya Project First Annual Report, p. 4-6.

very progressive district commissioner and chief of a location who is very supportive of the work. Also the presence [in] Karurumo of rural health training center is a big asset as this training center under the provincial medical officer and Ministry of Health has offered us logistic support and general backup facilities such as the emergency use of a telephone, a good water supply and electricity furnished by a generator.

Summary of Criteria for Selection of Embu District-Kyeni Location for Field Study Area:

1. Stability - on basis of preliminary evidence, there is a low migration rate, moderate labor translocations for adult males (a characteristic of the Kenyan situation) when the areas are closer to Nairobi, [and it is a] politically and socially stable area.

2. Homogeneity - small tribal differences, mainly Kiembu - few members of other tribal groups. The area poses an environmental gradient, from semi-arid to more fertile and better watered highlands. The Study will concentrate on mid-zone which is fairly homogeneous.

3. Appropriate Size - adequate [as] there are almost 1200 households in the study area, an adequate base from which to select the 600 households for the preliminary survey. Reproductively-active females, pre-school and school-age children are well represented.

4. Appropriate Range of Nutrient and Nutritional Intake Status - on the basis of reported anthropometric measurements from a rapid survey by the UC and University of Nairobi team [in 1982] and in discussions with local health personnel, we expect to find an appropriate range of nutrient intake and nutritional status. Well-nourished and moderately malnourished groups are available. Severe PEM is not widespread.

5. Infrastructure - local governmental and health personnel and community organizations and leaders have expressed interest and support of project; logistics appear adequate [and] health facilities [are] fairly near or at the periphery of the study area.

6. Potential Problems - anticipate unstable malaria situation (malaria increases during rainy season and decreases during dry season). Malaria is present

throughout Kenya. Intestinal parasites are also ubiquitous.

Complete census and mapping were accomplished as a guide to selection of households.

Mexico Project:

INN considered alternative sites and the PIs agreed on the Valley of Solis, State of Mexico, approximately 2.5 hours driving time from Mexico City. Their deliberations were reported[3] as follows:

> Our analysis of data available in the community (health clinic and survey data), discussions with community and health care personnel, etc., provides the following reasons why this site may be suitable:
>
> 1) This site is readily accessible to Mexico City, but with minimal migration
>
> 2) The staple diet is corn and beans
>
> 3) The existence of marginal malnutrition is suggested by the high incidence of low birth-weight infants, infant mortality, low weight-for-height of school children, dependency on subsistence agriculture, etc.
>
> 4) The Valley contains a number of communities which are very similar (ethnically, culturally, geographically, economically), so that it may be feasible to use several of these for our study.
>
> 5) Spanish is spoken by everyone
>
> 6) There is an existing out-patient clinic, currently run by Salud Publica and CODAGEM (a government agricultural assistance program). This clinic has a laboratory, personnel (including a social worker) known by the community and has been the center for several anthropometric and health surveys of the valley.

Further evaluation included mapping, census of four communities tentatively chosen for study, ethnographic studies, anthropometric measurement (60% of individuals in the four communities) and dietary studies to characterize the habitual diet (24% of households). Prevalent consumption of pulque (a fermented agave juice, 2-3% alcohol) was noted. During Phase II, two additional communities within the valley were added to the study because birth rates were lower than had been anticipated. Census was carried out in the

---

[3] Excerpt from Mexico project report 12/1/81 - 2/28/82.

added communities but other evaluative measures, if conducted, are not reported. Periodic translocation of men for employment in Mexico City was substantially higher than anticipated, not surprising in a community in transition.

b. Significance to Cross-Project Evaluation. The projects' final reports indicate that their original judgments as to stability and homogeneity were not fully supported by the field experience. As examples, Egypt found morbidity to differ geographically, food intake was markedly different in successive years in Kenya, and Mexico observed a changing labor migration pattern. This presents at once an opportunity and a problem. To the extent that there is adequate documentation, projects can explore the importance of unanticipated environmental and cultural variables for which they could not control. Within project data sets, even if the number of households/target subjects within a cell (say, for example, categories of drinking water source) is too small for statistical proof, findings may have useful explanatory power and may provide very important leads for further study. The existence of a large number of differences in potentially confounding factors between sites may, however, limit the ability securely to identify commonalities in measured outcomes across projects.

2. Selection of Households and Subjects

Originally, it was the intent that the Nutrition CRSP would focus on the household as the unit of study. Study of the family or household as a reference unit would permit simultaneous measurement of many functions, and, very importantly, investigation of the interactions of various functions and conditions. The household focus would allow examination of time-related processes and of cumulative influences on an individual and the group. In the final research plan, however, the range of functions to be evaluated and the scrutiny of intrahousehold transactions were greatly reduced from early plans, based on EEP advice to simplify the work and give priority to functions of greatest interest - morbidity, cognitive development, and reproduction. This meant that rather than selecting households and studying persons/events within them, selection was for specific categories of subject with delimited study of the households to which they belonged.

Investigators had agreed early on that the mother-child dyad should be a major focus, the mother being studied during gestation and mother and child during lactation. Toddlers were selected to receive particular study because of the expectation that morbidity would be highest in this age group and because cognitive development could be assessed by change (in contrast to simply state) measures. It was difficult to decide which other children were to be studied; a child age seven to nine years was chosen finally because most would be enrolled in school at that age, boys and girls are more alike than at other ages, and they have reasonable ability to report on their own. (Investigators would have liked to include an adolescent age group but reluctantly abandoned this group because of the high variability in onset and completion of puberty and the associated biological and social changes. The adult male head of a household was included as a principal decision-maker and because of the

importance of his contribution to family welfare, notably in relation to children. Target individuals were to be members of studied households.

a. The Sampling Frame. The hypotheses to be tested posit relationships between food energy intake and a series of functions. It is evident that the study population must include an appropriate range of the independent variable, from adequate to the lower limit of moderate malnutrition. (Longitudinal study of serious malnutrition is precluded on ethical grounds.) It must also provide a range of the dependent outcomes. There was no research requirement that the population be representative of the larger local population, only that the sample be adequate in range, and number, to test the hypotheses.

For the purpose of the design, certain characteristics of a family/household determined inclusion/exclusion from the sample frame.

1) Only a small proportion of non-fertile families, not actively reproducing, could be included so as to avoid inordinate increases in sample size requirements (for the reproduction component).

2) Households containing members with severe nutritional or other diseases, apparent on clinical examination at the time of initial screening, would be excluded.

3) Households representing atypical cultural, ethnic, or genetic characteristics for the selected community would be excluded.

Based on preliminary estimates of variance, it was concluded that data would be needed on about 100 individuals of each target category. For efficiency, investigators sought households that would provide multiple targets, but the number of households was necessarily two- to three-times the number of individual target subjects in each category (See Table 2).

Selection of households was driven primarily by the need to assure 100 early pregnancies. This meant that single parents, childless households, and older adults largely were excluded from study. The number of nontarget subjects of various ages, on whom a significant amount of data were gathered (body size, weekly morbidity) is substantial and will provide an exceptional opportunity for future analyses.

Table 2

Composition of Households in CRSP Projects

	Egypt	Kenya	Mexico
No. of Households*	312	292	320
No. Persons in Household, mean	7.5**	7.28**	10***
, range	3-20	3-17	3-19

* Total households enrolled; number useful for analysis is less. See section B, below. Data are from project final reports: Egypt, chapter 3 C.& D.; Kenya, Table 5.5; Mexico, chapter 19.
** Mean of households in final sample studied.
*** Mode

The original plan called for the population to be stratified according to one or more nutrition variables and the sample selected to provide equal numbers from each stratum. The projects do not report on stratification measures and appear to have used SES as a proxy for nutrition intake/status.

b. Project Sample Population.

Egypt project:

SES was found to vary geographically and recruitment of households was done geographically which also "provided the necessary variation in socioeconomic status" (Egypt Final Report [EFR], chapter 3). Households eligible for consideration included one or more of the following individuals: adult female, aged 15-45 years, with mate present in the household, in the first trimester of pregnancy or likely to become pregnant within the first year of Phase II; child who would be 18 months of age in the first year of study; and/or child who would be seven to nine years of age during the first year of study. Households were excluded if children's ages were not known accurately or if the household stability through the projected study period was threatened due to death or divorce. Selection focussed on households with a toddler or schooler and assumed a 3-year span between births.

The first enrollment did not provide sufficient pregnancies. The sample was increased several months later by enrollment of newlyweds and households in which the youngest child was older than 12 months.

Kenya Project:

Preliminary registration of 823 households from three villages was according to criteria designed to provide the required number of target subjects. To be eligible, the household had to consist of a monogamous pair living together in which the woman was below 40 years of age. A birth interval of 25 months was estimated from preliminary data and used as one indicator of probable pregnancy; age of living children was another. (Kenya Final Report [KFR], chapter 5) Additional information was gathered, including information on food intake, and the sample was reduced to 292 households based on criteria relating to pregnancy. There is no reference to other selection bases (eg. SES, nutrition intake) if they were applied. Households were grouped according to geographic location (clusters) for convenient accessibility by field staff. Attrition rate was 12%, but no additional recruitment was required to generate the agreed on numbers of target subjects.

Mexico Project:

Initial criteria for selection of households from four communities were presence of lead male and female and one target child, predominantly schoolers. (Mexico Final Report [MFR], chapter 4.IV) A second enrollment added households having a woman less than five months pregnant or a toddler approaching 18 months of age. Additional recruitment was based on pregnancy-testing at three-month intervals and toddlers reaching entry age. No other selection criteria are specified in the MFR.

As noted above, additional villages were added to the sample, from the Solis Valley. The fifth "community" was an outlying neighborhood of the fourth community; the sixth community was the next settlement over from one of the first four. Selection criteria for households, other than pregnancy, are not stated.

c. Implications for Cross-Project Analysis. Because of the different enrollment schemes, cohorts of target subjects do not necessarily overlap in time; parents were, however, always studied during the one-year period of study of a child. Because some households had more than one special target (second trimester pregnancy plus infant to six months, toddler, schooler), some were studied well beyond one year and others were not. In all cases, the pregnancy sample is the latest in time in as much as enrollment was anticipatory or awaited proof. Generally, the schooler cohort is the earliest in time. These differences present a problem where environmental conditions changed from one year to the next. This is particularly a difficulty in the Kenya sample because of the occurrence of drought with associated food scarcity, more evident in the first period of Phase II than in the second. SES conditions also may have changed over time and projects generally attempted to assess SES factors at entry and exit of a study household.

As food intake was not a selection criterion, the existence of an adequate range for hypothesis testing is not assured. (discussed in Section VI.D.,

below). This may equally be true of anthropometric measures, the likeliest proxy measure for food intake. (See Section VI.F., below) The ends of the intake scale, operationally defining "adequate" and "moderately inadequate" intake, may well differ among projects. The CRSP may shed light on how "adequate" and "moderately inadequate" should be defined in functional terms and perhaps conclude that few, if any, subjects meet the definitions in one or more of the sites.

B. Size of the Cross-Sectional and Longitudinal Samples

1. The Cross-Sectional Sample

All tables in this section are based on the original data as archived by the three country projects. Table 3 shows the total number of subjects on whom selected measures (food intake, anthropometry, psychology, or morbidity) were collected. No criteria are used to select these individuals; the numbers are a simple count of individuals identified by the projects as belonging to the indicated target group, and having one or more of the specified type of record. (Thus, it is possible that these numbers may overstate the actual number of individuals if coding errors exist for the subject identifiers.) The counts for each of the four measures may not always include the same individuals (if, for example, some individuals were measured for anthropometry but not for psychology). Thus, these numbers indicate the maximum number of individuals that may be used for cross-sectional analyses.

2. The Longitudinal Sample

The number of longitudinal targets were counted for each project using the following criteria:

Longitudinal toddler: Having at least one food intake record prior to age 19.0 months and at least one food intake record after age 29.0 months.

Longitudinal schooler: Having a food intake record between the ages of 7.0 years and 8 years 1 month, and another food intake record at least 10 months later.

Longitudinal pregnancy/lactation: A lead female with at least one food intake record prior to her infant's birthdate minus 5.0 months, and at least one food intake record after her infant's birthdate plus 5.0 months.

Thus, by this definition, a longitudinal target must have food intake records spanning at least 10 months, although no minimum number of records is required. The number of each of these three types of longitudinal subjects is shown in Table 4.

A longitudinal household is one having one or more of the three targets defined above. The entry date for a longitudinal household is the date of the earliest of the following three events:

1. Longitudinal toddler becomes 18.0 months of age.
2. Longitudinal pregnancy begins (infant birthdate minus 6.0 months).
3. Longitudinal schooler begins (first food intake record after 7.0 years of age).

TABLE 3
NUMBER OF TARGETS IN THE CROSS-SECTIONAL SAMPLE

	Egypt	Kenya	Mexico
Food intake:			
Lead males	289	285	271
Lead females	290	286	284
Schoolers	123	222	195
Toddlers	157	133	149
Infants	130	149	98
Anthropometry:*			
Lead males	220	289	211
Lead females	236	291	268
Schoolers	155	183	182
Toddlers	202	130	154
Infants	81	143	134
Psychology:*			
Lead males	64	244	200
Lead females	115	248	252
Schoolers	115	144	181
Toddlers	57	114	133
Infants	75	121	108
Morbidity:*			
Lead males	126	236	320
Lead females	182	271	329
Schoolers	74	151	156
Toddlers	139	144	137
Infants	110	139	115

* Anthropometry = body weight.
Psychology = Wechsler for lead males, lead females, and schoolers, Bayley for toddlers and Brazelton for infants.
For Kenya and Egypt, morbidity does not count individuals never ill. It is reasonable to assume that the majority of individuals with food intake records but no morbidity records were never ill.

The exit date for a longitudinal household is the date of the latest of the following three events:

1. Longitudinal toddler becomes 30.0 years of age.
2. Longitudinal infant becomes 6.0 months of age (a longitudinal infant is the child of a target pregnancy as defined above).
3. Longitudinal schooler exits (12.0 months after the date of the first food intake record that was used to identify the longitudinal schooler (see above).

The longitudinal window for a household is thus the time between the entry and exit dates. A longitudinal lead male and a longitudinal lead female must be members of a longitudinal household.

Tables 5-8 show the number and entry dates of the longitudinal households as well as the three target types.

TABLE 4
NUMBER OF LONGITUDINAL SUBJECTS

	Number of subjects	Mean number of food intake records/subject*	CV
Kenya			
toddlers	91	23.4	6.3
schoolers	129	23.6	6.7
pg/lact	93	22.7	9.3
Egypt			
toddlers	88	19.1	19.7
schoolers	42	19.7	20.2
pg/lact	62	18.5	17.2
Mexico			
toddlers	55	23.0	11.5
schoolers	109	23.4	9.4
pg/lact	69	22.7	13.2

* Within the longitudinal window

TABLE 5
NUMBER AND ENTRY DATES OF LONGITUDINAL HOUSEHOLDS

		Egypt	Kenya*	Mexico**
1983	November	3		
	December	19		
1984	January	12	35	30
	February	14	50	9
	March	17	40	4
	April	13	25	10
	May	14	15	8
	June	9	13	11
	July	15	13	27
	August	9	7	17
	September	3	7	7
	October	10	2	8
	November	10	6	4
	December	3	1	4
1985	January	3		6
	February			4
	March			4
	April			1
	May			4
TOTAL HOUSEHOLDS		154	214	158

* Drought in Kenya was approximately from June through November, 1984
** Mexico Project food intake protocol changes occurred in October, 1984 and in January, 1985.

TABLE 6
NUMBER AND ENTRY DATES OF LONGITUDINAL SCHOOLERS

		Egypt	Kenya	Mexico
1983	November			
	December	6		
1984	January		24	30
	February	1	36	6
	March	8	12	4
	April	4	9	4
	May	3	5	5
	June		4	5
	July	5	11	27
	August	1	6	14
	September	1	6	4
	October	3	3	4
	November	4	7	5
	December	3	5	
1985	January	3	1	
	February			
	March			
	April			1
	May			
TOTAL SCHOOLERS		42	129	109

TABLE 7
NUMBER AND ENTRY DATES OF LONGITUDINAL TODDLERS

		Egypt	Kenya	Mexico
1983	November	3		
	December	16		
1984	January	14	5	
	February	10	9	
	March	11	25	
	April	11	19	6
	May	7	13	1
	June	1	9	4
	July	3	3	3
	August	1	1	2
	September		3	2
	October	4	3	6
	November	7	1	4
	December			3
1985	January			8
	February			7
	March			7
	April			1
	May			1
TOTAL TODDLERS		88	91	55

TABLE 8
NUMBER AND ENTRY DATES OF LONGITUDINAL PREGNANCIES

		Egypt	Kenya	Mexico
1983	November			
	December			
1984	January		7	
	February	4	11	4
	March	8	14	
	April	4	12	5
	May	5	11	3
	June	12	6	5
	July	9	12	3
	August	10	7	7
	September	4	5	5
	October	4	3	4
	November	2	3	3
	December		2	8
1985	January			4
	February			4
	March			2
	April			6
	May			6
TOTAL PREGNANCIES		62	93	69

The number of observations for longitudinal targets are counted using the standard protocol windows for infants (0 to 6.0 months of age) and toddlers (18.0 to 30.0 months of age). The window for schoolers is considered to be the 12.0 months following the first food intake between 7.0 and 8 years 1 month of age. The window for lead males and lead females is the same as the window for the household. The psychology measures for adults were considered state variables, and did not have to be measured within the longitudinal window. For schoolers, psychology measures were not required to be within the exact 12 month longitudinal window, but were accepted any time between 7 and 9 years of age.

Table 9 is in the same format as Table 3, but contains the counts of longitudinal targets rather than cross-sectional targets. Note that pregnant/lactating women are a subset of lead females. The counts in Table 9 indicate how many longitudinal targets have at least one of the specified measure. Tables 10-14 show frequencies of food intake measures for each of the target types. The protocol called for 24 food intake measures (2 per month for one year) for target toddlers, schoolers, and pregnant/lactating women. Lead males and lead females were to have at least 24 measures, and often had more if there were multiple non-overlapping targets in the household.

TABLE 9

NUMBER OF TARGETS IN THE LONGITUDINAL SAMPLE

	Egypt	Kenya	Mexico
Food intake:			
Lead males	153	214	155
Lead females	153	214	158
Pg/lact females	62	93	69
Schoolers	42	129	109
Toddlers	88	91	55
Anthropometry:			
Lead males	108	212	136
Lead females	150	214	158
Pg/lact females	61	93	69
Schoolers	40	183	109
Toddlers	86	129	55
Infants	51	93	69
Psychology:			
Lead males	42	208	155
Lead females	78	214	158
Pg/lact females	8	80	15
Schoolers	40	129	108
Toddlers	27	91	54
Infants	30	83	31

TABLE 10

FREQUENCY OF FOOD INTAKE MEASURES ON LEAD MALES
IN LONGITUDINAL HOUSEHOLDS

Number of measures (Protocol=24+)	Number of subjects Egypt (%)	Kenya (%)	Mexico (%)
1-2	1 (0.7)		3 (1.9)
3-4	1 (0.7)	1 (0.5)	6 (3.9)
5-6	3 (2.0)	2 (0.9)	4 (2.6)
7-8	3 (2.0)	1 (0.5)	5 (3.2)
9-10	5 (3.3)		5 (3.2)
11-12	8 (5.2)	3 (1.4)	5 (3.2)
13-14	12 (7.8)	2 (0.9)	9 (5.8)
15-16	13 (8.5)	6 (2.8)	6 (3.9)
17-18	19 (12.4)	7 (3.3)	10 (6.5)
19-20	33 (21.6)	13 (6.1)	13 (8.4)
21-22	19 (12.4)	27 (12.6)	17 (11.0)
23-24	19 (12.4)	69 (32.2)	33 (21.3)
>24	17 (11.1)	83 (38.8)	39 (25.2)
TOTAL	153 (100)	214 (100)	155 (100)

TABLE 11

FREQUENCY OF FOOD INTAKE MEASURES ON LEAD FEMALES
IN LONGITUDINAL HOUSEHOLDS

Number of measures (Protocol=24+)	Number of subjects Egypt (%)	Kenya (%)	Mexico (%)
1-2			
3-4			
5-6	2 (1.3)		
7-8	1 (0.7)		
9-10	1 (0.7)		1 (0.6)
11-12	3 (2.0)		1 (0.6)
13-14	11 (7.2)		
15-16	14 (9.2)	1 (0.5)	
17-18	20 (13.1)	2 (0.9)	
19-20	33 (21.6)	7 (3.3)	4 (2.5)
21-22	29 (19.0)	34 (15.9)	20 (12.7)
23-24	21 (13.7)	75 (35.0)	56 (35.4)
>24	18 (11.8)	95 (44.4)	76 (48.1)
TOTAL	153 (100)	214 (100)	158 (100)

TABLE 12

FREQUENCY OF FOOD INTAKE MEASURES ON LONGITUDINAL SCHOOLERS

Number of measures (Protocol=24)	Number of subjects Egypt (%)	Kenya (%)	Mexico (%)
1-2			
3-4	1 (2.4)		
5-6			
7-8			
9-10			
11-12			1 (0.9)
13-14	2 (4.8)		
15-16	4 (9.5)		1 (0.9)
17-18	5 (11.9)	1 (0.8)	
19-20	14 (33.3)	[illegible] (3.1)	6 (5.5)
21-22	6 (14.3)	[illegible] (2.5)	19 (17.4)
23-24	6 (14.3)	[illegible] (5[illegible].3)	60 (55.0)
>24	4 (9.5)	25 (19.4)	22 (20.2)
TOTAL	42 (100)	129 (100)	109 (100)

TABLE 13

FREQUENCY OF FOOD INTAKE MEASURES ON LONGITUDINAL TODDLERS

Number of measures (Protocol=24)	Number of subjects Egypt (%)	Kenya (%)	Mexico (%)
1-2			
3-4			
5-6	1 (1.1)		
7-8			
9-10			
11-12	2 (2.3)		1 (1.8)
13-14	8 (9.1)		
15-16	9 (10.2)	1 (1.1)	
17-18	13 (14.8)		3 (5.5)
19-20	21 (23.9)	2 (2.2)	1 (1.8)
21-22	19 (21.6)	22 (24.2)	18 (32.7)
23-24	10 (11.4)	56 (61.5)	21 (38.2)
>24	5 (5.7)	10 (11.0)	11 (20.0)
TOTAL	88 (100)	91 (100)	55 (100)

TABLE 14

FREQUENCY OF FOOD INTAKE MEASURES ON LONGITUDINAL PREG/LACT WOMEN

Number of measures (Protocol=24)	Number of subjects Egypt (%)	Kenya (%)	Mexico (%)
1-2			
3-4			
5-6			
7-8			
9-10	1 (1.6)		1 (1.4)
11-12	1 (1.6)		1 (1.4)
13-14	7 (11.3)		1 (1.4)
15-16	8 (12.9)	2 (2.2)	
17-18	11 (17.7)	3 (3.2)	2 (2.9)
19-20	15 (24.2)	9 (9.7)	4 (5.8)
21-22	14 (22.6)	24 (25.8)	14 (20.3)
23-24	4 (6.5)	44 (47.3)	37 (53.6)
>24	1 (1.6)	11 (11.8)	9 (13.0)
TOTAL	62 (100)	93 (100)	69 (100)

Tables 15-20 show the frequencies of anthropometry measures (using body weight as the variable) for six target types.

TABLE 15
FREQUENCY OF ANTHROPOMETRY MEASURES ON LEAD MALES
IN LONGITUDINAL HOUSEHOLDS
(Protocol = 12+ measures)

Number of measures	Number of subjects		
	Egypt (%)	Kenya (%)	Mexico (%)
0	45 (29.4)	2 (0.9)	19 (12.3)
1	42 (27.5)	1 (0.5)	30 (19.4)
2	32 (20.9)	3 (1.4)	34 (21.9)
3	17 (11.1)	2 (0.9)	34 (21.9)
4	12 (7.8)	3 (1.4)	24 (15.5)
5	4 (2.6)	7 (3.3)	7 (4.5)
6		7 (3.3)	7 (4.5)
7		11 (5.1)	
8		17 (7.9)	
9		24 (11.2)	
10		28 (13.1)	
11		25 (11.7)	
12		46 (21.5)	
>12	1 (0.7)	38 (17.8)	
TOTAL	153 (100)	214 (100)	155 (100)

TABLE 16
FREQUENCY OF ANTHROPOMETRY MEASURES ON LEAD FEMALES
IN LONGITUDINAL HOUSEHOLDS
(Protocol = 12+ measures)

Number of measures	Number of subjects		
	Egypt (%)	Kenya (%)	Mexico (%)
0	3 (2.0)		
1	4 (2.6)		1 (0.6)
2	5 (3.3)		4 (2.5)
3	10 (6.5)		19 (12.0)
4	17 (11.1)		39 (24.7)
5	13 (8.5)		12 (7.6)
6	16 (10.5)		8 (5.1)
7	26 (17.0)	2 (0.9)	6 (3.8)
8	20 (13.1)		6 (3.8)
9	15 (9.8)	4 (1.9)	5 (3.2)
10	8 (5.2)	6 (2.8)	8 (5.1)
11	3 (2.0)	23 (10.7)	12 (7.6)
12	4 (2.6)	28 (13.1)	16 (10.1)
>12	9 (5.9)	151 (70.6)	22 (13.9)
TOTAL	153 (100)	214 (100)	158 (100)

TABLE 17
FREQUENCY OF ANTHROPOMETRY MEASURES ON LONGITUDINAL SCHOOLERS
(Protocol = 5 measures)

Number of measures	Number of subjects		
	Egypt (%)	Kenya (%)	Mexico (%)
0	2 (4.8)		
1	6 (14.3)		
2	2 (4.8)		1 (0.9)
3	5 (11.9)		4 (3.7)
4	10 (23.8)		30 (27.5)
5	8 (19.0)		26 (23.9)
6	3 (7.1)		37 (33.9)
7	2 (4.8)	4 (3.1)	9 (8.3)
8	1 (2.4)	1 (0.8)	2 (1.8)
9	3 (7.1)	15 (11.6)	
10		17 (13.2)	
11		38 (29.5)	
12		48 (37.2)	
>12		6 (4.7)	
TOTAL	42 (100)	129 (100)	109 (100)

TABLE 18
FREQUENCY OF ANTHROPOMETRY MEASURES ON LONGITUDINAL TODDLERS
(Protocol = 12 measures)

Number of measures	Number of subjects		
	Egypt (%)	Kenya (%)	Mexico (%)
0	2 (2.3)		
1	6 (6.8)		
2	1 (1.1)		
3	2 (2.3)		
4	12 (13.6)		
5	11 (12.5)		1 (1.8)
6	12 (13.6)		
7	13 (14.8)	1 (1.1)	2 (3.6)
8	8 (9.1)		2 (3.6)
9	10 (11.4)	4 (4.4)	1 (1.8)
10	5 (5.7)	5 (5.5)	7 (12.7)
11	3 (3.4)	18 (19.8)	4 (7.3)
12	2 (2.3)	34 (37.4)	3 (5.5)
>12	1 (1.1)	29 (31.9)	35 (63.6)
TOTAL	88 (100)	91 (100)	55 (100)

TABLE 19
FREQUENCY OF ANTHROPOMETRY MEASURES ON LONGITUDINAL PREG/LACT FEMALES
(Protocol = 12 measures)

Number of measures	Number of subjects Egypt (%)	Kenya (%)	Mexico (%)
0	1 (1.6)		
1	2 (3.2)		
2	5 (8.1)		
3	7 (11.3)		2 (2.9)
4	6 (9.7)		1 (1.4)
5	9 (14.5)		2 (2.9)
6	7 (11.3)		5 (7.2)
7	6 (9.7)		6 (8.7)
8	5 (8.1)	1 (1.1)	3 (4.3)
9	8 (12.9)	1 (1.1)	9 (13.0)
10	2 (3.2)	3 (3.2)	14 (20.3)
11	1 (1.6)	1 (1.1)	20 (29.0)
12	1 (1.6)	13 (14.0)	6 (8.7)
>12	2 (3.2)	74 (79.6)	1 (1.4)
TOTAL	62 (100)	93 (100)	69 (100)

TABLE 20
FREQUENCY OF ANTHROPOMETRY MEASURES ON LONGITUDINAL INFANTS
(Protocol = 7 measures)

Number of measures	Number of subjects Egypt (%)*	Kenya (%)	Mexico (%)
0	9 (17.1)		
1	5 (9.4)		
2	4 (7.5)		2 (2.9)
3	3 (5.7)		3 (4.3)
4	1 (1.9)	3 (3.2)	1 (1.4)
5	2 (3.3)	11 (11.8)	4 (5.8)
6	6 (11.3)	35 (37.6)	19 (27.5)
7	14 (26.4)	30 (32.3)	22 (31.9)
>7	18 (34.0)	14 (15.1)	18 (26.1)
TOTAL	62 (100)	93 (100)	69 (100)

* Due to discrepancies in the Egypt Project infant IDs on various files, it was not possible to match the anthropometry IDs with the birthdate IDs. Thus, the information in this table is obtained from household ID only, and may understate the number of individual infants, due, for example, to households with twin target infants.

C. Food Composition Data Bases

1. Development of the Project Nutrient Data Bases

Each of the three projects developed a food composition data base for use in converting the collected food intake data into nutrient intake data. As shown in Table 21, the numbers of food items and nutrients in the tables varied somewhat among the projects. The original CRSP protocol called for a minimum of four nutrient variables: energy (kcal), protein (grams), fat (grams), and carbohydrate (grams). All projects included these four variables in their nutrient data bases. In addition, all projects calculated iron intakes (mg) and animal protein intakes (grams). Mexico and Egypt Projects calculated intakes of several additional nutrients, as shown in Table 21.

Two projects (Kenya and Egypt) developed new nutrient data bases, using a combination of published data and data from analyses of a small number of frequently consumed local foods, performed by Medallion Laboratories.[1] Mexico Project expanded an existing nutrient data base, developed by the INN. These data bases, in common with all tables of food composition, reflect, at best, the average composition of foods actually consumed; values often are derived from a small number of samples analyzed by methods of differing quality.

It should be noted that much of the published food composition data for the three countries is quite old. Although some limited analyses were funded as a part of the Nutrition CRSP, there is still much work to be done. Not only has the food supply changed in recent years, but analytic methodologies have undergone significant advances. We recommend that countries consider updating their composition data as a top priority.

For all projects, certain problems existed with specific nutrient data, as discussed below:

a. Energy. The accepted method of calculating energy values for food composition tables is to multiply the protein, fat, and carbohydrate values by specific energy values for each. These approximate 4, 9, and 4 kcal/gram respectively, but vary somewhat by the type of food due to differences in the array of organic compounds present and in their digestibility. Complete values have been published (6). However, the Medallion analyses reported energy values using the approximate values (4,9,4), rather than the specific values by food group. Kenya Project data are based on a recalculation of the energy using specific energy values, rather than on the values reported by Medallion. The situation was more complex for the Egypt Project, since Medallion had also analyzed dietary fiber values. For this project, Medallion reported energy values based on the 4,9,4 factors, but first subtracting dietary fiber from the carbohydrate values. Thus, in order to make the energy values comparable across projects, as well as within the Egypt Project nutrient data base, it was necessary to recalculate energy for these food items, using the specific factors and total carbohydrate.

---

[1] Medallion Laboratories, 9000 Plymouth Ave., Minneapolis, MN 55427.

TABLE 21

NUTRITION CRSP

SUMMARY OF FOOD COMPOSITION DATABASES FOR THE THREE PROJECTS

	EGYPT	KENYA	MEXICO
DATABASE SIZE:	955 Database records 206 Food codes	180 Ingredient codes including 172 "ingredient" foods and 8 "hotel" foods. 154 Standard recipes	509 records in database
NUTRIENTS IN DATABASE:	KCAL PRO (animal), PRO (plant), FAT (animal), FAT (plant), CHO Ca, Fe (animal), Fe (plant), Na Thiamine, Niacin, vitamin A, B-carotene, vitamin C	KCAL, PRO, Animal PRO, FAT, CHO, Fe	KCAL, PRO, CHO, FAT Ca, Fe niacin, riboflavin, thiamine, vitamin C
FOOD CODING STRATEGY:	4 digit food code appended by a 2 digit preparation code	3 digit ingredient codes and a 4 digit dish code (for recipe code). The low order digit of the dishcode is incremented for subsequent preparations of a dish within the 2 day survey period.	4 character alphanumeric food code composed of a 1 or 2 letter food class code appended by a 2 digit code sp[illegible] for a food in that class.
EDIBLE PORTION / AS PURCHASED:	AS PURCHASED	EDIBLE PORTION	AS PURCHASED

Consumption of alcoholic beverages was reported by the Mexico and Kenya projects. Yielding 7 kcal/gram, alcohol can contribute importantly to energy intake. A large portion of consumption by the Mexican population studied is as pulque, a locally produced beverage from fermented agave juice; some locally brewed beer appears in Kenya records. The composition of these local brews is likely to be more variable than commercial products but the variance, if known, is unreported.

b. Carbohydrate. Carbohydrate was calculated "by difference" for all projects - that is, the remaining weight of a food item once the water, fat, protein, and ash had been subtracted. This is the method generally used by publishers of food composition tables in the United States, and also by the FAO food composition tables. The major exception is the British food composition table (7). Paul and Southgate report analyzed carbohydrate values, expressed as simple sugar equivalents; the tables do not include ash values so it is not possible to calculate carbohydrate by difference. The British values were used for a small number of items on the Kenya nutrient data base. As far as can be determined, all other carbohydrate values reflect carbohydrate by difference.

Validity of the energy values assigned to "carbohydrate by difference" has been questioned on the basis that values were determined from studies in which the intake of course fiber-rich foods was low. If digestibility is poorer when intake of these foods is high, energy intake is overestimated. The Mexico project particularly questions the INN value of 4.13 kcal/g carbohydrate in maize; the accepted specific energy value is 4.03 kcal/g which may also be an overestimate. In the Kenya data base, the caloric yield of legume carbohydrate was computed as 3.78 kcal/g based on a coefficient of digestibility of 90%. The question of digestibility is addressed below.

c. Iron. The iron values used in all nutrient data bases reflect the total iron content of a food item. No attempt was made to adjust for bioavailability (although estimates of deficiencies did assume a mean bioavailability). Furthermore, no attempt was made by any project to incorporate the iron content of water into the nutrient data base. Thus, depending on the iron content of the local water supply (which analyses showed was not insignificant for Kenya villages), the actual iron intakes may be underestimated. Finally, the level of iron contamination from sources such as dirt on food items cannot be estimated. There is substantial variation in published iron values for many food items, and one source of this variation may be the amount of contamination iron. Thus, it is the opinion of ME that iron intake values are at best an estimate of actual iron intakes for the CRSP populations.

d. Vitamins. Both Egypt and Mexico Projects have included several vitamins on their nutrient data bases. The loss of vitamin activity with preparation, storage and processing is always difficult to estimate. USDA uses standard retention percents for calculation of the vitamin content of prepared dishes (8). Egypt Project incorporated these values into their nutrient data base, but it is not known if the factors are completely applicable to the standard preparation methods used in Kalama. However, the format of the Egypt nutrient data base is ideal for the use of vitamin retention factors - for every food

item, multiple entries are carried reflecting all common preparation methods (e.g., baked, fried, stewed, etc.). It is not known if retention factors were used by Mexico Project, but the lack of multiple preparation methods on the nutrient data base excludes the possibility of using specific retention factors for foods with multiple preparation methods. It is the opinion of ME that vitamin intake values should be viewed as very crude estimates.

e. Animal sources of nutrients. All three projects have reported animal protein intakes. Egypt Project has also separated animal and plant sources of fat and iron. Mexico Project has reported animal and plant energy. Egypt and Kenya Projects have added animal sources of nutrients to their respective nutrient data bases, and have partitioned the nutrients in mixed dishes when appropriate. For example, if an item on the data base is composed of a mixture of plant and animal ingredients, such as an infant cereal fortified with milk powder, the relative proportions of animal and plant protein are indicated on the data base. The approach taken by the Mexico Project was quite different - a SAS program was written to classify food items as either plant or animal, depending on the major ingredients in the item. Thus, all nutrients in a food item are either plant or animal, never a combination of the two. Although this approach introduces some error in the determination of animal and plant nutrients in a diet, it is not believed that the magnitude of the error is large in the present instance, because there are relatively few mixed food items on the Mexico nutrient data base, and most of these are reported infrequently.

2. Validation of the Nutrient Data Bases

Several methods were used by the projects and management to validate the values on the nutrient data bases. Validation efforts addressed only the proximate nutrients (energy, protein, fat, carbohydrate, and to a lesser extent, animal protein). Thus, all other nutrients must be considered unvalidated.

a. Comparison to analytic values of composites. (Kenya and Egypt)

Management gave all projects the opportunity to collect local food items for nutrient analysis by contract with Medallion Laboratories. Projects were also urged to collect samples of composite meals, so analyzed values could be compared to those calculated from the nutrient data base. Only Egypt Project took full advantage of this offer, and collected composite meals for seven types of individuals (lead male, pregnant female, etc.). The results of these analysis were very useful in identifying several potential errors on the Egypt nutrient data base, although errors in the sample collection, particularly for diets high in fat, are also suspected. Kenya Project collected samples of several common mixed dishes, and these analytic values were compared to average calculated nutrient values for a dish. However, the exact ingredient content of each dish was not recorded (since in several cases, the samples were composites of multiple dishes of the same type), so comparison to values for a specific individual dish were not possible.

b. Comparison to hand-calculated values (all projects, by ME).

Management hand-calculated nutrient values for several diets from all projects. Numerous errors were identified in this process, and communicated to the projects for correction.

c. Calculation of 4,9,4 energy values and comparison to energy values on the nutrient data bases (by ME).

All three nutrient data bases were examined by calculating the energy value of each food item from the protein, fat, and carbohydrate content, using the general energy factors of 4 kcal/g for protein and carbohydrate, and 9 kcal/g for fat. This calculated energy content should closely match the energy value on the data base for most food items. Large differences were highlighted, and sent to the projects for resolution. Several large errors were corrected as a result of this process.

3. Problems with "As Purchased" versus "As Consumed" Nutrient Values

There are innumerable problems with ensuring that the food items on the nutrient data base reflect the same form of the food item for which the portion sizes were collected and which the subject actually consumed.

A major issue concerns the weight of refuse. Nutrient values in food composition tables always reflect the analyzed content of only the edible portion of a food item (for example, the nutrient content of bones is not included in the nutrient content of chicken), but the weight of the food item may include the refuse. Thus, the nutrient content of 100 grams of chicken will be different (lower) if the 100 grams includes the bones, than if the 100 grams is the meat and skin only. Refuse weights are estimated for food items (and may be zero), and nutrient values may be reported both for weights including refuse (as purchased) and for weights not including refuse (edible portion). Clearly, whatever form is chosen for the nutrient data base must be matched by the measurement of the subject's portion sizes. If reported weights include refuse, then the nutrient data must reflect a nutrient density that corresponds to the same amount of refuse.

A related issue concerns the portion of the food item that is considered edible. For example, chicken feet may be considered edible in some cultures and not in others. Thus, the nutrient data for a food item may vary depending on the individuals being studied.

A third issue concerns the preparation of the food item. If the weight measurements reflect the loss or absorption of water, it is critical that these changes be reflected on the nutrient data base. For example, the nutrient density (kcal/g weight) of cooked rice is clearly less than that of raw rice. Furthermore, the nutrient density of a liquid that has cooked uncovered for some time is greater than that of the original item.

All these issues must be considered in matching the collection of the weights of the food items consumed with the expression of the nutrient values on the nutrient data base. It clearly is critical that these weights match exactly

the form of the food on the nutrient data base. Projects addressed these concerns in different ways.

Kenya Project collected only edible portion data - average refuse factors were calculated and applied to the weights of food items in the field. Thus, food item weights that were entered onto the data files were for edible portions only, a fact which was reflected on the nutrient data base. Care was taken to specify the exact form of the food item (raw versus cooked), so the item could be matched to the equivalent form on the nutrient data base.

Mexico Project collected "as purchased" data on food items, and the nutrient data base contains nutrient values for weights of food items that include refuse. In general, raw forms of food items were used, both in collecting the food data and on the nutrient data base. Conversion factors were derived for some commonly used foods that gain water in home preparation (rice, pasta, and beans), based on analysis of the moisture content of food ready for consumption and the weight of starting product. The weight "as prepared" was needed particularly where, for example, leftover cooked beans was the ingredient in a second dish, refried beans. These conversion factors are the average of figures for 32 households in the study community without consideration of within or between household variation.

Egypt Project also collected "as purchased" data on food items. Even in cases where the weight of a food item was actually measured in the form "as consumed", the weight of the "as purchased" (i.e., uncooked and containing refuse) form was calculated by the interviewer before the data were entered onto the data files.

4. Determination of the Nutrient Content of Recipes (Mixed Dishes)

Since many (in some cases, most) of the food intake of the CRSP households was from mixed dishes, it was necessary to devise a method of recording the ingredients in these recipes, and the subsequent allocation of the dish to household targets. Once again, the procedures differed among the projects.

Kenya Project personnel weighed each ingredient (or estimated the weight if the interviewer was not present) prior to its inclusion in a mixed dish, including all water added, and then weighed the final dish prior to its consumption. All weights were edible portions (refuse was subtracted). From this information, the nutrient content per 100 grams of the dish could be calculated. The amount consumed by each individual was also weighed (or estimated, as necessary).

Mexico Project estimated (and whenever possible, weighed) the quantity of each ingredient prior to its inclusion in a recipe. The weight of the final cooked dish was not measured. The weight of portions consumed by individuals was estimated, and included refuse that was served to individuals (such as bones). There is some question as to when other refuse was included. Refuse such as peels were included in the weights prior to estimation of the mixed dish ingredients. However, the portions measured for individuals do not include this type of refuse. Thus, two problems may be present for the Mexico nutrient data: one, that the nutrient density of recipes as consumed actually

may be different than that calculated for the recipe as prepared; two, that the nutrient density of the mixed dishes may be somewhat greater than calculated from the ingredients, due to loss of water during cooking.

For Egypt Project, ingredients in mixed dishes were allocated to each individual by interviewers in the field. The lead female estimated portion sizes as fractions of the amount prepared (e.g., the lead male consumed one-fourth, etc.), using her own portion as a reference in some cases. Thus, the interviewer determined the weight of each ingredient in a recipe, and then calculated the percent of the ingredient that was consumed by each individual. This calculated portion size for each ingredient was then entered on the data file. Furthermore, all ingredients were recorded in the "pre-preparation" form, even if this required conversion from the "as consumed" form. Thus, the precision of the nutrient intake data is likely to be substantially less than for the other two projects, due to the implied inaccuracy of the lead female's estimates of portion sizes as compared to a portion size estimated from the utensil from which the food item was consumed (glass, bowl, etc.).

5. Indicators for Missing Data

For all projects, missing nutrient values for food items were avoided whenever possible. In some cases, values had to be estimated from the nutrient content of similar food items. Both Kenya and Mexico Projects filled in their nutrient data bases completely. Egypt Project has recently archived a nutrient data base with no missing values for energy, but a number of missing values for vitamins and minerals; but the data base used in calculating food intake values for the final report was incomplete for several food items. The presence of missing values is not flagged on the Egypt Project nutrient intake data files.

Furthermore, for Egypt Project, there were several food items for which nutrient values were later judged to be incorrect. Many of the inaccuracies were due to incorrect energy values, so it was decided that energy intakes for each individual would be calculated (using the 4,9,4 factors) from the protein, fat, and carbohydrate of the person's diet.

Mexico Project developed programs to check extensively for mismatched food items. This was the only project to key-enter the names of every food item in a diet, as well as the food item numeric code. Thus, it was possible to determine if the name in the diet matched the name on the nutrient data base. All mismatches for either code or name were reviewed, and corrected either at the University of Connecticut, or sent to Mexico for correction.

Kenya Project was the only project to carry a flag whenever a nutrient value was believed incomplete. This situation occurred whenever a recipe was reported by an individual, but the corresponding ingredient information could not be located.

All three projects developed flags to indicate a toddler was being breast-fed (and therefore the nutrient data were likely to be incomplete).

Egypt Project developed a flag to indicate the number of meals consumed at home by the lead male. If the number was less than three, it could be assumed there was an outside source of nutrient intakes, and that the calculated intakes were incomplete. This procedure was necessary due to the relatively large number of meals consumed outside the home by Egypt Project lead males.

6. Digestibility and Bioavailability

Values for the energy yield of foods, including those referred to as "Specific Energy Values", are based on data generated from human studies in which the foods were fed as components of a typical Western diet relatively low in total fiber (6). Projects have a uniform concern that the diets typical of their study populations are much higher in fiber and that accepted energy values overestimate the energy yield of their diets. The reasoning is that the presence of large amounts of poorly digested carbohydrates ("fiber") induces physiologic responses that result in excess fecal loss of energy (even though the accepted values may be correct when the same foods are eaten in the smaller amounts customary in Western diets).

Fiber in foods consists primarily of a mixture of carbohydrates which are not digested by human digestive enzymes, but which are susceptible to digestion by enzymes of the enteric microflora. (The fiber lignin, present in generally small amounts, is not a carbohydrate and is resistant to digestion by human and bacterial enzymes.) Various analytical methods detect some or all of the fiber types, and estimation of amount in foods varies accordingly. At the present, published values for fiber are open to question.

Even if fiber content of CRSP foods were fully categorized and quantified, an energy value could not be assigned with security. Some of the byproducts of bacterial action (the short chain fatty acids) can be absorbed through the wall of the colon and provide energy to the human host; others cannot. Presence of a rich supply of energy substrate favors bacterial growth with bacterial utilization of some nitrogenous compounds that otherwise would have been available for absorption/reabsorption by the host. Also, a bulky diet passes more quickly through the intestine than does a refined diet, with an opportunity for increased loss of any or all dietary components (including, importantly, minerals such as calcium, iron, and zinc).

Conventional digestibility trials involve measuring the gross energy (a measure of heat released when a substance is burned in pure oxygen) of food consumed and of the feces (corrected for the amount of energy lost due to the body's inability to oxidize nitrogen completely). The projects had originally considered measuring digestibility of representative diets by animal assay at the University of Arizona. The animal tests were dropped because of limited time and funds.

Human studies require strict control of intake and output not possible under usual field conditions. The Egypt Nutrition Institute did, however, undertake a study of toddlers using the village clinic as a field laboratory. A personal communication from Dr. Wafaa Moussa reports digestible energy to have ranged from 82 to 92% for the six toddlers studied.

The Mexico project used the food composition data base from the INN in which the energy value of carbohydrate from maize is listed as 4.13 kcal/g, the energy value of pure starch. This may have been a typographical error in the original INN table, as the specific energy value is 4.03 (based on 97.5% digestibility) in the source from which the other INN values are taken (6). A lead adult in the study population might have consumed 17 tortillas/d, containing 320 g carbohydrate (17x40g X 47% carb). Energy yield from tortilla carbohydrate would be 1322 kcal by the INN table and 1290 kcal by the specific energy value. If digestibility were, say, 5% lower (92.5% vs. 97.5%) the energy value would be 3.82 kcal/g or 1222 kcal in the example. The difference between the figures calculated by the Mexico project and the extreme assumption is 100 kcal/d. The mean energy intake of Mexican women is reported to be about 2500 kcal/d, so the error of overestimation in this example is about 4% for the day.

The FAO/WHO/UNU (9) recommended that "when diets contain moderately large amounts of fiber, derived from fruit, vegetables, and wholemeal bread, values for available energy calculated as above should probably be reduced by about 5%." This recommendation is consistent with the data and calculations presented above.

Correction for assumed digestibility of energy should have little bearing on any analyses using food intake in the CRSP data sets unless subsequent evaluation reveals marked variation in the consumption pattern of fibrous foods within a project population.

Digestibility of protein is a second issue. The energy lost due to excess nitrogenous matter in feces is, of course, included in tests of digestible energy. The absorption of protein *per se* is, however, of concern in evaluating protein intakes of the populations for comparisons with reference standards. How this is done is described in section VI.D., below.

Absorption of iron is also problematic. The factors that affect absorption of food iron include presence of inhibitors (fiber, phytate, oxalate) and facilitators (vitamin C), as well as the form of the iron (heme vs. nonheme). Discussion of this issue is beyond the scope of the present report. Full evaluation would require close examination of food intake patterns by eating event, and would be worth undertaking if the food data base were, or could be made, reliable with respect to iron.

D. Food Intake

1. Introductory Comments

The measure of food intake should correctly estimate the actual food consumed by each target subject and by the household. For most analytical purposes of the CRSP, it should also estimate well the subject's "usual" food intake, the average intake across time periods of analytic interest. It is recognized that a single day's measure of intake, even if meeting the first criterion, will not meet the second. On the basis of published information (10,11) about intra-individual variability of food energy intake, at the design stage it was accepted that six independent 24-hour observations would be required to assure that the second condition was met. The research plan thus called for two days of quantified food intake each month, permitting analyses by 3-month time intervals. The sampling frame was to be designed so that all days of the week would be included in the data set. Trained interviewers were to be rotated among households. Adherence to a correctly preplanned community schedule would help insure that random events, such as holiday feasts and illness episodes, would be included proportionately in the data set. In practice, no project followed the intended plan to the letter.

The present section of the report presents a comparative summary of the observed energy and nutrient intakes and an examination of the methods actually employed, the difficulties encountered, the unanticipated sources of variation, and other factors that may have introduced either biases or random errors in the final data sets comprising estimated nutrient intakes. The section will provide some guidelines on the expected impact of these biases and random errors on the types of analyses of interest in the CRSP. Issues relating to the food composition data bases used have been addressed in a previous section; obviously these issues carry over to the computed nutrient intakes but they will not be readdressed here.

2. Distribution of Observed Intakes

Table 22 presents a summary of the distributions of observed intakes. These are portrayed graphically in Figures 2-7. In these summations only subjects included in the longitudinal sample defined by ME (see section VI.B.) are included. For each subject, the mean intake across all observations has been computed. For the expression of intakes per kg, all body weight data for the individual were averaged to yield a single estimate of the individual's weight. In toddlers and schoolers, neither the intakes nor the weights used in the calculations are age-adjusted. The statistics represent the distributions of these mean values.

The data presented in Table 22 exclude information from toddlers who are still breast fed, from Egypt project records with missing meals, and from Kenya project records with missing nutrient values. For Mexico, only data collected after Jan 1, 1985 are included. Energy intakes for Egypt have been calculated with 4-9-4 conversion factors pending revision of the food composition tables.

In comparing data from the three country projects, the most striking observation is the sharply lower energy intakes reported for Kenya, whether

expressed as absolute intake or intake per kg body weight. In contrast, there is a general similarity in mean intakes per kg body weight reported for the Egypt and Mexico children. Mexican adults are reported to have higher energy intakes/kg than in Egypt. (This might reflect different levels of physical activity.) The second striking feature is an apparent difference in fat and protein intakes across projects. In the pre-CRSP design phase, sites in different geographic areas were selected in anticipation of finding differences in the nature of the diets consumed, recognizing that this variation might be very important in interpreting the relationships between total food (energy) intake and outcome functions. The data in Table 22 confirm that indeed the macronutrient composition of the diets does differ among sites.

An early concern, during the data collection phase of the CRSP, was that the reported interindividual variability of intake per kg in Egyptian toddlers was unexpectedly high. The data in Table 22 and in the Figures suggest that this still may be true. In the exploration of interviewer bias discussed later, it was found that one interviewer played a dominant role in the collection of food intake for the "outlier" toddlers. The data also suggest that a similar issue may hold for Mexico. The upper range of toddler intakes per kg body weight is surprisingly high. This issue will be explored again in the section addressing the partitioning of variance in dietary data.

Table 22. Observed Distribution of Energy and Nutrient Intakes, ME-defined Longitudinal Sample

A. EGYPT PROJECT

Group	N	Energy	Protein	Carbohydrate	Fat	Animal Prot
		kcal/d	g/d	g/d	g/d	g/d
Toddlers	88	1217 ± 327	39.9 ± 11.9	189.7 ± 53.1	33.1 ± 10.1	18.4 ± 6.9
Schoolers	42	1864 ± 433	63.3 ± 16.5	248.9 ± 69.9	46.2 ± 12.8	26.8 ± 9.0
Men	152	2810 ± 451	99.2 ± 19.7	441.3 ± 73.1	72.1 ± 18.2	43.0 ± 17.0
Women	153	2285 ± 336	78.5 ± 13.3	362.9 ± 58.5	57.7 ± 11.8	32.5 ± 9.5
Preg	62	2228 ± 357	77.5 ± 14.1	347.2 ± 57.5	58.9 ± 15.1	33.5 ± 10.2
Lact	62	2277 ± 389	79.6 ± 17.5	357.0 ± 60.8	58.7 ± 15.1	34.8 ± 15.2
		kcal/kg/d	g/kg/d	g/kg/d	g/kg/d	g/kg/d
Toddlers	88	109.0 ± 33.7	3.58 ± 1.18	17.0 ± 5.46	2.96 ± 0.97	1.64 ± 0.64
Schoolers	42	83.1 ± 23.5	2.82 ± 0.87	13.3 ± 3.75	2.06 ± 0.69	1.20 ± 0.46
Men	152	41.6 ± 7.40	1.47 ± 0.29	6.59 ± 1.28	1.04 ± 0.25	0.63 ± 0.20
Women	153	36.8 ± 8.80	1.25 ± 0.30	5.91 ± 1.47	0.91 ± 0.26	0.50 ± 0.16

B. KENYA PROJECT

Group	N	Energy	Protein	Carbohydrate	Fat	Animal Prot
		kcal/d	g/d	g/d	g/d	g/d
Toddlers	91	844 ± 184	22.9 ± 5.40	166.5 ± 37.0	13.5 ± 4.73	3.2 ± 2.46
Schoolers	129	1447 ± 266	42.8 ± 8.70	286.6 ± 52.4	21.2 ± 5.66	2.5 ± 1.98
Men	214	1954 ± 439	58.2 ± 13.5	379.8 ± 89.2	30.5 ± 8.59	5.2 ± 3.95
Women	214	1672 ± 339	49.2 ± 11.1	328.9 ± 64.5	25.5 ± 7.97	4.1 ± 3.20
Preg	93	1580 ± 377	46.6 ± 12.2	315.4 ± 71.5	23.5 ± 7.88	3.6 ± 2.72
Lact	93	1776 ± 467	53.9 ± 15.5	351.8 ± 88.9	25.9 ± 9.22	3.6 ± 2.84
		kcal/kg/d	g/kg/d	g/kg/d	g/kg/d	g/kg/d
Toddlers	91	84.1 ± 18.1	2.28 ± 0.55	16.59 ± 3.17	1.34 ± 0.47	0.34 ± 0.24
Schoolers	129	72.5 ± 11.9	2.14 ± 0.40	14.36 ± 2.39	1.06 ± 0.26	0.12 ± 0.09
Men	214	35.9 ± 8.4	1.10 ± 0.26	6.98 ± 1.69	0.56 ± 0.16	0.09 ± 0.07
Women	214	32.6 ± 6.7	0.96 ± 0.22	6.41 ± 1.31	0.49 ± 0.15	0.08 ± 0.06

C. MEXICO PROJECT

Group	N	Energy	Protein	Carbohydrate	Fat	Animal Prot
		kcal/d	g/d	g/d	g/d	g/d
Toddlers	54	1119 ± 285	33.8 ± 8.52	193.9 ± 56.7	23.1 ± 8.37	9.6 ± 5.5
Schoolers	79	1828 ± 462	54.1 ± 13.4	336.0 ± 89.2	29.5 ± 10.2	10.2 ± 6.5
Men	104	3070 ± 747	84.3 ± 20.6	496.9 ± 130.9	41.2 ± 15.2	14.9 ± 10.3
Women	143	2572 ± 643	73.1 ± 17.9	455.1 ± 119.4	38.1 ± 13.9	12.5 ± 8.6
Preg	68	2539 ± 533	72.1 ± 16.7	451.9 ± 99.5	38.1 ± 14.9	12.8 ± 9.8
Lact	68	2631 ± 612	73.9 ± 18.4	466.1 ± 111.5	39.4 ± 15.0	12.8 ± 9.8
		kcal/kg/d	g/kg/d	g/kg/d	g/kg/d	g/kg/d
Toddlers	54	102 ± 32.8	3.10 ± 0.96	17.63 ± 6.25	2.08 ± 0.84	0.87 ± 0.48
Schoolers	79	86 ± 21.2	2.55 ± 0.65	15.81 ± 4.14	1.37 ± 0.45	0.47 ± 0.29
Men	104	47 ± 12.2	1.30 ± 0.33	7.63 ± 2.10	0.64 ± 0.24	0.23 ± 0.15
Women	143	44 ± 13.0	1.26 ± 0.36	7.83 ± 2.36	0.65 ± 0.25	0.21 ± 0.14

FIGURE 2. Histogram of Distribution of Energy Intakes of Toddlers, Kcal/d

Percentage

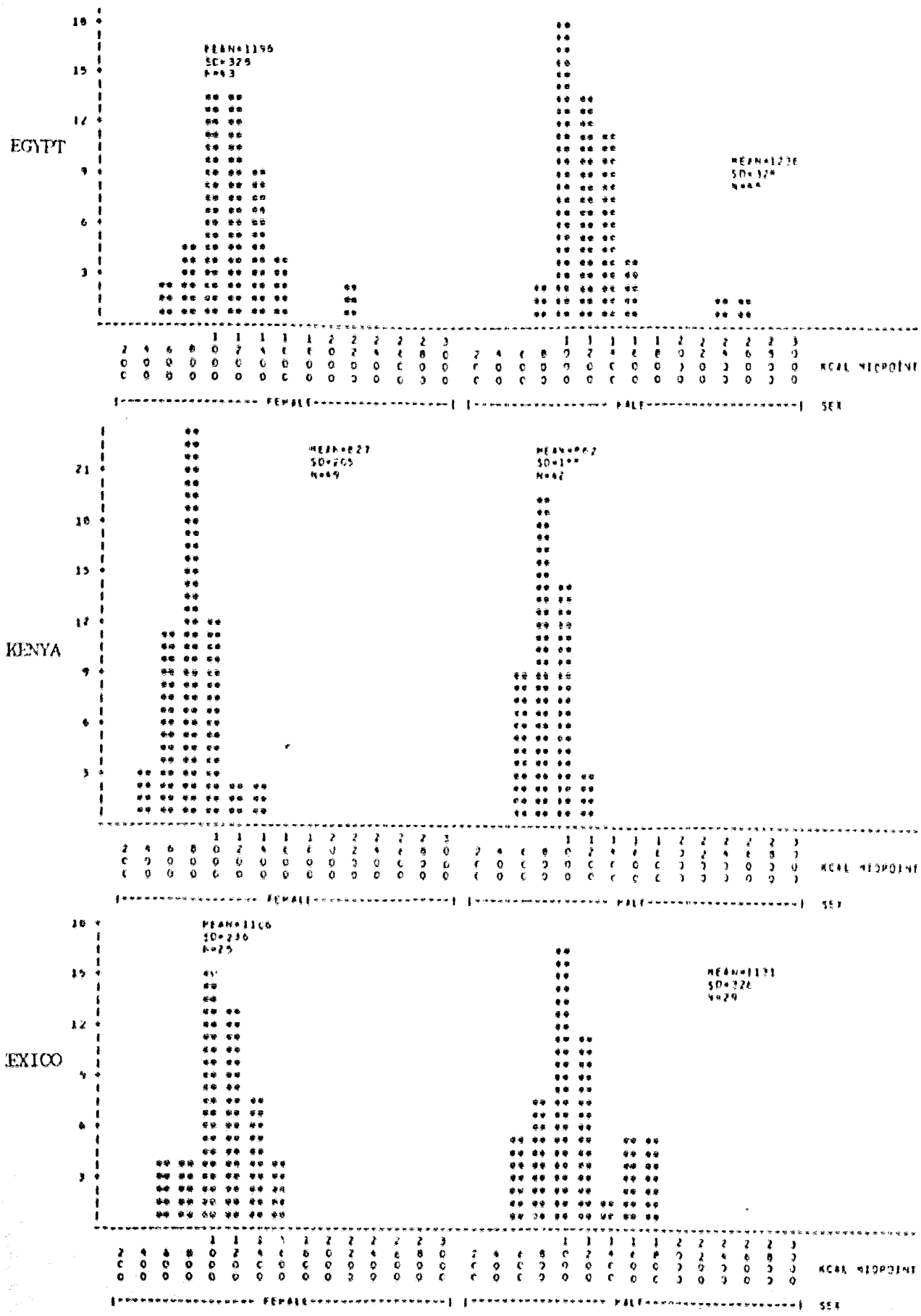


FIGURE 3. Histogram of Distribution of Energy Intakes of Toddlers, Kcal/kg/d

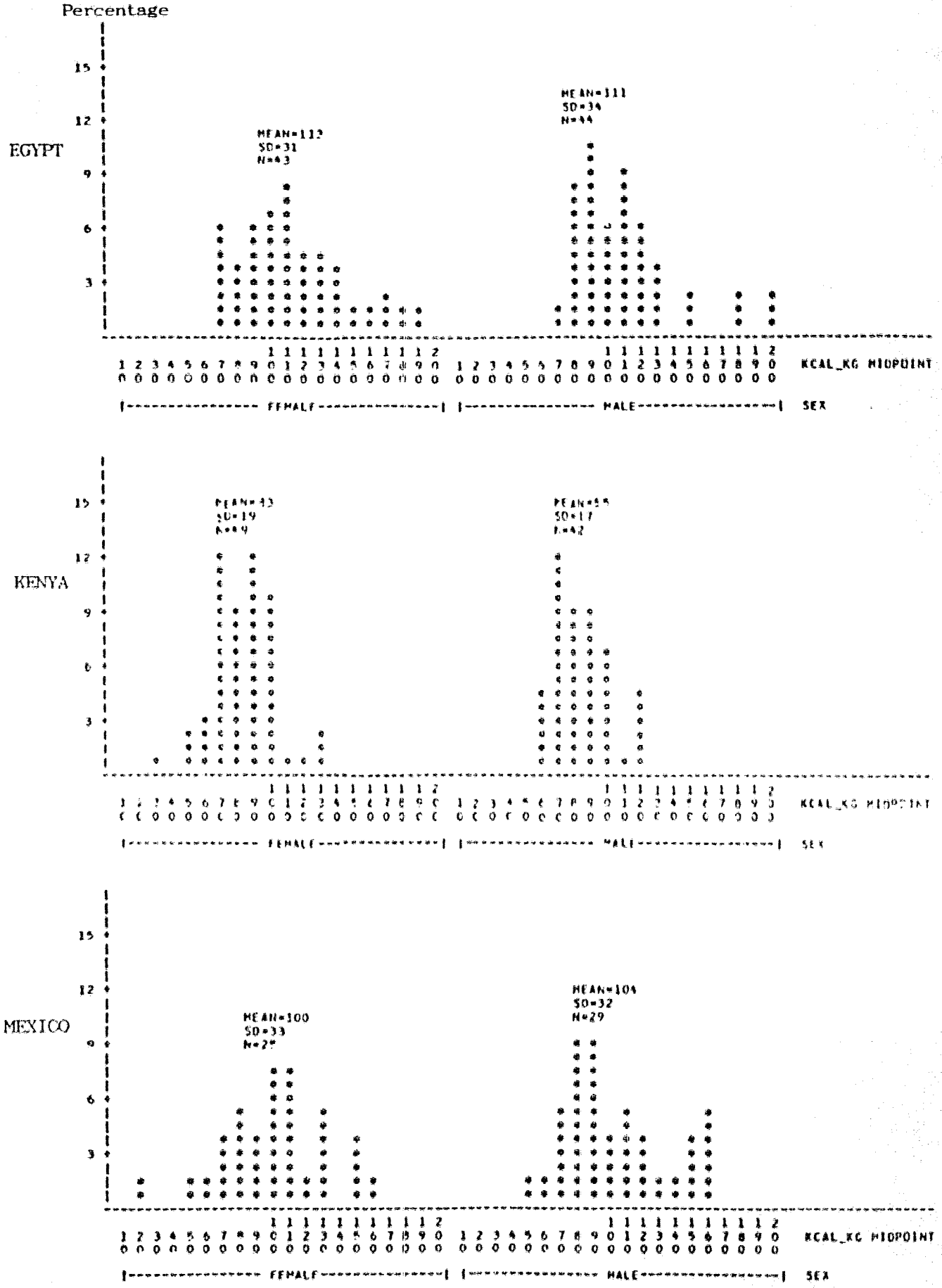


FIGURE 4. Histogram of Distribution of Energy Intakes of Schoolers, Kcal/d

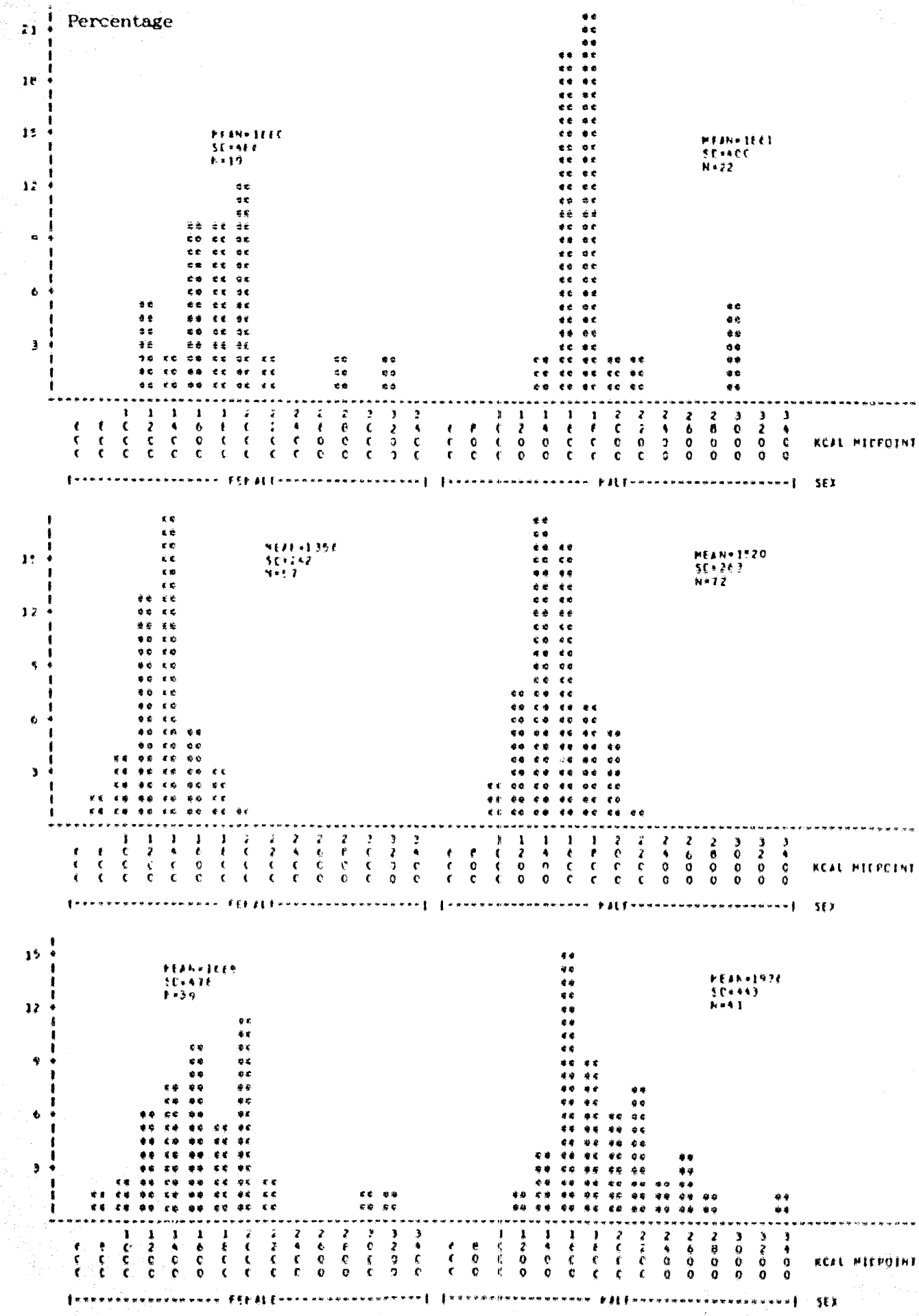


FIGURE 5. Histogram of Distribution of Energy Intakes of Schoolers, Kcal/kg/d

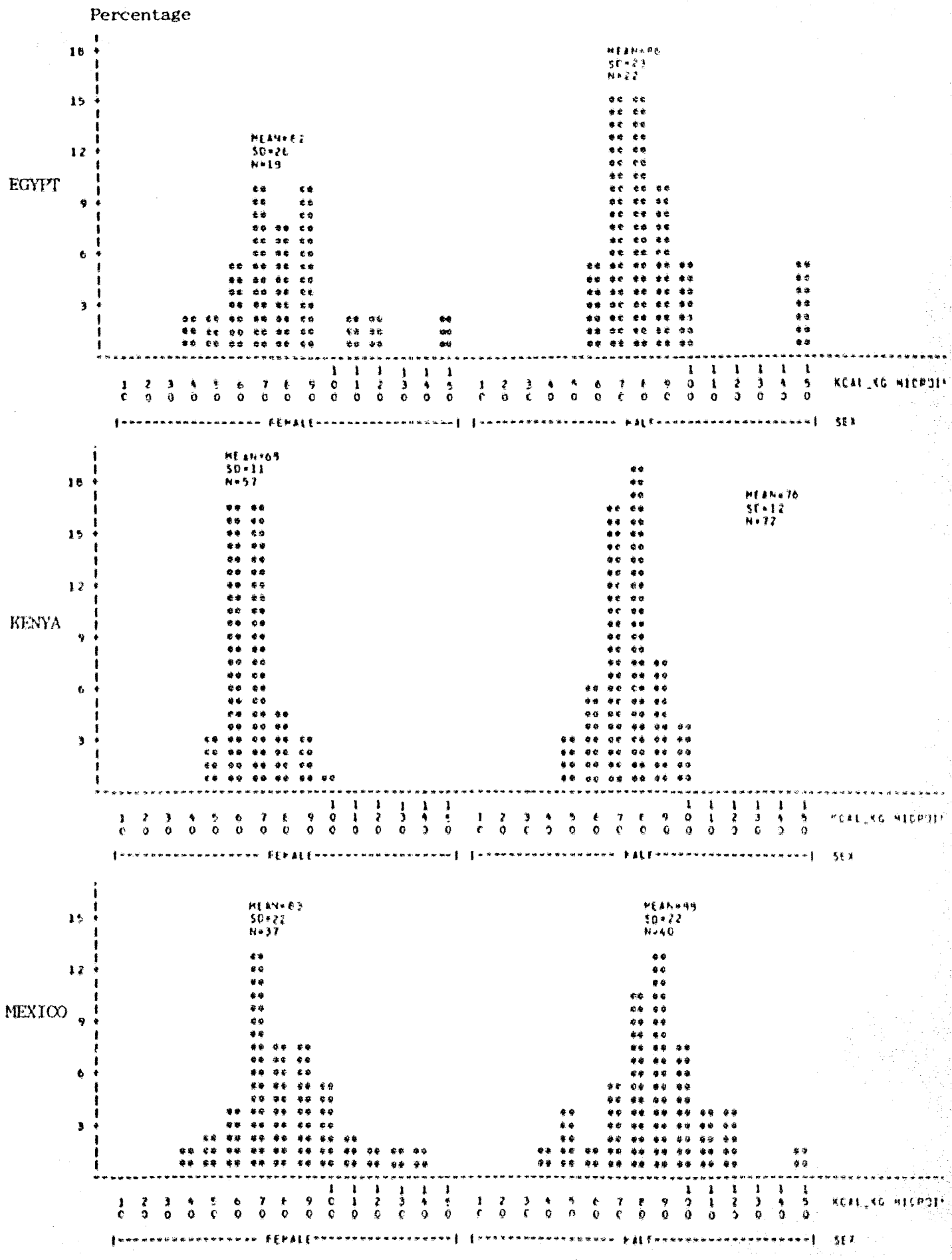


FIGURE 6. Histogram of Distribution of Energy Intakes of Adults, Kcal/d

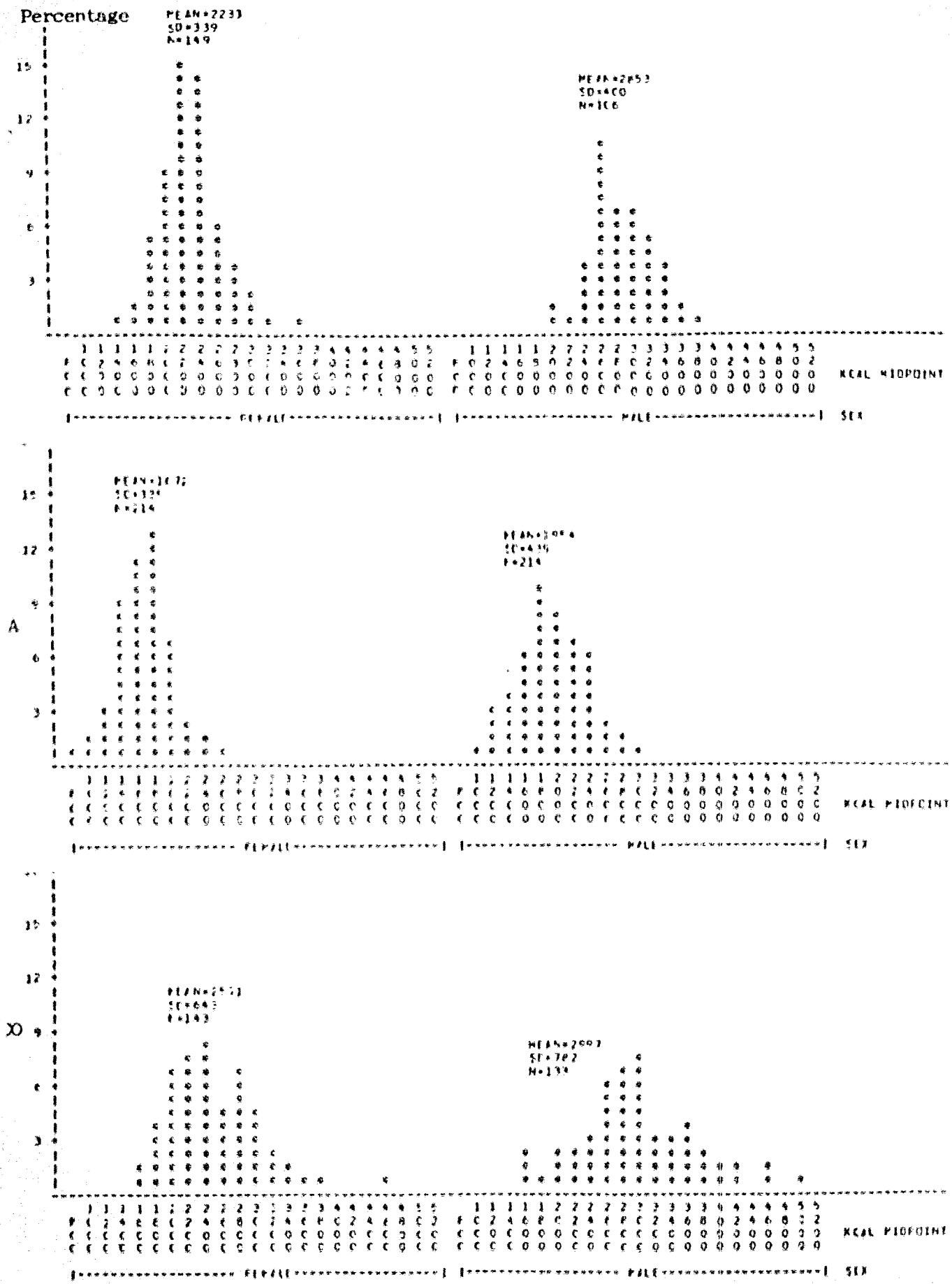
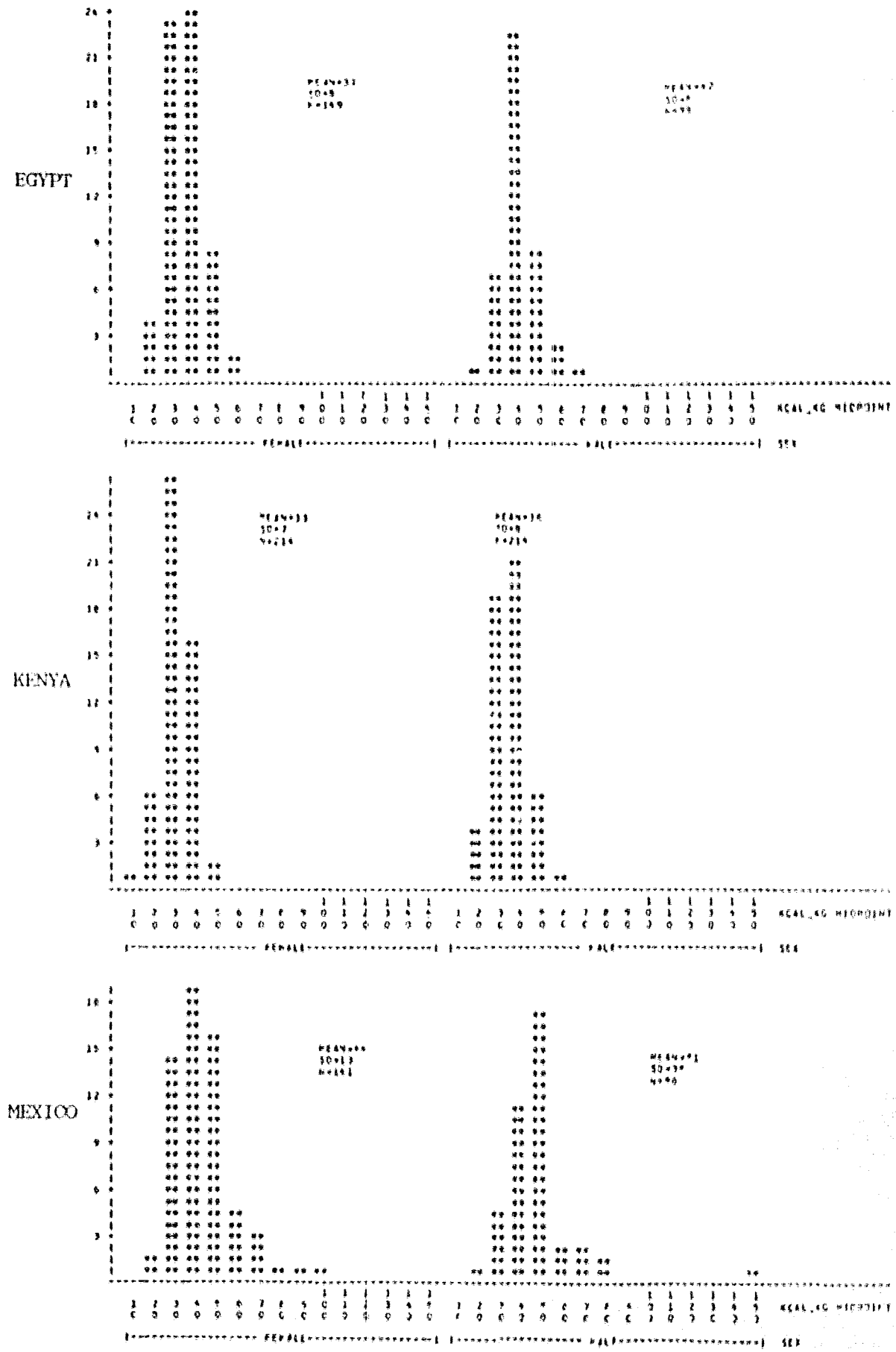


FIGURE 7. Histogram of Distribution of Energy Intakes of Adults, Kcal/kg/d

Percentage



3. Apparent Nutritional Adequacy of Protein Intakes

Intake of most nutrients covaries with energy intake. That is, as the level of total food intake (measured as energy) moves up or down, the levels of constituent nutrients also moves up or down. A concern in the design of the CRSP projects was that associations between energy intake and function might be confounded due to limitation of specific nutrients. Across projects, the relationships between energy and nutrients might be expected to be different according to dietary pattern. In the present section, the question posed is whether or not it is probable that protein intakes are limiting.

Table 23 presents a probability assessment of the adequacy of reported protein intakes for the three projects. The methods of this approach have been described in a National Academy of Sciences publication (10). In essence, this approach asks what is the probability that an observed intake is above or below the requirement of a randomly selected individual of the same class (age, sex, body weight, physiologic state). Given a description of the distribution of nutrient requirements, the probability is assigned as the area under the distribution curve to the right of the intake level. In SAS this can be computed as

```
1 - PROBNORM((INTAKE-MEANREQ)/SDREQ)
```

where MEANREQ and SDREQ are the mean and SD of the requirement distribution for the nutrient in question (here requirements are assumed to be normally distributed). If these probabilities are then averaged for the whole population and multiplied by 100, an estimate of the prevalence of inadequate intakes is obtained. It is this value that is presented in Table 23.

The requirement distributions were based on those developed by the FAO/WHO/UNU (9). The primary requirement estimates were in terms of egg or milk protein. Digestibility of dietary protein for the Mexico and Egypt projects was assumed to be 82%, based on data from studies with seemingly similar diets (9). In the case of the Kenya project, digestibility was considered on an individual basis on the assumption that the digestibility of animal source proteins was 100% relative to that of egg or milk and 80% for vegetable source proteins. In the absence of specific data on amino acid composition, it was assumed that the amino acid scores of diets in all three country projects would be about 70% for toddlers, 90% for schoolers and 100% for adults. Again these assumptions were based on a review of seemingly similar dietaries (9). For Egypt and Mexico, the requirement estimates were increased to account for digestibility and amino acid score,

$$\text{REQadjusted} = \text{REQegg or milk} \times \frac{100}{\text{Digestibility}} \times \frac{100}{\text{Amino Acid Score}}$$

In the case of Kenya, the observed intakes were adjusted downward to estimate utilizable protein equivalent to milk or egg.

$$\text{INTAKEmilk or egg equiv} = \text{INTAKEobserved} \times \frac{\text{Digestibility}}{100} \times \frac{\text{Amino Acid Score}}{100}$$

The intake data used in these analyses represent all intake data available, averaged for each subject across all observations for that subject (only intakes after January 1, 1985 for the Mexico project). In Kenya, records showing an incomplete nutrient intake calculation, and in Egypt, records showing missing meals were excluded. Both cross sectional and longitudinal subjects were included. Weights were averaged over time for each subject (thus in the toddler assessment, neither intake nor weight was specifically age adjusted). This approach takes no account of any periodic swings in dietary intake. It represents an assessment of the individuals' usual diets as described by their mean intakes over the whole period of study.

Table 23. Probability Assessment of Observed Protein Intake

Country	Age/Sex Group	Intake	Requirement	Predicted Prevalence of Inadequate Intakes
		g/kg/d	g/kg/d	% of individuals
Egypt	Toddlers [1]	3.64 ± 1.41 [2]	1.61 ± 0.17	0.4
	Schoolers	2.80 ± 0.88	1.10 ± 0.14	1.8
	Adult Men	1.46 ± 0.34	0.73 ± 0.09	1.2
	Adult Women [3]	1.24 ± 0.30	0.73 ± 0.09	3.6
Kenya	Toddlers	1.36 ± 0.36 [4]	0.92 ± 0.11 [4]	12.5
	Schoolers	1.67 ± 0.38	0.81 ± 0.10	0.7
	Adult Men	0.92 ± 0.24	0.60 ± 0.075	8.3
	Adult Women [3]	0.85 ± 0.22	0.60 ± 0.075	13.3
Mexico [5]	Toddlers	3.00 ± 0.95	1.61 ± 0.17	4.4
	Schoolers	2.57 ± 0.66	1.10 ± 0.14	0.6
	Adult Men	1.29 ± 0.32	0.73 ± 0.09	3.7
	Adult Women [3]	1.24 ± 0.35	0.73 ± 0.09	5.4

[1] Excludes intakes while breast feeding continues
[2] Mean ± Standard Deviation
[3] Pregnant and lactating women not differentiated
[4] Both intake and requirement are expressed as the equivalent of milk or egg protein
[5] Includes only data collected after January, 1985 (new dietary method)

The results in Table 23 suggest that relationships between energy intake and outcome functions are unlikely to be confounded in any major way by covarying protein intake across a range of intake that would be deemed inadequate by existing concepts of protein requirements. This may or may not be true for other nutrients. A preliminary examination of the probable adequacy of iron intakes (see project reports) suggests that iron intake, depending on bioavailability, might be a confounding factor.

4. Sampling of Days - A Potential Error Source

All projects found it efficient to record food intake for two consecutive days each month. It was recognized that this implied some loss of statistical representativeness since adjacent days would be expected to be more similar than the intended independent days; it was a necessary compromise. No project collected information relevant to all 7 days of the week. Each avoided interviews on at least one "special" day, but the impact of this depended upon the dietary methodology. Thus, in the Egypt project no interviews were conducted on Fridays, the religious holiday, but Saturday interviews captured Friday intakes. The missing intake day was actually Thursday, a week day with intake likely to be similar to the other week days. The Mexico project conducted no interviews on Sundays; Monday interviews provided recalled information for Sunday evening and Friday. Saturday schedules ended late Saturday afternoon. Thus the missing intake periods were indeed those which can be expected to be different. The bias would be expected to be appreciably greater in the Mexico data base than in Egypt. For energy, the difference is exacerbated since weekend days are reported to be periods of appreciably higher alcohol intake, particularly for adult males. The impact in Kenya is likely to have fallen between these extremes. Weekend intake days were missed or at least seriously underrepresented, as in Mexico, but alcohol intake on the weekend is believed to be a much smaller problem in Kenya.

There were other differences worthy of note. The Egypt project reported that they had excellent cooperation in gathering information during the Ramadan period of religious observance. The Mexico project collected no dietary information on feast-days. It is believed that the Kenya project followed the same practice as did Mexico.

A different type of sampling bias that might result in either bias or random errors in the final data sets relates to the sampling of days during which illness was present and food intake may have been altered. This is discussed in the morbidity section of the present report.

5. Potential Respondent Error and Bias

The food-intake informant in almost all cases was the lead female in the household. This means that, with the exception of information about her own intake, the information obtained was from a secondary source with potential inaccuracies attributable to incorrect or incomplete knowledge. It also implies that apportionment of food used, in a situation such as Egypt, was a function of her assessment of portions. In Egypt, this was recorded as proportions of what she herself ate; in Kenya, an attempt was made to standardize the estimates of portions through use of marked reference bowls and utensils. It must be assumed that there is an error term that varies across age and sex groups within projects and also between projects. It is not clear that this error will be consistent across individuals and hence present a bias in population mean estimates; it may operate systematically within a household across time and hence result in a bias in the estimate of a particular individual's intake.

A second potential effect exists in all project data sets. Except in Kenya, food intake was unrecorded where the lead female was absent or declined to be interviewed for reason of illness or other. Rescheduling of the food intake measurement, which generally was done, could not correct for potential bias if household members eat differently in the absence/illness of the female head. There is no way of checking the potential error arising from this source.

Intake data in all projects must be accepted as more or less quantitative. The problem of individual cookery practices and food yield has been discussed in the previous section on Food Composition. All projects collected data on the volume or weight of food prepared for the family. Amount consumed by individuals was obtained by a combination of observation and recall of measured portions in the Kenya and Mexico projects. The Egypt project asked the lead female to estimate the portion of a mixed dish she consumed and intake of other family members was recorded as fractions or multiples of her consumption level; thus a portion of the individual food intake record cannot be quantified precisely.

Individual food intake records must include food eaten away from home and take account of food carried from home but not eaten. It is universally true that it is easier for a mother to know what she has served to her children than what they have eaten. Interviewers are said to have observed toddler intake "when possible"; the possibility would vary among the projects by methodology and mobility of the toddler. By schedule, Kenya interviewers spent the most time in a household, Egypt interviewers the least. School-age children vary in the reliability of their self-reported intake. No project has provided documentation of a fully validated comparison of intakes of these age groups by the established method and continuous direct observation (child-following), so the magnitude of error, if any, is unknown. All projects experienced some difficulty in obtaining information about lead males' intake away from home; the data set is less complete for this category of subjects as a result, and may be biased in reflecting predominantly intake by men who ate household food exclusively or were, or appeared to be, more willing respondents.

6. Estimation of Alcohol Intake

The Kenya project reports the likelihood that alcohol intake is underreported within the food intake records. Given that the Mexico project did not collect dietary data pertinent to the day of highest alcohol consumption, it is clear that there was underreporting. In the Egypt project, this was likely to have been a minimal problem given that few persons consume alcoholic beverages. In Mexico, because of the reported pattern of consumption across classes of individuals, the bias in estimated energy intake can be expected to be greatest in men but also present in women and chil'ren. In Kenya, the bias is likely to be restricted largely to adult men.

The Mexico and Kenya projects have independent estimates of alcohol consumption. For Kenya, reported patterns of intake are available for almost all of the adults in the longitudinal sample. For some analyses of the multiple regression type, this supplementary information can be used with considerable statistical advantage as an additional independent variable and

hence at least partially adjust for the bias in the food intake-based energy intake estimates. It is not possible to blend the two sources of information into an adjusted estimate of intake. In the Kenya and Mexico data sets, individuals can be classified as consumers and non-consumers of alcohol, based on their own statements. The basic question then arises as to what extent this binary information can be used to compensate for the bias incurred by not actually measuring the amount of alcohol consumed.

If alcohol is related either directly or indirectly to both the independent and dependent variables, then exclusion will typically cause the analysis of other variables to be biased. The magnitude of this bias is related to the strength of relationship between alcohol and the other variables in the analysis. In analyses, for example, of lead males, this bias is likely to have a substantial influence when total calories or a function of total calories is part of the analysis since alcohol consumption, in some situations, is an important part of the energy intake. Employing the binary variable (did or did not drink alcoholic beverages) in place of the underlying variable (total amount of alcohol consumed) is preferable to ignoring the influence of alcohol altogether. Inclusion of an imperfectly measured variable induces a certain amount of confounder bias but less than the bias incurred by excluding the variable entirely. Omission of a variable like alcohol which is highly correlated with other important variables of interest is likely to lead to substantial and unknown biases with respect to understanding other relationships within the data set.

When adults drink to the level of inebriation, the food intake of children may suffer. According to the MFR (Tables 23.4 and 23.5), 12% of women and 80% of men acknowledge drinking to the point of inebriation with varying frequency. Information to address this potential effect on intake is not available in either the Mexico or Kenya data bases. This might be seen as an issue akin to that arising when a mother is ill - food intake interviews are rescheduled and we have no direct knowledge of the impact of the illness on the intakes of others.

7. Methodology Changes - Mexico Project

As discussed in the Mexico project report, the procedures for collection of food intake were altered during the course of the project. Specifically, two changes took place a few months apart. On October 1, 1984, the interviewer protocol and reporting forms were altered, the new protocols and forms being designed to capture family as well as individual intakes. Three months later, on January 1, 1985, a reorganization of interview responsibilities was implemented and former supervisors became the operational interviewers for food intake data, and the previous system of rotation of interviewers was discontinued in an effort to improve relationships with the households. These changes are detailed in the project report.

Unfortunately both changes impacted on the estimated food intake of target individuals. An analysis of the change in estimated intake among toddlers, schoolers and lead females is presented in the Mexico report. That analysis was conducted with cohorts of subjects who had been studied in periods 1 and 2

and in periods 2 and 3 using paired analysis techniques. The order of magnitude of possible effects is shown in Table 24.

Table 24. Estimated Effect of Dietary Methodology Changes
Mexico Project *
(kcal/d)

Target Group	Period 1 to 2		Period 2 to 3	
	N	Mean Change	N	Mean Change
Toddlers	32	+ 145	58	+ 162
Schoolers	71	+ 223	52	+ 320 (Males)
			56	+ 374 (Females)
Lead Females	101	+ 245	101	+ 344

* For full analyses, see Mexico Project Report. These estimates of change cannot be generalized across the whole sample. The analyses were designed to determine whether or not there was a statistically significant effect rather than to estimate a population bias measure.

In a further exploration of the impact of the methodology changes a cohort of adult males who had been in the Mexico study through both methodology changes, and for sometime thereafter, was identified. The criterion of selection was the existence of food intake records before July 1, 1984 and after Nov 30, 1985. Twenty-eight subjects were identified. For each of these subjects, all intake data prior to October 1, 1984 were pooled to obtain a single mean intake estimate. Similarly, data for the period October 1 to December 31, 1984 and for the period after January 1, 1985 were pooled to obtain mean intakes for each period for each subject. Since the periods were of different lengths and hence included different numbers of intake observations (actual number of observations per subject were 1-13 for period 1, 2-6 for period 2 and 6-34 for period 3), the confidence intervals around these subject-period means differed; this was ignored in the comparisons presented in Table 25.

The analyses presented in Table 25 not only document that the methodology change affects estimates of energy intake for adult men, as demonstrated for other target groups in the project analyses presented in Table 22, but extend this to nutrient intake and measures of "diet quality". At least for these 28 men, it appears that the methodological change had differential effects on different aspects of the diet.

As was noted in the original project analyses, the present analyses do not rule out the possibility that the methodology changes were concurrent with a general change in the nature of the diet consumed in the community, although this seems extremely unlikely.

These observations have major implications for design and interpretation of data analyses. Until and unless the source of the difference in intake estimates is ascertained, it is difficult to render judgment on which methodology is more valid. The Mexico project presents validation comparisons

for the final methodology. In analyses presented in the Management report, only food intake data collected with the final methodology are reported.

Table 25. Effect of Methodology Changes on Estimated Nutrient Intakes Among 28 Adult Males
Mexico Project

Nutrient	Mean Intakes			Difference Between Periods		
	Period 1	Period 2	Period 3	2 - 1	3 - 2	3 - 1
Energy (kcal/d)	2334 * (165)	2763 (165)	3067 (116)	540 *** (155)	323 (176)	732 *** (165)
Animal Energy (kcal/d)	49.7 (10.?)	112.0 (21.2)	117.3 (27.2)	57.6 *** (18.4)	8.2 (19.0)	67.6 *** (21.5)
Total Protein (g/d)	57.2 (3.79)	70.9 (3.63)	81.2 (3.27)	15.7 *** (3.44)	10.9 (4.22)	24.0 *** (4.17)
Animal Prot (g/d)	3.38 (0.68)	6.75 (1.18)	6.61 (1.60)	3.06 ** (1.05)	0.03 (1.16)	3.22 (1.27)
Fat (g/d)	21.0 (1.93)	33.4 (3.22)	42.3 (3.10)	13.2 *** (3.16)	9.93 *** (2.72)	21.3 *** (2.91)
Carbohydrate (g/d)	384 (27.5)	422 (24.4)	481 (18.0)	57.1 ** (21.0)	64.3 ** (25.3)	97.5 *** (26.8)
An Energy **** (% kcal)	2.68 (0.63)	4.72 (1.00)	4.01 (0.86)	1.77 (0.66)	-0.64 (0.46)	1.32 (0.67)
Prot Energy (% kcal)	22.9 (1.52)	28.4 (1.45)	32.5 (1.31)	6.29 *** (1.37)	4.36 ** (1.69)	9.59 *** (1.67)
An Prot Energy**** (% kcal)	1.35 (0.27)	2.69 (0.47)	2.64 (0.64)	1.22 ** (0.42)	0.01 (0.47)	1.29 ** (0.51)
An Prot % **** (% total prot)	5.80 (1.16)	9.64 (1.70)	7.70 (1.58)	3.31 ** (1.17)	-1.80 (0.93)	1.91 (1.14)

N = 28 subjects represented in periods 1 and 3, 25 also with data in period 2

* Mean (Standard Error)

** Statistically significant at 0.02 to 0.002 level

*** Statistically significant at at least 0.002 level

**** Animal source energy and protein calculated before revision of food composition tables; contrast between periods is believed to be valid but absolute levels are incorrect.

8. Interviewer Effects

All projects used multiple interviewers to collect the food intake data. However, the method of assigning interviewers to households varied among the three projects. Kenya Project used two teams of interviewers for each of the four clusters. Since each team consisted of a supervisor and 4-5 enumerators, households were typically interviewed by 8 to 10 different interviewers. Rotating schedules ensured that each household was interviewed by multiple interviewers, thus minimizing the effect of any interviewer bias on the mean

intake of the individual (but leaving open the possibility of interviewer biases contributing to the within-person variability of intake). By comparison, the Egypt Project found it was not feasible to rotate interviewers, due to objections from the subjects. Therefore, households had a main interviewer who was replaced only in the case of an illness or vacation. In general, interviewers were assigned households in the same or adjacent blocks. Mexico Project rotated interviewers primarily within villages, but to a lesser extent than for the Kenya Project - a household typically had 2 to 3 different interviewers after the 1/1/85 protocol change.

Interviewers had differing backgrounds, both between projects, and within projects. Kenya Project enumerators were fieldworkers without any nutrition training, supervised by a trained worker. Egypt Project interviewers all had a degree in dietetics but varied in experience. After the 1/1/85 protocol change, Mexico Project interviewers were all licensed nutritionists. Increasing social distance between interviewer and subject may lead to unrepresentative, inflated estimates of intake as subjects improve, or report improved, diets during days of observation.

As an initial screen for interviewer effects, Table 26 presents the range of mean energy intakes estimated from data collected by individual interviewers. Only interviewers with at least 10 observations (days) are included.

Table 26. Range of Mean Energy Intakes by Interviewer

	Kenya	Egypt *	Mexico
Number of interviewers	20	29	19
Range, Lead males	34% **	66%	29%
Range, Lead females	27%	65%	29%
Range, Toddlers	56%	116%	48%

* Egypt energy intakes calculated with old food composition data
** Calculated as the mean for the highest interviewer minus the mean for the lowest interviewer, as a percent of the overall mean for that target type.

If there were complete rotation of interviewers among all households, and a sufficiently large number of collections by each interviewer for each household, the range of mean observed energy intakes among interviewers should be small if there is no interviewer difference. For practical reasons, this type of collection was not possible for any project. It follows that at least part of the range seen is attributable to sampling variation (inadequate sample size for a fully reliable estimate of the true interviewer mean). This may be exacerbated by true differences between communities which confound with differences between interviewers (interviewers were seldom rotated among communities or clusters so this is not a major consideration). The ranges overestimate true interviewer effects. For Mexico and Kenya, this screen does not establish either the presence or absence of interviewer effects.

The large interviewer ranges seen for Egypt strongly suggest an interviewer effect. The ranges do reflect to some extent the almost complete lack of rotation of interviewers among households. Inter-household variation will be included in the ranges to a much greater degree than in the Mexico and Kenya data. Nevertheless, the ranges are unexpectedly high.

To further examine the Egypt Project data, toddler inter-interviewer variation was selected. Table 27 presents the mean energy intakes reported by the 22 interviewers who collected at least 10 days of intake data from non-breastfed toddlers. The mean intake overall for these toddlers was 1173 kcal per day.

Table 27. Mean Toddler Energy* Intakes by Egypt Project Interviewers

Interviewer	Mean energy	Number of days
50	1897	209
51	1188	148
53	900	113
54	1305	189
55	1121	174
56	836	33
57	1020	260
58	947	144
59	1055	150
60	665	42
61	1185	126
64	1135	28
65	724	82
66	952	30
67	1021	90
68	1286	25
150	1035	66
151	1464	16
152	964	33
154	1066	12
156	1366	53
158	1331	39

* Energy intakes computed using old food composition data base.

Interviewers with numbers from 50 through 59 were senior dietitians, with the lowest number being the most senior. Numbers 60 through 62 were Ph.D. students, and the remaining numbers were junior dietitians. Senior dietitians were given households nearest the Health Center (primarily blocks 1 and 2), although they did collect dietary data from other blocks as well.

The lowest interviewer (60) and the highest (50) were examined in more detail. Interviewer 60 was the second lowest of the 29 interviewers for lead males, but not among the lowest 3 interviewers for lead females, schoolers, or households. Interviewer 50 was the highest of the 29 interviewers for

households and schoolers, as well as for toddlers, and the second highest interviewer for lead males and lead females.

As a more definitive approach, the mean intakes reported by different interviewers for the same toddler were compared. Only interviewers reporting at least 2 days of intake for the target subject were included. The target toddlers selected for interviewer 50 were those with reported energy intakes above 150 kcal/kg (the outlying high intake toddlers shown in Fig 3 were included). For interviewer 60, all toddlers interviewed 2 or more times were included. In addition to interviewers 60 and 50, the two interviewers with the largest numbers of reporting days for each of the selected toddlers are included. The results are presented in Table 28.

Table 28. Energy Intakes* Reported by Different Interviewers for the Same Toddler
Egypt Project

Toddler ID	Int. #	N	Mean	Int. #	N	Mean	Int. #	N	Mean
70189407	60	2	971	55	21	1101	56	2	759
90061410	60	2	564	51	4	1207	67	6	943
90092406	60	6	379	65	4	480	152	3	837
90114406	60	4	541	65	2	655	158	3	1191
100031408	60	6	694	150	4	973	52	2	1020
110041405	60	11	642	56	3	1205	154	4	1010
121059407	60	8	649	53	2	725	65	2	709
10023406	50	2	2609	156	7	1569			
10087410	50	11	1948	61	4	1187	67	10	1343
10136406	50	11	2465	153	4	1115			
20019407	50	16	1602	150	3	1113			
30015407	50	18	2600	150	2	930			
30192406	50	8	1928	55	2	2952	150	2	1174
70068404	50	7	2627	56	2	1147	65	2	504

* Energy intakes computed using old food composition data base.

From the data in Table 28, interviewer 60 is the lowest of the three interviewers for six of the seven toddlers. Interviewer 50 is the highest of the two or three interviewers for six of the seven toddlers. She was the dominant interviewer for the high intake outlying toddlers in Figure 3.

From these limited examinations, it appears very likely there is a true difference among interviewers in the Egypt Project intake data. Because of the design of interviewer assignments in that project, the effect of the interviewer bias would operate spuriously to increase the between subject variability as well as potentially leading to abrupt changes in apparent intake during those periods that there was substitution of interviewers. This effect has been taken into account in the Egypt project analyses.

The original screening (Table 26) suggests that there may be some interviewer effects in the Kenya and Mexico studies. Here the effects would operate to increase the within-person variance more than the between subject differences. That is, because of rotation of interviewers, there may be minimal bias of the estimates of a subject's mean intake over time but increased "noise" in the levels of intake from day to day. This conclusion should be tempered by the observation of some very high toddler intakes per kg body weight (Fig 3). There may be an important interviewer effect not detected in the preliminary screen conducted by ME.

If these assessments are correct, interviewer bias will impact differently in the different data sets. It is a matter that should be examined in greater detail in later analyses.

9. Partitioning of Variance in Dietary Data

As noted previously, in the original design phase of the Nutrition CRSP, the expectation was that relationships between food intake and other variables would be based on three month averages, perhaps moving averages. The necessity to pool data arose from the recognition that there is major day-to-day variation in intake within individuals. Most of the relationships that would be of interest were between intakes persisting over moderate periods of time, rather than on a single day, and an outcome variable. Based on limited data available in the literature at that time (12), the assumption was that energy intake might exhibit a within-person variability (SD) equivalent to about 25% of the mean intake. It was hoped that the sample selection would maintain a variability between people of at least that amount since the literature available suggests that in normal populations the within and between-person variances were about equal (10). By pooling data over 3 months (6 measurements of intake), the within-person variance of three month mean would be about 10% for energy and the variance ratio ($SD_{within}^2/SD_{between}^2$) would be about 0.16. It was felt that this would be satisfactory for analyses, giving an acceptable level of attenuation of statistical relationships.

In Phase I of the field studies, the Egypt and Mexico projects examined the partitioning of variance in data collected as the food intake methodologies were developed and tested. It appeared that the variances were reasonably close to the target assumptions. In the present report, the partitioning of variance in the final research data sets has been examined and described for each of the projects. The implication of the results of these examinations for the interpretation of statistical analyses within individual projects and for the interpretation of comparisons between projects is then discussed.

a. Partitioning of variance in the Kenya project.

i. Toddler intake data.

Table 29 presents the results of ANOVA runs (partitioning of variance between the subject or "model" effect and the residual error in a one-way ANOVA). Standard Deviations have been calculated without taking into account the unbalanced design. The data base for these analyses consisted of 2603

observations of one day food intakes of 109 toddlers each of whom contributed at least 12 observations between the ages of 17.0 and 32.0 months. The average number of observations per child was 23.9 with a range from 12 to 30. All toddlers included also had recorded birth dates and at least 6 measurements of body weight. The selection coincides with the longitudinal sample and includes also specific rules for the correction of sex classification and exclusion of outliers (see ME Annex describing these rules). For these analyses, weight on the day that food intake was collected was imputed from the linear regression of weight on date run for each child.

It is readily apparent from the data in Table 29 that there was substantially more within-person variance in the Kenya food intakes than had been expected in the original design. Expressed as the Coefficient of Variation (CV), the within-person variability for energy intake was 47.9% while the between-person variability was 19.3% (compare to design assumptions of 25% each). The result of these departures was a greatly increased variance ratio (6.1 compared to 1.0) and as discussed below, a greatly weakened statistical power. In the Table, the variance ratios for all of the nutrients measured are presented. They are generally similar except for fat which has a substantially smaller ratio. The within-person variance has not been inflated because of progressive change in weight (and hence intake) across time; when intakes were expressed per kg imputed body weight, the variance ratios were similar to those of absolute intakes. For some nutrients, the analyses have been run for intakes expressed per 1000 kcal (a measure of "dietary quality"). The variance ratios tend to be higher. Part of the variance between subjects has been removed by controlling for total amount of food consumed and the variability in the nutrient density of foods from day to day is exposed. The effect is particularly striking for total protein.

Table 29. Estimates of the Partitioning of Variance in Kenya Toddler Intakes

Nutrient Vector	Mean	SDInter	SDIntra	Vratio
Energy				
kcal/d	856.1	165.2 (19.3)*	410.3 (47.9)	6.16
kcal/kg/d	83.61	15.24 (18.2)	39.96 (47.8)	6.87
Total Protein				
g/d	23.25	4.77 (20.5)	13.63 (58.6)	8.14
g/kg/d	2.26	0.454 (20.0)	1.29 (57.0)	8.10
g/1000 kcal/d	26.94	1.95 (7.2)	7.90 (29.3)	16.35
Animal Protein				
g/d	3.49	2.25 (64.4)	5.79 (165.9)	6.62
g/kg/d	0.342	0.225 (65.7)	0.555 (162.3)	6.07
g/1000 kcal/d	4.25	2.31 (54.4)	6.03 (141.9)	6.81
Fat				
g/d	14.27	5.06 (41.0)	9.68 (67.8)	3.65
g/kg/d	1.39	0.488 (35.1)	0.963 (69.3)	3.89
g/1000 kcal/d	16.85	3.96 (23.5)	8.93 (53.0)	5.10
Iron				
mg/d	6.46	1.49 (23.0)	4.51 (69.8)	9.12
mg/1000 kcal/d	7.76	1.33 (17.4)	3.98 (51.3)	9.01
Carbohydrate				
g/d	167.5	31.3 (18.7)	84.7 (50.6)	7.34
Vegetable Protein				
g/d	19.76	4.10 (20.7)	12.51 (63.3)	9.31

* Coefficient of Variation = 100 x SD/Mean

An attempt was made to identify any major sources of the within-person variance in the Kenya toddler food intakes. As noted above, it was not attributable to differences in body weight across time. Age does seem to affect the total energy intake. From regression analyses, the mean change in energy intake per month of age appears to be about 24 kcal.

Two other potential sources of the within-person variance were considered: that there was a marked difference in variability of intake attributable to household food shortages in the drought period, and that the intake methodology had the effect of attributing intake to the wrong day in a pair. As a first approach, scatter plots of data were visually examined. What was striking about these plots was not a marked change during the "famine" period but rather the very high variation throughout the whole period and even between adjacent days. Whatever is the source of the variance, it does not appear to be the food shortages and potential "feast and famine" situation of the drought period.

To further examine this question the data base was searched for pairs of data - data collected on adjacent days from the same subject. From a total of 2603 records, 1260 such pairs were identified. The intakes recorded for the two days were then compared and the difference was tested to see if it differed from 0. The results are presented in Table 30.

Table 30. Comparison of Intakes Recorded on Adjacent Days
Kenya Toddlers

Nutrient Vector	Mean Intake		Difference in Intake *		
	Day 1	Day 2	(Day 1-Day 2)	T	Prob>|T|
Absolute Intake					
Energy kcal	861.0 ± 434.5	856.8 ± 446.7	4.20 ± 489.0	0.30	0.76
Protein g	23.2 ± 13.8	23.5 ± 15.1	-0.31 ± 15.9	-0.69	0.49
Fat g	14.3 ± 10.4	14.3 ± 11.2	-0.01 ± 11.1	-0.03	0.97
Intake /1000 kcal					
Protein g	26.8 ± 8.15	27.1 ± 8.10	-0.30 ± 8.66	-1.21	0.22
Fat g	16.8 ± 9.38	14.3 ± 11.2	-0.13 ± 9.98	-0.45	0.65

* N = 1260 pairs

There is no evidence of a systematic bias between days of a pair. The very high variability of the difference between adjacent days was surprising, particularly with regard to the difference in nutrient density. This would suggest rather major changes in the nature of the diet consumed, as well as the amount, even on adjacent days. In a further exploration, an Analysis of Variance (PROC ANOVA, SAS PC) was run including subject and day (1 or 2) in the model (2520 records for 109 subjects). As expected, day explained negligible variance. Day 1 and 2 were then averaged to give a single 2 day estimate. Again an ANOVA was run with only subject in the model (1260 records for 109 subjects). From both runs the unexplained variances were examined to see whether pooling the two days encompassed within the pairs had a different effect than one would expect from pooling sets of two random days from the subjects record. The results for energy intake are shown in Table 31. The observed residual error from pooled data was slightly larger than predicted from the unpooled data ANOVA. A reasonable interpretation of this is that the variance of the data within pairs of days is generally comparable (marginally smaller) than the within-person variance in the data set in general. (From the series of analyses it can be inferred that the correlation between the adjacent days was about 0.5.) There is very little similarity between days; given the very large within-person variance across time, this is a surprising finding; it might have been expected that there were highs and lows in intakes attributable to external moderate term cycles (with greater similarity between adjacent days than across these cycles). That does not appear to be the case. These models were run with data sets including only the pairs of data (1260 pairs).

Table 31. Comparison of Observed Unexplained Variance in Pooled Two Day Data With Predicted Variance of Random Two Day Means
Energy Intake, kcal/d
Kenya Toddlers

Estimation Basis	Variance Not Explained by Subject	
	Of 1 Day Data	Of 2 day Data
ANOVA, Separate Days	166315.70 *	83157.85 **
ANOVA, Pooled Means of Pairs of Days	-	108785.74 *

* Observed Mean Square Error from ANOVA, Model Energy = SUBJID
** Calculated as Mean Square Error / 2 for 2 day data

As a test of the hypothesis that major swings in intake were attributable to external factors changing in a cyclic manner, partitioning of variance was compared between the period of drought effects and the period after the drought. That is, the hypothesis in hand was that, with limited and perhaps spasmodic food supply in the aftermath of the drought, variability would be affected. In this analysis, data collected between July 1 - Dec 31, 1984 were compared with data collected after March 30, 1985. Subjects were not identical in the two data sets and age was not controlled; subjects included contributed at least 6 observations in the period examined. The results of these analyses, for selected variables, are presented in Table 32.

Table 32. Comparison of the Partitioning of Variance in Two Time Periods, a Period of Food Shortage and a Period of Stability
Energy, kcal/d

Variable / Time Period	Records	Subjects	Mean	SDinter	SDintra	Variance Ratio
Energy, kcal/d						
July 1-Dec 31, 1984	1157	105	850.6	188.5	393.3	4.35
After Mar 30, 1985	193	22	944.3	148.6	384.9	6.70
Protein/1000 kcal/d						
July 1-Dec 31, 1984	1157	105	26.19	2.04	7.43	13.25
After Mar 30, 1985	193	22	30.24	2.64	6.36	5.85
Fat/1000 kcal/d						
July 1-Dec 31, 1984	1157	105	16.65	4.21	8.76	4.34
After Mar 30, 1985	193	22	15.47	2.82	6.18	4.79

Includes only subjects with at least 6 records in the specified time period; subjects not same in the two periods.

From Table 32, it may be inferred that during the period of food shortages, the day to day variability of total intake was generally similar to that seen

in a period of relative stability and indeed with that seen through the whole period (Table 29). Although the period of food shortage does not explain the high within-person variance in the total intake data set, fat intake expressed in relation to energy intake (a measure of the composition of the diet) may show a higher variability between subjects, and within subjects, in this period.

The analyses in Table 32 were not controlled for age. As noted previously, intake appears to change by about 24 kcal/month of age. Since there is clearly an age bias between the groups compared in Table 25, this would explain a part of the difference in mean intakes observed. Obviously, age at the time of food intake and the date of collection of intake are related. Linear regressions were run to examine this with the following results (PROC REG, SAS PC):

KCAL = 787.12 + 6.93 MONTH	Rsquare = 0.0046 Slope of MONTH, P < 0.0006
KCAL = 625.19 + 13.66 AGEFD	Rsquare = 0.0143 Slope of AGEFD, P < 0.0001
KCAL = 391.93 + 23.32 AGEFD - 10.42 MONTH	Rsquare = 0.0175 Slope of AGEFD, P < 0.0001 Slope of MONTH, P < 0.0036

where KCAL = reported energy intake, kcal/d
AGEFD = age in months when intake was estimated
MONTH = month in which intake was measured with Jan, 1984 = 1

If the analyses were also controlled for Subject, the slopes for age and month were marginally higher (23.41 and -10.51) and the Rsquare was unchanged. Subject explained very little variance after age and month of observation had been taken into account. The original slope estimate was the mean of regressions on individuals.

It was noted that both body weight and energy intake exhibited wide within-person variability. It might be hypothesized that some external factor(s) characteristic of the child's environment, affected the stability of both food intake and weight. In a test of the association between these variabilities, a regression of the SD of all intakes observed for each individual on the SD of all weights for the same individual was run. With 109 subjects in the data set, the R square (0.0008) offered no support for the suggestion that the same external factor might be driving variability (instability) in both of these parameters. Associations between intake and weight within individuals across time have been examined within the Kenya project report.

ii. Schooler intake data.

Analysis of variance of intake for schoolers included in the longitudinal sample yielded the data shown in Table 33. The picture was very similar to that seen in the younger children. Again variance ratios for all vectors examined were very high but that for fat intake was somewhat lower. It is clear from the SDintra estimate for animal protein that there is skewing of

the data. No attempt was made to normalize the data for these analyses (untransformed data have been used in project analyses).

Table 33. Estimates of the Partitioning of Variance in Kenya Schooler Intakes

Nutrient Vector	Mean	SDinter	SDintra	Vratio
Energy, kcal/d	1444	238.6 (16.5)*	573.8 (39.7)	5.78
Total Protein, g/d	42.7	7.76 (18.1)	19.43 (45.5)	6.27
Animal Protein, g/d	2.46	1.73 (70.3)	4.37 (177.6)	6.41
Fat, g/d	21.1	5.11 (24.2)	10.32 (48.9)	4.08

Longitudinal sample only: Subjects = 141; Mean observations/person = 21.3
* Coefficient of Variation = 100 x SD/Mean

iii. Adult intake data.

The partitioning of variance observed in the intake data sets for adult men and women are presented in Table 34. Subjects selected were from households defined by ME as being in the longitudinal sample. Again evidence of skewing is present (see animal protein) but no transformation of data was examined. The variance ratios observed tend to be slightly higher in adults than was seen in children. That for animal protein was much higher. In adult women, the variance ratio for fat was again lower than that for the other dietary variables. However, this was not the case for adult men.

Table 34. Estimates of the Partitioning of Variance in Kenya Adult Intakes

Nutrient Vector	Mean	SDinter	SDintra	Vratio
		1. FEMALES *		
Energy, kcal/d	1665	300.8 (18.0)**	762.7 (45.8)	6.43
Total Protein, g/d	49.0	9.78 (20.0)	25.83 (52.7)	6.99
Animal Protein, g/d	4.03	2.72 (67.5)	8.21 (203.7)	9.11
Fat, g/d	25.3	7.20 (28.5)	15.49 (61.2)	4.62
		2. MALES *		
Energy, kcal/d	1975	379.9 (19.2)	911.2 (46.1)	5.75
Total Protein, g/d	58.7	11.36 (19.4)	31.02 (52.8)	7.45
Animal Protein, g/d	5.05	3.09 (61.2)	11.21 (191.6)	13.18
Fat, g/d	30.6	7.55 (24.7)	18.59 (60.8)	6.07

Household sample only:
* Males Subjects = 214; Mean observations/person = 24.6
* Females Subjects = 214; Mean observations/person = 24.6
** Coefficient of Variation = 100 x SD/Mean

iv. Interrelationship of variances - mothers and toddlers.

Very high within-person variance is characteristic of all the target groups. It has been examined in detail for toddlers but remains largely unexplained. Although it was suggested earlier that interviewer effects may be present in the data set, these would not explain the very high variance between adjacent days when the same interview team would have been involved. As will be shown later, the variance of household intake across time is of a similar relative magnitude to that seen above. It seems unlikely that the high within-person variance can be attributed to simple errors in allocation of food consumed in the household. Unfortunately, there is no way to determine how much of this variance is real variation in intake and how much might be attributable to unidentified random error sources in the methodology for estimation of food intake.

As a further approach to examining the nature of the variance, mothers and toddlers examined in the same households on the same days were compared to see if deviations moved in phase (both were low or both were high on the same day). Toddlers included in the longitudinal data sets and lead females in the same household were selected. From these records, those days on which both the toddler and the lead female had a recorded intake were selected. For each subject linear regressions of intake on date were run and the residuals were computed for each observation. This was performed for energy intakes and for fat intakes. Residuals were examined as kcal/d and g/d respectively or as relative measures (% of the predicted value for that day). Pearson Product Moment correlations were run for maternal and toddler data.

For energy intake, the correlation between the residuals for maternal and toddler intakes was 0.28 and that for the relative deviation was 0.27. In a linear regression model predicting the toddler residual from the maternal residual, the Rsquare was 0.079 (as expected from correlation). Controlling for household did not change the Rsquare. Expressing the residuals as a proportion of predicted values lowered the Rsquares marginally (0.075, 0.075).

Similar results were obtained when fat intake was examined. Correlation coefficients were 0.29 and 0.28 and Rsquares without inclusion of household were 0.085 and 0.080 and with household included in the model they were 0.085 and 0.079.

In these same analyses, the Pearson correlation between predicted intakes of the mother and toddler were 0.20 for energy and 0.39 for fat. Using actual recorded intakes, the correlations were slightly higher, 0.26 for energy and 0.32 for fat.

These analyses suggest that the day to day variations in intake of mother and toddler are essentially independent (about 8% of the variance can be explained). That is surprising since it would seem to refute two interesting hypotheses: that the variation is attributable to widely fluctuating household food availability or that there is a systematic methodological error operating for all persons in the household. It remains possible that, in the presence of a shortage of food, mothers decrease intake to permit children to eat. If this were the case, it might be expected that there would be at least a

correlation in the signs of the deviations in intake even though there was not correlation in the magnitude of the deviations. To test this hypothesis, one further analysis was run. Residuals were categorized by sign of the deviation (0 = negative value; 1 = 0 or positive value), a two by two contingency table was built and Chi Square analysis applied. Results are shown in Table 35. There is an association between the directions of deviation of maternal and toddler intakes on a day to day basis. Both tend to move in the same direction. Although statistically highly significant with the large n involved, the effect is not large; note that the direction of deviation is more often negative than positive (although the sum of deviations expressed in actual magnitudes approximates 0). The analysis sheds little light on the reason for the very large within-person variances seen in the Kenya data sets.

Table 35. Contingency Table Examination of Direction of Same Day Deviations in Maternal and Toddler Intakes - Kenya

		Direction of Deviation in Maternal Intake					
		ENERGY INTAKES			FAT INTAKES		
		Negative	0 or Positive	Total	Negative	0 or Positive	Total
Direction of Deviation in Toddler Intake	Negative	846	503	1349	880	491	1371
	0 or Positive	471	567	1058	435	601	1036
	Total	1317	1090	2407	1315	1092	2407
		Chi Square = 79.2 P = 0.000			Chi Square = 117.3 P = 0.000		

b. Partitioning of variance in the Mexico project. (Corrected food composition data)

i. Toddler intake data.

As was done for the Kenya project, data were selected from subjects included in the longitudinal data set (as defined by ME). In the case of the Mexico project, recognizing that the method change had a marked effect upon mean intakes, only data collected after December 31, 1984 (i.e. with the final methodology) were examined. To be included in the data set, a particular subject was required to contribute at least 8 observations within the time defined by the target window (18-30 months). This resulted in the creation of a data set comprising 947 observations from 50 toddlers with an average of 18.94 observations each. The results of the ANOVA partitioning of variance are shown in Table 36.

Table 36. Estimates of the Partitioning of Variance in Mexico Toddler Intakes (Final Methodology)

Nutrient Vector	Mean	SDinter	SDintra	Vratio
Energy, kcal/d	1097	271.1 (24.7)*	355.2 (32.4)	1.72
Animal Energy, kcal/d	137.2	85.16 (62.1)	113.1 (82.4)	1.77
Total Protein, g/d	33.1	7.36 (22.2)	12.39 (37.4)	2.83
Animal Protein, g/d	9.69	5.05 (52.1)	8.24 (85.0)	2.67
Vegetable Protein, g/d	23.4	7.01 (30.0)	10.07 (43.0)	2.06
Fat, g/d	23.6	7.99 (33.9)	11.99 (50.8)	2.25
Iron, mg/d	11.3	3.07 (27.2)	4.98 (44.1)	2.63

Longitudinal sample only: Subjects = 50; Mean observations/person = 18.94
* Coefficient of Variation = 100 x SD/Mean

The contrast of the variance ratios with those seen in the Kenya data sets is striking (compare with Table 29). It is noteworthy that the variance ratios are not particularly different between total energy (or total protein) and animal source energy (or animal source protein). The fact that Mexico project analyses consistently demonstrated stronger relationships between these variables and outcome measures is not explicable as an artefact attributable to different relative error terms in the variables (as was implied might be a contributing factor in the fat intakes in Kenya). Conversely, what is striking is the sharp increase in the between subject variability in the animal source vectors (see CV's in above Table). This would suggest that whatever use of animal source foods may mark (intake of one or more specific nutrients or characteristics of the household and associated factors), the relative range among individuals is appreciably greater for the animal source nutrients. This is what would appear to be driving the Mexico analyses. It will be important to explore this phenomenon further to determine what other variables associate with use of animal source foods. In such further analyses, or in any interpretations of the present analyses, note should be taken of the major skew that exists within the animal source vectors (the CV's approach 100%; clear evidence of a skew). The distributions were not normalized for these analyses.

It is to be noted that the relative variabilities and the variance ratios seen for total energy, total protein, total fat and iron in toddlers (and in adult females below) are in keeping with the general expectations of the original design and with variabilities reported in industrialized countries. Comparative information for animal source energy and protein is not available.

ii. Lead female intakes and comparison of methodologies

As for toddlers, data for lead females were selected from the longitudinal sample and the longitudinal time windows as defined by ME. In the case of adult women, two samples were drawn. One data set comprised only data collected after December 31, 1984 (final method) while the other was made up of data collected prior to January 1, 1985 (a mix of earlier methods). An outlier screen was applied and intake records with more than 5000 kcal/d were

rejected from the analyses. It was required that subjects included contributed at least 8 intake records (the actual ranges were 8 to 32 for the final methodology and 8 to 24 for the earlier methodologies). The two data sets were subjected to ANOVA procedures and the variance components were estimated. The results are presented in Tables 37 and 38. No attempt was made to separate pregnant or lactating women from their nonpregnant, nonlactating compatriots nor was any attempt made to match individuals in the two data sets.

Table 37. Estimates of the Partitioning of Variance in Mexico Adult Female Intakes (Final Methodology)

Nutrient Vector	Mean	SDinter	SDintra	Vratio
Energy, kcal/d	2524	520.0 (20.6)*	629.1 (24.9)	1.46
Animal Energy, kcal/d	161.2	117.5 (72.9)	155.9 (96.7)	1.76
Total Protein, g/d	71.5	14.52 (20.3)	21.19 (29.6)	2.13
Animal Protein, g/d	12.7	7.80 (61.4)	12.67 (99.8)	2.64
Vegetable Protein, g/d	58.8	13.53 (23.0)	17.90 (30.4)	1.75
Fat, g/d	39.0	13.15 (33.7)	18.29 (46.9)	1.94
Iron, mg/d	28.9	6.77 (23.4)	9.03 (31.2)	1.77

Longitudinal sample only: Subjects = 129; Mean observations/person = 18.5
Not separated by physiologic state
* Coefficient of Variation = 100 x SD/Mean

On a comparative basis, the analyses in Table 37 parallel those in Table 36 suggesting that stability of eating patterns is similar in mothers and toddlers. There is somewhat less variability in energy intake among women than was seen in toddlers (but note that toddler analyses were not controlled for age and neither were controlled for weight; these might contribute some variance). Again the variation between subjects and within subjects increased sharply with animal source foods.

Table 38. Estimates of the Partitioning of Variance in Mexico Adult Female Intakes (Early Methodologies)

Nutrient Vector	Mean	SDinter	SDintra	Vratio
Energy, kcal/d	2048	477.5 (23.3)*	647.6 (31.6)	1.84
Animal Energy, kcal/d	87.5	71.05 (82.1)	121.8 (139.2)	2.94
Total Protein, g/d	55.6	12.94 (23.2)	20.61 (37.1)	2.54
Animal Protein, g/d	6.79	4.99 (73.5)	10.82 (159.4)	4.71
Vegetable Protein, g/d	48.8	12.46 (25.5)	16.78 (34.4)	1.81
Fat, g/d	22.9	7.51 (32.8)	13.27 (57.9)	3.12
Iron, mg/d	23.6	6.24 (26.4)	8.44 (35.8)	1.83

Longitudinal sample only: Subjects = 110; Mean observations/person = 15.9
Not separated by physiologic state
* Coefficient of Variation = 100 x SD/Mean

Method comparisons were not undertaken for toddler data because of the likelihood that systematic age differences covarying with method would confound interpretation. For the adult women, comparison of analyses in Tables 37 and 38 present some sharp contrasts. Some care in interpretation is necessary since the "old methods" actually comprises data collected with two methodologies which have been demonstrated to yield different estimates of intake (see section 7). This method effect would impact on variance in the "old methods" data set. This could be explored by repeating the present analyses but using only data collected with the first methodology; the second method was used for too short a time to warrant ANOVA approaches. As has been noted before (see section 7), the comparison of Tables 37 and 38 suggests that there is a major bias associated with choice of method. It probably affects all vectors, as might be expected, but it does not appear to operate equally for all nutrients and their variances. In section 7, it appeared that the bias operated differentially for animal source foods (see nutrient:energy ratios in that section). Those analyses were performed on intakes calculated before the food composition tables were corrected to better estimate animal sources of energy and protein. The present analyses, based on corrected food composition data, extend those observations. The comparison of Tables 37 and 38 clearly suggest that the change in methods had greater impact on the variabilities of animal source energy and protein than on any other vector except fat intake. Examination of the variance ratios suggests that for animal source nutrients and for fat, attenuation of relationships would be much greater for the old methods than the new methodology. That is, it is at least possible that some of the apparently strong relationships seen between animal energy or animal protein and outcome variables are method-specific. Given that it is uncertain which of the three methodologies used is actually the most precise, and given the suggestion that the methods differentially affect estimation of dietary components, it will be important to explore in greater detail the exact impact of the methodology changes on the estimation of specific classes of food and to then explore the relationship of these classes of foods to the households in which they are reported. It is clear also that it will not be a simple matter to develop a statistical adjustment which would permit use of data collected by both dietary methodologies in the same analyses -- at least not unless and until it is better known what food component(s) of the diet are most affected by the methodology.

c. Partitioning of variance in Egypt project.

i. Toddler intake data.

The Egypt project provided 2408 intake records for toddlers in the age range 18 to 30 months. To these records the following exclusion rules were applied: must have been weaned, intakes could not be in excess of 4500 kcal and after those exclusions, at least 8 records must be available. The resultant data set provided a total of 1891 records from 118 toddlers contributing 8 to 27 records (mean = 16.0). ANOVA was performed as for the other projects. Estimates of the components of variance are presented in Table 39. It is to be noted that for these intake records, energy had been calculated as (g protein x 4) + (g fat x 9) + (g carbohydrate x 4) pending revision of the food composition data base. The records analyzed were those used by the Egypt project in preparation of their final report.

Table 39. Estimates of the Partitioning of Variance in Egypt Toddler Intakes (4-9-4 Calculations)

Nutrient Vector	Mean	SDinter	SDintra	Vratio
AGE 18-30 MONTHS				
Energy, kcal/d	1207	289.9 (22.5)*	450.1 (37.3)	2.41
Total Protein, g/d	39.6	9.85 (24.9)	19.67 (49.7)	3.99
Animal Protein, g/d	18.6	4.56 (24.5)	16.16 (86.9)	12.53
Vegetable Protein, g/d	21.0	6.13 (29.0)	9.71 (46.2)	2.50
Fat, g/d	33.4	8.70 (26.0)	18.44 (55.2)	4.49
AGE 24-30 MONTHS				
Energy, kcal/d	1245	282.7 (22.7)	429.3 (34.5)	2.31
Total Protein, g/d	40.6	9.71 (23.9)	19.24 (47.4)	3.93
Animal Protein, g/d	18.2	3.61 (19.8)	15.84 (87.0)	19.25
Vegetable Protein, g/d	22.4	6.60 (29.5)	9.77 (43.6)	2.19
Fat, g/d	33.0	7.47 (22.1)	16.20 (49.3)	4.70

* Coefficient of Variation = 100 x SD/Mean

The between subject variances, expressed as CV's, are comparable to those seen in Mexican toddlers and close to design expectations. This is moderately surprising since it is known that the data set includes an observer effect that could be expected to increase the SDinter to some degree. The within-person variances are again comparable in range to those seen in Mexico and much lower than those seen in Kenya. However, the ratio of variances is generally higher for Egypt than for Mexico and that for animal protein is particularly surprising. To test the possibility that this might be a function of inclusion of milk or formula during the weaning period, the analyses were repeated using only data collected between 24 and 30 months for weaned toddlers. This might be presumed to reduce the proportion receiving immediate post-weaning diets. The results are shown in Table 39. This analysis drew on data from 88 subjects who contributed an average of 10.3 observations each (range 8 to 14). The picture was very similar to that seen in the whole group.

ii. Maternal intake data.

The partitioning of variance in adult females was examined first in the total group (with requirements that at least 8 observations be contributed and that the reported energy intake not be greater than 5000 kcal). From an initial pool of 5068 observations, 21 were rejected on the basis of outlying energy values and 68 were rejected as belonging to sets with fewer than 8 observations. The analyzed data set comprised data from 246 adult women contributing an average of 19.8 observations (range 8 to 41). A subset of data drawn from women known to be in the second or third trimester of pregnancy or in the first six months of lactation (birthdate - 6 months to birthdate + 6 months). This data set was drawn from 97 women who contributed an average of 16.9 observations (range 8 to 26). Both sets of analyses are presented in Table 40.

Table 40. Estimates of the Partitioning of Variance in Egypt Adult Female Intakes (4-9-4 Calculations)

Nutrient Vector	Mean	SDinter	SDintra	Vratio
ALL WOMEN				
Energy, kcal/d	2268	302.4 (13.3)*	672.4 (29.6)	4.94
Total Protein, g/d	78.0	10.67 (13.6)	33.78 (43.3)	10.02
Animal Protein, g/d	32.1	6.47 (20.2)	30.53 (95.1)	22.24
Vegetable Protein, g/d	45.9	7.21 (15.7)	15.61 (34.0)	4.68
Fat, g/d	57.1	9.45 (16.5)	29.25 (51.2)	9.58
PREGNANT OR LACTATING WOMEN				
Energy, kcal/d	2212	313.9 (14.2)	680.8 (30.8)	4.73
Total Protein, g/d	76.2	11.98 (15.7)	34.52 (45.3)	8.30
Animal Protein, g/d	32.2	7.81 (24.3)	31.32 (97.3)	16.08
Vegetable Protein, g/d	44.0	6.60 (15.0)	15.92 (36.2)	5.81
Fat, g/d	57.7	11.34 (19.7)	29.90 (51.8)	6.94

* Coefficient of Variation = 100 x SD/Mean

Adult women exhibit a lower between-person variation, relative to the mean, for most nutrient vectors than do the toddlers. The relative within-person variabilities also tend to be reduced (except for animal protein), but the variance ratios tend to be higher for adult women than for toddlers. These subjects depart quite appreciably from the original design assumptions, making it more difficult than anticipated to detect associations (the departure is not as severe as was seen in Kenya but more so than for Mexico). It is of interest to note that the estimated vegetable protein intake has the highest between-person CV and a variance ratio that approximates that for energy. The project detected low order associations between vegetable protein intake and outcome measures in pregnancy/lactation while no associations were detected for either total protein or animal protein. This is not surprising given the variance ratios portrayed above.

d. Implications for interpretation of interproject comparisons. The import of the findings on partitioning of variance for statistical analyses is illustrated, for selected examples, in Table 41, drawing upon the data of Kenya as reported in Table 29. Here the predicted attenuation of bivariate correlation and regression coefficients, based on the observed variance ratios, is displayed. Details of the calculation of these attenuation factors will be found in the LSHO report (11). Application of the factors may be illustrated by an example. If the true relationship between energy intake and a body weight parameter were described by a correlation coefficient of 0.4, Table 41 suggests that with 2 days of data (1 month's observations), the observed correlation might be about 0.2. With 3 months of data (6 observations) it might rise to about 0.28, and if a six month mean were used it might be 0.32. If the original regression slope were 0.2, the observed slopes might be only 0.05, 0.10 and 0.13. Thus, both the estimate of the relationship and the power to detect it as statistically significant is sharply reduced by these high variance ratios. It implies also that the ability to detect relationships across short periods of time is seriously

blunted. Some of the analyses presented by the Kenya project involve mean intakes over six month periods. Relationships were not seen when shorter intervals were used.

Table 41. Examples of the Predicted Attenuation of Bivariate Correlation and Regression Analyses as a Function of Number of Observations Pooled in the Estimate of Intake, Kenya Project Data

Nutrient Vector	Correlation Coefficient			Regression Slope		
	2 days	6 days	12 days	2 days	6 days	12 days
Energy /d	0.49	0.70	0.81	0.24	0.49	0.66
Protein /d	0.44	0.65	0.77	0.20	0.42	0.60
Anim Protein /d	0.48	0.69	0.80	0.23	0.48	0.64
Fat /d	0.59	0.79	0.87	0.35	0.62	0.77
Carbohydrate /d	0.46	0.67	0.79	0.21	0.45	0.62
Iron /d	0.42	0.63	0.75	0.18	0.40	0.57

A very serious implication of these observations is the need to guard against the false negative interpretation of analyses. That is, if a relationship is not seen, it may be because it does not exist. It may also be because the attenuation is so great that it is not detected in conventional analyses.

A second problem is apparent from the analyses presented in the foregoing pages. The variance ratios, and hence degrees of attenuation of relationships, vary between nutrient vectors within projects and within nutrient vectors between projects. The implication is clear. If, within a project, the real biological relationships between several nutrient vectors and an outcome variable were comparable, the observed relationships would differ because of different degrees of attenuation. Analogously, between projects, different nutrient vectors may appear as significant predictors of outcomes as a result of differences in variance ratios as well as because of true differences in the biological relationships. This may be a particular problem when simple correlation analyses are used in the selection of variables to be entered into regression models; the correlations would be attenuated as indicated above and this would influence the selection decision. Once entered into multivariate regressions, it is difficult to predict the effect of the variance ratios (11). The observation of different nutrient-function relationships across projects, or the presence of relationships in analyses for one project and absence in analyses for another must be interpreted with the variance ratios in mind.

Examining the variance ratios alone begs the question of why the variance ratios differ. Although the foregoing discussions have emphasized the variance ratios and the CV's in relation to original design targets, examination of the absolute variance components (expressed in the tables as SDinter and SDintra) is of interest. The variance ratio involves both components. A question of interest is which differs between projects and why

does it differ. Intraindividual variances would differ if there are real differences in the stability of eating patterns or if there were a random methodological error in recording intake that differed between projects. Interindividual variances would differ if indeed there were differences in the "homogeneity" of the populations studied or if there were systematic biases operating across individuals within one or more of the projects. Table 42 presents a comparison of some variance components across projects. It would appear that, for toddler intakes, between-subject variance differs between projects more than does the within-person variance (fat intake in Kenya does not fit this pattern). That is, the differences in variance ratios described above are less likely to reflect differing methodological errors across the projects than differences in the range of subjects themselves. Specifically, the Kenya subjects would appear to be much more homogeneous than either the Egypt or Mexico subjects. (See earlier discussion of the planned sampling design which was intended to ensure an adequate variance between subjects.) This conclusion must be tempered by the potential effect of interviewer biases on the between-subject variance (see discussion of Egypt project above). Nevertheless, the results do suggest again that the power to detect relationships between energy intake and outcome variables differs between projects. The differences between projects in the SDinter and SDintra of the nutrient vectors presented in Table 42 for adult females are not as consistent as would seem to the case for toddlers. One cannot suggest as easily that it is one component or the other that differs between projects.

Table 42. Comparison of Variance Components for Selected Dietary Measures Across Projects

Project	Toddler Intake Data			Maternal Intake Data*		
	Energy kcal/d	Protein g/d	Fat g/d	Energy kcal/d	Protein g/d	Fat g/d
	SDintra (Within-person variance)					
Egypt	450	19.7	18.4	672	33.8	29.3
Kenya	410	13.6	9.7	763	25.8	15.5
Mexico	355	12.4	12.0	629	21.2	18.3
	SDinter (Between-person variance)					
Egypt	290	9.9	8.7	302	10.7	9.5
Kenya	165	4.8	5.1	301	9.8	7.2
Mexico	271	7.4	8.0	520	7.8	13.2

* Not separated by physiologic state

These similarities and differences between projects would be worthy of further exploration both as a basis for further understanding of the data and as an approach to "external validation" of the dietary methodologies. It would be interesting, for example, to explore correlations between maternal intake and toddler intake in the Egypt project as was done for Kenya data. Under the Egypt methodology, mothers estimated toddler intakes as a proportion of their own. Any association of errors in estimation might be detectable by such exploration.

Finally, as was noted above, examination of the standard deviations for a number of the nutrients suggests that the primary distributions carry major skewing. This is very apparent, for example, in the Mexico toddler animal source energy and protein but is seen also in a number of other nutrient estimates across projects. In the preliminary analyses reported by projects, it does not appear that specific consideration has been given to the potential impact of such skewing of the dietary variables (e.g. by normalizing transformations or by use of nonparametric approaches). This might be worthy of exploration in later analyses.

10. Household Food Intake

Project reports do not include comment on household intake. Using the Kenya project data set, ME staff developed an approach for deriving and expressing household level intake data. This is described below. At this stage, the derived variable has not been used in analyses.

After exploring a number of possible approaches to expression of the household intake (developing an appropriate denominator to take into account the differences in demography across households) it was decided to express intake of the household as the ratio to estimated household BMR. This was predicted from FAO/WHO/UNU equations (9). The term "household caloric intake ratio" has been applied to this variable.

The household caloric intake ratio for a given day is defined as the total household intake of calories on that day divided by the household basal metabolic rate. Household intakes of calories and other nutrients (proteins, fats, carbohydrates, and iron) were calculated summing total meal nutrients consumed within the household plus snack nutrients consumed by individuals plus nutrients from other foods consumed outside the house (or with origin outside the house, as in the drought foods) minus nutrients from foods consumed by guests. Since household meals represented the largest portion of nutrient intakes, every effort was made to approximate dish nutrients when recipes or ingredient weights were unavailable. Standard recipes were used to assign dish nutrients in these cases (which numbered 307 out of 37,153 dish observations). In only four of these cases were we unable to assign nutrient values to a consumed dish.

Household basal metabolic rate (BMR) was defined as the sum of individual basal metabolic rates (expressed in kilocalories) for all individuals in the household, where individual age, sex, and weight data were used in the FAO/WHO/UNU BMR equations (9). For each household and for each observation day, total caloric intake was divided by the household BMR to arrive at the "household caloric intake ratio".

Since the household BMR variable is an aggregate of individuals' BMRs, it is sensitive to the household composition on any given observation day. Census data were used to determine the number of household members (as well as their age and sex) in the aggregate household BMR variable. Since the census was conducted only once every three or four months, an assumption was made about the composition of the household when no census data were available. In

particular, the composition of the household was assumed to be the same as it was on the most recently conducted census.

In a few cases, census data were discarded due to obvious coding errors. The most recent previous census was used to determine household composition in these cases. It should be noted that in many other cases, the "number of persons in household" variable (form 612, columns 60-61) did not match the actual number of persons for which data were collected on a particular census day. The discrepancy may have been due to the use of previous census forms to fill in this variable, or the failure of the interviewer to thoroughly probe the respondent as to all members in the household. In the former case, the "number of persons" variable would have been low by the number of extra members in the household due to births or visitors. In the latter case, the "number of persons" variable might have been accurate, but the actual count of records (one per person) would have been low by the number of persons missed by the interviewer. Except for obvious coding errors, household BMR calculations were made from the actual number of records reported on each census day.

Since there was a significant degree of missing data with respect to the weight variable (and to a lesser extent with birthdate and sex variables) protocols were developed so that every household member had estimated values for all variables to be used in the BMR equations. In this way, underestimation of household BMR due to missing data was kept to a minimum.

Mean values for age by person type were used as the default values for those individuals with missing age data. For example, the mean age of all 'lead females' who had age data was 32.2 years. Male 'non-relatives' averaged 20.7 years, while female 'non-relatives' averaged 19.7 years of age.

Mean values for weight by person type were used as the default values for those individuals with missing weight data and for whom the person type was well-defined with respect to sex and age. For example, the mean weight of all 'lead males' who had weight data was 55.06 kilograms. For person types that are overly aggregated with respect to sex or age, such as 'other relatives', mean values for weight for 12 distinct sex/age groups were used. For example, a 45 year old female 'other relative' with missing weight data, was assumed to weigh 51.7 kilograms.

In a certain number of cases, the sex of the individual was not reported. The defaults used were clear cut for ten of the person types (e.g.'lead males', 'lead females', etc.). For 'other relatives', sex was assigned randomly to individuals with missing data, in the proportion in which they existed in the 'other relatives' population. For example, approximately two-thirds of the 'other relatives' were female and one-third was male. The same procedure was followed in the 'non-relative' category, where the distribution was approximately half male and half female.

For a few cases the person type of the individual was missing. In these cases, the default values used were either 'non-target female offspring' or 'non-target male offspring'. These values were assigned randomly with equal probability. Non-target offspring was chosen as the default person type

since, with over 50 percent of the population, it is the largest group. This choice also has the least likelihood of forcing the data, since all age and sex groups are represented.

Table 43 shows the caloric intake ratios, total caloric intakes, and household BMRs for several households for their entire time in the study. The table is presented here to demonstrate a small part of the final data set. Summary statistics for the caloric intake ratio variable for several households with at least five observations are presented in Table 44. The mean caloric intake ratio by household varied from 0.65 to 2.75. In 14 of the 247 households, the ratio was below 1.0 and in only five of the households was it above 2.0. The number of observation days per household ranged from 8 to 44.

Table 45 contains information concerning the distribution of excessively low intake ratios (less than 0.5) across time. Relatively larger percentages of the very low caloric intake ratios occur in the months corresponding to the drought in Kenya. Further work should explore those households with unusually low caloric intake ratios (apparent intakes below the aggregate BMR of household members).

Application of an analysis of variance procedure to these data permitted an estimation of the partitioning of variance. There were a total of 8,194 observations (1 per household per day) in this data set. Observations with missing dish nutrients or from households with less than 5 total observations were not included. The mean caloric intake ratio for the whole sample was 1.40. The within-household coefficient of variation was 38.3% and the between-household coefficient of variation was 17.1%. This very high within-household or day-to-day variation is very similar to that described for individual intakes in the Kenya data set (see preceding section).

Table 43. Energy Intake Ratios for Several Kenya Households

CURRHH	MONTH	DAY	YEAR	KCLRATIO	THHKCAL	HHNMR	NRECORDS	MISSDISH
1000	2	11	1984	1.11	10288	9279	8	0
1000	2	12	1984	1.02	9504	9279	8	0
1000	3	1	1984	1.13	10487	9279	8	0
1000	3	2	1984	1.31	12112	9279	8	0
1000	4	17	1984	1.03	9564	9279	8	0
1000	4	18	1984	0.99	9221	9279	8	0
1000	5	10	1984	1.83	16964	9292	8	0
1000	5	11	1984	0.94	8757	9292	8	0
1000	6	11	1984	1.02	9509	9292	8	0
1000	6	12	1984	1.23	11466	9292	8	0
1000	7	5	1984	1.19	11074	9292	8	0
1000	7	6	1984	0.94	8743	9292	8	0
1000	8	6	1984	1.43	13305	9292	8	0
1000	8	7	1984	1.32	12241	9292	8	0
1000	9	3	1984	0.70	6518	9292	8	0
1000	9	4	1984	1.31	12135	9292	8	0
1000	10	11	1984	0.86	7963	9292	8	0
1000	10	12	1984	1.21	11263	9292	8	0
1000	11	5	1984	1.25	11878	9504	9	0
1000	11	6	1984	0.97	9191	9504	9	0
1000	12	13	1984	1.83	17391	9504	9	0
1000	12	14	1984	1.95	18501	9504	9	0
1000	1	24	1985	1.03	9822	9504	9	0
1000	1	25	1985	1.71	16299	9504	9	0
1000	2	11	1985	0.75	7098	9504	9	0
1000	2	12	1985	0.80	7636	9504	9	0
1000	3	30	1985	0.94	8972	9504	9	0
1000	3	31	1985	0.81	7676	9504	9	0
1002	2	11	1984	2.02	20486	10101	9	0
1002	2	12	1984	1.14	11561	10101	9	0
1003	2	11	1984	1.37	10285	7500	6	0
1003	2	12	1984	0.66	4977	7500	6	0
1003	3	1	1984	0.97	7300	7500	6	0
1003	3	2	1984	0.33	2473	7500	6	0
1003	4	17	1984	1.53	11471	7500	6	0
1003	4	18	1984	0.85	6345	7500	6	0
1003	5	10	1984	0.84	6284	7500	6	0
1003	5	11	1984	1.05	7900	7500	6	0
1003	6	11	1984	1.65	12390	7500	6	0
1003	6	12	1984	1.50	11247	7500	6	0
1003	7	5	1984	1.33	9973	7500	6	0
1003	7	6	1984	1.25	9349	7500	6	0
1003	8	6	1984	0.97	7281	7500	6	0
1003	8	7	1984	1.45	10876	7500	6	0
1003	9	3	1984	0.17	1271	7500	6	0
1003	9	4	1984	0.28	2089	7500	6	0
1003	10	11	1984	0.14	1015	7500	6	0
1003	10	12	1984	1.01	7554	7500	6	0
1003	11	5	1984	0.70	5284	7500	6	0

Table 43 (Cont.) - Energy Intake Ratios for Several Kenya Households

CURRHH	MONTH	DAY	YEAR	KCLRATIO	THHKCAL	HHBMR	NRECORDS	MISSDISH
1003	11	6	1984	0.98	7336	7500	6	0
1003	12	17	1984	0.71	5316	7500	6	0
1003	12	18	1984	0.69	5174	7500	6	0
1003	1	24	1985	1.03	7752	7500	6	0
1003	1	25	1985	0.79	5924	7500	6	0
1005	1	26	1984	2.68	31190	11631	10	0
1005	1	27	1984	0.82	9564	11631	10	0
1005	3	1	1984	1.21	14117	11631	10	0
1005	3	2	1984	0.90	10499	11631	10	0
1005	4	17	1984	1.42	16565	11631	10	0
1005	4	18	1984	1.09	12707	11631	10	0
1005	5	21	1984	0.99	11467	11631	10	0
1005	5	22	1984	0.87	10075	11631	10	0
1005	6	21	1984	1.13	13165	11631	10	0
1005	6	22	1984	0.68	7961	11631	10	0
1005	7	12	1984	0.74	8574	11631	10	0
1005	7	13	1984	0.49	5672	11631	10	0
1005	8	25	1984	1.30	15135	11631	10	0
1005	8	26	1984	1.12	13081	11631	10	0
1005	9	25	1984	0.87	10135	11631	10	0
1005	9	26	1984	0.83	9691	11631	10	0
1005	10	14	1984	1.14	13270	11631	10	0
1005	10	15	1984	0.96	11150	11631	10	0
1005	11	5	1984	0.66	8649	13088	11	0
1005	11	6	1984	0.71	9301	13088	11	0
1005	12	17	1984	1.17	15378	13088	11	0
1005	12	18	1984	0.93	12215	13088	11	0
1005	1	24	1985	1.31	17207	13088	11	0
1005	1	25	1985	0.69	9068	13088	11	0

KCLRATIO is the ratio of total kcal intake/household BMR.
THHKCAL is the total caloric intake for that day.
HHBMR is the aggregate household BMR.
NRECORDS is the number of persons used in calculating HHBMR.
MISSDISH is the number of dish records thrown out due to missing data on the weight consumed or recipe ingredients, etc.

Table 44. Energy Intake Ratio, Summary Statistics by Household
Kenya Project

OBS	CURRHH	ERN	ERMEAN	ERSTDEV	ERCV	ERMIN	ERMAX	INTMEAN	PERSMEA
1	1000	28	1.17	0.33	28.52	0.70	1.95	10914	8
2	1003	24	0.93	0.43	46.24	0.14	1.65	6953	6
3	1005	24	1.03	0.42	41.18	0.49	2.68	12326	10
4	1006	40	1.03	0.34	32.68	0.24	1.79	11382	10
5	1007	44	1.57	0.74	46.99	0.54	5.09	5808	3
6	1009	43	1.45	0.53	36.26	0.03	2.47	11759	8
7	1010	33	1.68	0.77	45.86	0.70	5.08	9768	5
8	1011	8	2.42	0.54	22.49	1.67	3.15	8068	3
9	1013	26	1.39	0.61	43.93	0.60	3.18	11114	7
10	1014	38	1.13	0.32	28.37	0.62	1.96	14045	11
11	1015	42	1.09	0.38	34.71	0.47	2.03	16530	13
12	1016	42	1.82	0.55	30.31	0.76	3.33	10999	6
13	1017	24	1.64	1.55	94.80	0.35	8.67	7059	4
14	1018	40	1.46	0.53	36.52	0.58	2.71	10565	7
15	1019	44	1.51	0.50	33.40	0.31	2.60	9160	7
16	1020	24	0.89	0.47	52.48	0.21	1.71	9903	10
17	1021	18	1.32	0.33	25.03	0.73	1.81	5495	4
18	1022	44	1.53	0.69	45.35	0.49	3.48	12117	8
19	1023	44	1.19	0.50	41.92	0.02	2.34	10850	10
20	1024	42	1.63	0.51	31.10	0.71	2.99	10480	6
21	1025	22	1.76	0.57	32.26	0.65	3.02	15672	8
22	1026	40	1.63	0.45	27.52	0.62	2.57	11955	7
23	1027	42	1.64	0.51	30.99	0.89	2.76	8498	6
24	1028	17	1.16	0.35	30.28	0.63	2.11	10878	9
25	1029	24	1.29	0.44	34.09	0.58	2.74	15056	11

Observation days with missing dish nutrients were deleted from this table as were households with less than five observation days.

CURRHH is the household identification number.
ERN is the number of caloric intake ratio (energy ratio) observations per household.
ERMEAN is the mean energy ratio per household.
ERSTDEV is the standard deviation of the energy ratio variable.
ERCV is the coefficient of variation of the energy ratio.
ERMIN is the minimum value of the energy ratio per household.
ERMAX is the maximum value of the energy ratio per household.
INTMEAN is the mean kcal intake per household
PERSMEAN is the mean number of persons per household.

Table 45. Distribution of Low Energy Intake Ratios* Over Time
Kenya Project

YEAR	MONTH	FREQUENCY	PERCENT	CUMULATIVE FREQUENCY	CUMULATIVE PERCENT
84	01	2	0.7	2	0.7
84	02	5	1.8	7	2.6
84	03	7	2.6	14	5.1
84	04	7	2.6	21	7.7
84	05	17	6.2	38	13.9
84	06	12	4.4	50	18.3
84	07	13	4.8	63	23.1
84	08	26	9.5	89	32.6
84	09	31	11.4	120	44.0
84	10	29	10.6	149	54.6
84	11	30	11.0	179	65.6
84	12	13	4.8	192	70.3
85	01	16	5.9	208	76.2
85	02	11	4.0	219	80.2
85	03	6	2.2	225	82.4
85	04	4	1.5	229	83.9
85	05	10	3.7	239	87.5
85	06	5	1.8	244	89.4
85	07	7	2.6	251	91.9
85	08	7	2.6	258	94.5
85	09	4	1.5	262	96.0
85	10	4	1.5	266	97.4
85	11	6	2.2	272	99.6
85	12	1	0.4	273	100.0

* Low calorie intake ratios taken as 0.5 or below

The drought was during the spring and summer of 1984. Food shortages were reported to have become acute by September 1984 and a food distribution system operated from September to December, 1984.

E. Resting Metabolic Rate

1. Introduction

The purpose of what follows is to provide an overview of the RMR data collected in the three country projects. This overview includes both descriptive statistics and comparisons with BMR as predicted by the FAO/WHO/UNU prediction equations (9). The report explores the variations in RMR across time and, in the Kenya data set, the concurrent changes in food intakes and RMR (without presenting analyses to examine direction or strength of any relationships observed). Here implications for Nutrition CRSP hypotheses are considered. Finally, the report identifies remaining problems to be addressed before definitive analyses of these data sets are undertaken.

In these analyses, data for only two projects, Kenya and Mexico, are presented. Although RMR data were collected in Egypt and might have been subjected to similar preliminary analyses, at the time of preparation of this report two methodologic problems that were identified had not yet been corrected. These are described in section 6b. below together with suggested correction procedures. The Kenya data set also had included errors attributable to the effect of humidity. As used here, that data set was adjusted to remove the effect of that error source. Section 6c. discusses the possible presence of a second error source in the Kenya data set. In both the Kenya and Mexico data sets, exclusion rules have been applied to remove outlying values of variables involved in the computation of RMR. Details of these exclusions are provided in section 6a.; all tabulations include notation as to whether or not exclusion rules were applied in the particular analysis reported.

Analyses are presented for adult male subjects, for lead females (not classified by physiologic state) and for schoolers. In examining data sets and in comparing observed and predicted values, only lead males were considered. This approach was adopted since the lead male data sets are the most easily interpretable (no consideration of potential changes attributable to growth or pregnancy).

2. Variables and Measurement

The protocol required measurement of RMR every 3 months in lead males and lead females who were neither pregnant nor lactating, and at the 5th and 8th month of pregnancy and the 1st, 3rd, and 6th months of lactation. Measurement of RMR of schoolers was to be every 3 months if projects elected to include this age group. Measurements of toddlers and infants were determined technically infeasible under field conditions. The variables to be reported and archived were body temperature, weight and height, and data from test minutes 18, 19, 20, 28, 29, and 30 as follows: oxygen consumption, ml/min STP; respiratory quotient; heart rate; RMR, kcal/day; and RMR. kcal/kg.

All projects chose to purchase the Beckman Metabolic Measurement Cart for RMR determination. The MMC was used by the Kenya and Egypt projects; details of the difficulties encountered and correction procedures devised by Kenya

project staff with ME assistance will be found in the Kenya Final Report, chapter 12.

The Mexico project was unable to make their MMC operational; they substituted another apparatus, the portable Max Plank Respirometer, for volume measurement and used the Beckman OM-2 analyzer for analysis of oxygen. Carbon dioxide was not measured and RQ is assumed to be 0.85 for calculation of energy values.

The problems resulted in substantial loss of data as well as significant costs to projects and ME. Questions important to the scientific community - notably population variation in RMR and the effects of substandard energy intake - cannot now be answered with the confidence anticipated.

3. Summary Statistics and Comparison with FAO/WHO/UNU Predicted Values

Table 46 summarizes the distribution of replicate observations on the same individuals within the Kenya and Mexico data sets for adults after application of data exclusion rules.

Table 47 presents summary statistics for the observed RMR, the RMR as predicted by the FAO/WHO/UNU equations, and the pair-wise difference between observed and predicted values. In the paired analysis, the Kenya RMR values are significantly below those of the FAO/WHO/UNU equations while Mexico RMR's appear to be comparable.

The finding of a significant difference between observed and predicted RMR in Kenya, but not in Mexico, might be due to the methodologic difference (Kenya, Beckman MMC and Mexico, Max Plank). In the final stages of the Kenya project, a direct comparison of the Beckman MMC and Max Plank estimates was undertaken through paired estimations in 13 subjects. (This comparison has been reported in detail by the Kenya project. See their Final Report, chapter 12.) The mean difference between the two estimates was about 200 kcal/d with the Max Plank being higher. If one examines the data in Table 47, the difference in the observed RMR - predicted RMR is about -200 kcal/d, a figure that is close to the apparent effect of method of estimation. Using data published by Duggan, Alwar and Milner (13) for 20 black Kenyan children, 14-79 months of age, studied after recovery from measles, RMR predicted by the FAO/WHO/UNU equations was 55.6 kcal/kg compared to measured RMR of 65.4 kcal/kg. These data again suggest that the lower than predicted RMR values seen in the CRSP adult males were an effect of methodologic rather than biologic or environmental differences. It should be anticipated that similarly "low" values will be seen in the Egypt population when those data become available for analyses since the same Beckman MMC equipment was used.

Table 46. Distribution of Replicated RMR Observations Over Entire Study Period

Number of Observations	Kenya Males	Mexico Males
1	51	66
2	45	31
3	61	34
4	46	19
5	24	18
6	5	7
7	1	4
Total Number of Subjects:	233	179
Total Number of Observations:	665	466

Table 47. Comparison of Observed RMR and Predicted BMR (Adult Males in Kenya and Mexico)

Variable	Kenya			Mexico		
	N	Mean	SD	N	Mean	SD
Observed RMR (kcal/d)	233	1309	161	179	1637	278
Predicted RMR (kcal/d)	234	1526	103	178	1647	119
Observed - Predicted RMR	232	-214	136	178	-8	257

* Statistically different from 0; P |T| = 0.0001

** Predicted BMR was not age adjusted in Kenya analyses; predictions based on FAO/WHO/UNU equations for young male adults as a function of body weight. All data pooled by subject.

The comparison of observed and predicted values for adult males was pursued through regression analyses. In the first analyses, observed RMR was regressed on body weight. In the second analyses, observed RMR was regressed on BMR predicted by the FAO/WHO/UNU regression equations (a function of weight). The results are summarized in Table 48. (In these analyses, Mexico data have not been subjected to outlier rejection procedures.)

There was a surprisingly low order of relationship between RMR and body weight in either of these data sets. For the Kenya data set, using the conventional function of weight to the power 0.75 or log weight had little impact on the Rsquare. By contrast with the present observations, the FAO/WHO/UNU regression equations were reported to carry correlation coefficients of .6 to .8 (implied Rsquares of .36 to .64). Given the very weak relationship to body weight, the lack of clear relationship to the FAO/WHO/UNU regression equations is to be expected since these are a function of weight and age (age not included in Kenya comparison). As might be expected, if single measures (rather than pooled data) are used in the regressions, the Rsquare values fall slightly.

Table 48. Regression of RMR on Weight and on Predicted BMR of Adult Males

Regression of Observed RMR on Body Weight

A. All observations treated as independent

KENYA: Observed RMR = 540.6 + 14.22 Weight

N = 678
Rsquare = 0.2325
SE intercept = 55.45
SE slope = 0.994

MEXICO: Observed RMR = 1035.1 + 8.995 Weight

N = 527
Rsquare = 0.0687
SE intercept = 95.98
SE slope = 1.444

B. Data pooled for each subject

KENYA: Observed RMR = 536.7 + 13.51 Weight

N = 232
Rsquare = 0.2929
SE intercept = 76.93
SE slope = 1.384

MEXICO: Observed RMR = 981.42 + 9.63 Weight

N = 178
Rsquare = 0.1116
SE intercept = 131.10
SE slope = 1.989

Regression of Observed RMR on FAO/WHO/UNU Predicted BMR

KENYA: Observed RMR = -90.44 + 0.929 Predicted BMR

N = 682
Rsquare = 0.2825
SE intercept = 99.37
SE slope = 0.0650

MEXICO: Observed RMR = 473.90 + 0.6999 Predicted BMR

N = 525
Rsquare = 0.0664
SE intercept = 189.339
SE slope = 0.115

The data fail to establish whether or not the international predictions are applicable to these populations. Although suggesting that there is no bias in the group mean estimate of RMR for Mexico (see Table 47), the regression analyses, and the large SD of the differences in Table 47, strongly suggest that there are wide departures between observed and predicted values. In the case of Kenya, there is a significant bias toward lower values but this is probably explicable by methodologic differences. The fit with FAO/WHO/UNU predictions is better than in the case of Mexico, but not much better. (It might improve if the age adjustment were included in the calculations.)

4. Variation in BMR Across Time

In Table 49, a crude estimation of the partitioning of variance is presented. For this analysis, only subjects with 3 or more replicated estimations of BMR were included Data were examined by subject first to obtain estimates of the subject mean and SD. The mean and SD of the subject means was then obtained, providing a crude estimate of the between-subject SD. The mean of the subject SD's was then taken as an estimate of the within-person SD. Both the Kenya and Mexico data sets had been adjusted to remove apparent outliers. The analyses suggest that, in each project, within-person variability across time is of similar magnitude to between-person variability. In the Mexico project, both of these are substantially larger than were seen among Kenya males.

Table 49. Crude Estimates of the Partitioning of Variance
(Adult Male Subjects)

Project	N	Mean	Between-Person Standard Deviation	Within-Person Standard Deviation
Kenya	137	1312	158	145
Mexico	82	1633	229	241

The contrast between the Mexico and Kenya data sets is made clear by a simple statistic generated during one of the screening procedures. From a total of 44 subjects with 4 or more replicates in the Mexico data set, 20 (45.4%) exhibited a range of 600 kcal/d or more among the replicate measures. In the Kenya data set, there were 76 subjects with 4 or more replicates. Of these, only three (3.9%) exhibited a range as high as 600 kcal; of those, the ranges were high in two cases because of clear outliers that had not been removed. It is absolutely clear that variability of data is much greater in the Mexican data set than in the Kenya data. In series of repeated measures of BMR under controlled laboratory conditions, the expected CV of within-person variation is in the order of 3-4%; clearly a higher CV would be expected under less controlled field conditions but the magnitude of day-to-day variation reported above is surprising and disturbing in terms of its impact on analyses (single measures of BMR must be taken as an unreliable estimate of the subjects true, persisting BMR unless it can be demonstrated that the variation seen is real in the sense of being a response to an environmental influence, including dietary effects). As will become clear from the following discussion, a part of the within-person variance in the Kenya data set is explicable by the effect of food restriction during the period of the "famine". In the Mexico data set, no such explanation was found.

a. Time and measurement - day bias: Mexico data. The observation of large within-person variation (relative to the between-person variation) reminds one of the arguments of Sukhatme and Margen. Although those arguments have had many interpretations, a fundamental claim is that individuals exhibit substantial variation in their energy needs across time (a large within-person variance). (This is not in accord with reports that under controlled conditions, BMR shows very limited within-person variability across even extended periods of time.) One inference taken from the arguments is that

this is the result of a regulated process and serves to maintain the homeostasis of energy balance. It was of interest, therefore, to explore in somewhat greater depth, the nature of the variation and its relationship to food intake. (In the following analyses, both RMR data sets have been subjected to outlier deletion rules.)

An initial examination of the Mexican data involved rank ordering of observed RMR's for each subject followed by a visual search for associations among variables. There were no obvious associations between estimated RMR and room temperature or recorded vapor pressure of water (humidity). Humidity had been a source of error in both the Egypt and Kenya data sets.

Data for each individual were then plotted by date of observation, to see if there were discernible time trends or patterns that might explain the high within-person variability. Clear patterns were not visually apparent. (The Mexico project had noted that RMR determinations in the first month or two after start-up were low; a check showed that those observations were not driving the following analyses.) As a further examination the directions of change between adjacent observations were analyzed with results as shown in Table 50. The directions of deviations are consistent with the notion of random variance in a binomial distribution; trends were unlikely to be present.

Table 50. Comparison of Directions of Change in RMR Within Persons Across Time (Mexico Data)

Number of Observations	Possible Pattern *	Binomial Distribution	Observed Frequency	Expected Frequency	Chi Square
4	D D D D	0.0625	0	1.125	
(n = 18)	D D D U	0.25	5	4.5	
	D D U U	0.375	10	6.75	
	D U U U	0.25	2	4.5	
	U U U U	0.0625	1	1.125	4.15
5	D D D D D	0.03125	0	0.34375]	**
(n = 11)	D D D D U	0.15625	3	1.71875]	
	D D D U U	0.3125	3	3.4375	
	D D U U U	0.3125	3	3.4375	
	D U U U U	0.15625	2	1.71875]	
	U U U U U	0.03125	0	0.34375]	0.540

Overall Chi Square = 4.69, df = 7, p = 0.703

* D = downward change; U = upward change. Order of changes not considered in this analysis.

** Categories collapsed to avoid spuriously high test values.

Three other examinations were undertaken. Bias of RMR measurements within individuals was examined in relation to the hypothesis that measurements made on particular days differ in some systematic manner ("measurement day effect"). If the hypothesis holds, measurements for all subjects examined on the same day would tend to be higher (or lower) than expected for the subject. For this purpose, adult males and females (not separated by physiologic state)

with 4 or more RMR estimates were selected. The overall mean of the individual's observations was calculated and the individual's daily observations were then classified as above (=1) or not above (=0) the overall mean. The patterns of 1's and 0's were then examined for each of the different types of days: days with 2 subjects measured, 3 subjects measured, 4 subjects measured, etc. In this manner a series of binomial distributions were generated. If there were no measurement-day effect then it would be expected that the patterns of 1's and 0's would follow those of a binomial distribution adjusted to the number of measurements and likelihood of a positive deviation. The ratio of 1's to total observations on each day was taken as an estimate of the probability of a positive deviation among persons measured on each class of day. The results in Table 51 strongly suggest that there was a bias in RMR's associated with particular days of measurement. In turn, this is strongly suggestive of an uncontrolled methodological effect.

Table 51. Examination of Measurement Day Effect on RMR - Mexico Data
(Males and Females Included)

Number of Observations Per Day	Number of Categories	Number of People	Proportion Positive	Degrees of Freedom	Individual Chi Square	Probability
2	3	79	0.557	1	6.2790	0.0122
3	4	75	0.440	2	6.7087	0.0349
4	5	58	0.509	3	5.6417	0.1304
5	6	28	0.486	4	11.4423	0.0096
6	7	31	0.597	5	20.6290	0.0003
		Total Model:		13	50.7007	0.0000

Given an apparent effect of measurement day, two further examinations were made to see if this were attributable to either the sequence in which RMR was measured (a proxy for the time of day) or the day of the week on which RMR was measured. In each case the full data set (after application of outlier rules) was examined. Sequence was inferred from a ranking of ventilation meter readings; day of the week was developed from reported dates. Simple means and standard deviations were calculated. The results are reported in Tables 52 and 53.

Table 52. Effect of Day of the Week on RMR Estimates - Mexico

Day of the Week	Adult Males			Adult Females		
	N	RMR		N	RMR	
Monday	100	1662	(318.8)*	248	1337	(238.0)
Tuesday	76	1605	(365.0)	199	1340	(228.9)
Wednesday	88	1666	(329.3)	178	1367	(239.7)
Thursday	96	1617	(287.8)	241	1346	(244.2)
Friday	94	1623	(319.8)	219	1352	(243.5)
Saturday	10	1572	(283.6)	23	1333	(221.9)
Sunday	7	1485	(293.2)	6	1390	(199.9)

* Standard Deviation

Table 53. Effect of Sequence of Measurement During the Day on RMR Estimates - Mexico

Sequence of Measurement	Adult Males			Adult Females		
	N	RMR		N	RMR	
1	130	1598	(322.5) *	244	1304	(234.0)
2	89	1603	(309.0)	233	1354	(234.7)
3	81	1659	(293.7)	189	1343	(229.5)
4	61	1585	(330.8)	151	1340	(229.3)
5	49	1681	(309.6)	112	1400	(261.0)
6	31	1699	(407.9)	72	1365	(263.4)
7	10	1700	(238.6)	54	1370	(237.1)
8	13	1731	(318.0)	24	1397	(219.7)
9	2	1812	(454.0)	19	1340	(199.9)
10	3	1919	(523.5)	8	1463	(326.6)
11	2	1812	(454.0)	5	1444	(152.2)

* Standard Deviation

It may be concluded from these simple analyses that a likely explanation of the large within-person variation (and part of the between-person variation) in the Mexico data set is attributable to a methodological source associated with a bias in measurements on particular days. The specific source of this bias has not been identified. Further studies are needed before the data set can be fully exploited as a variable in other analyses. In the meantime, it may be concluded that the variance seen is not an indication of support for the Sukhatme-Margen hypothesis concerning true biological variation in response to regulatory processes. An important implication of these findings is the need for considerable care in the interpretation of data analyses involving recorded RMR measurements or variables based on such measurements. In the statistical sense, these measures must be seen as including substantial error terms (presumed at present to be random across time) which will impact on data analyses.

b. Season and food shortage effect: Kenya data. In preliminary exploration of the Kenya data set a different approach was possible and indeed was mandated. Here a natural event, the "famine", imposed an exogenous pressure of restricted food intake on the body systems. If the processes of energy metabolism are capable of homeostatic responses to maintain energy balance in the face of such pressures, one might expect to find an effect of restricted intake on measured RMR - a fall during the period of food shortage and a rise during the recovery period. This hypothesis was examined through a series of paired analyses (comparing RMR estimates for the same subject in two phases of the the food shortage/recovery period). Toward this purpose, and based on assessments offered by the Kenya project (see Kenya Project Report), periods were defined as shown in Table 54.

The term "famine" has been used in describing this period in Kenya; in contrast to famine as usually so-designated in the literature and as occurred in Ethiopia, the situation in project site in Kenya should be seen more as a period of severe food shortage than as a typical "famine". The period defined as "prefamine" might better be considered to be a period during which the signs of an impending food shortage were increasing in the community (i.e. a transitional period in which effects might be present but not grossly visible). In the "late famine" period, the project had taken the initiative to introduce a food distribution program in response to perceived serious need. Actual food distribution began on September 15, 1984. These aspects are described in full in the Kenya Project Report. Household food intake data suggest that the period of food shortage persisted through February, 1985.

Table 54. Operating Definitions of "Famine" Periods - Kenya Study

Period Designation	Boundary Dates
Prefamine Period	Prior to June 30, 1984
Famine Period	July 1, 1984 to Nov. 30, 1984
Early	July 1, 1984 to Sept. 14, 1984
Late	Sept. 15, 1984 to Nov. 30, 1984
Postfamine Period	Dec. 1, 1984 and later
Early	Dec. 1, 1984 to April 30, 1985
Late	May 1, 1985 and later

Within these time periods, replicated estimates of RMR and other variables for the same individual were pooled. Analyses were conducted without weighting. Probability values shown in the following tables relate to the change in measures between periods within individuals and were based on a two-tailed test.

Table 55 displays the numbers of repeated measurements available for analysis. Table 56 presents the mean weights by famine period for adult men and women and for schoolers. On paired analyses significant weight changes attributable to famine were demonstrated in adult men and women but not in schoolers who were still growing. (As suggested in the Kenya project report, gains in schoolers may have been less than would have been expected in the absence of the famine. This hypothesis was not tested.)

Table 55. Frequency of Replicated Measures, Kenya Project

Number of	Adult Male			Adult Female			Schoolers		
Measurement	RMR	Weight	RMR/Wt	RMR	Weight	RMR/Wt	RMR	Weight	RMR/Wt
0	7	6	8	1	1	1	3	3	3
1	51	48	50	20	19	21	33	37	38
2	45	45	47	46	41	46	49	40	44
3	61	61	61	46	47	48	25	29	25
4	46	49	45	64	61	65	26	29	28
5	24	24	23	38	46	36	9	7	7
6	5	5	5	22	19	20	1	1	1
7	1	1	1	8	10	8	1	1	1
8	0	1	0	1	2	1	0	0	0

Table 56. Paired Analyses of Famine Effects on Body Weight *
Kenya Project

Group	Comparison Periods	N	Mean Change (kg)	SD	Prob |T|
Adult Men	Prefamine to Famine Period	86	- 1.63	4.16	0.0005
	Prefamine to Late Famine	76	- 1.70	4.34	0.0010
	Famine to Postfamine Period	109	+ 2.16	4.14	0.0001
	Late Famine to Late Postfam.	55	+ 2.54	4.49	0.0001
	Prefamine to Postfamine Per.	95	+ 0.15	3.44	0.6804
	Prefamine to Late Postfamine	49	+ 0.97	3.63	0.0677
	Early to Late Famine Period	15	+ 0.69	4.53	0.5668
	Early to Late Postfamine Per.	92	+ 0.65	5.24	0.2361
Adult Women	Prefamine to Famine Period	124	- 1.52	4.40	0.0002
	Prefamine to Late Famine	104	- 2.30	4.03	0.0001
	Famine to Postfamine Period	137	+ 0.63	3.83	0.0506
	Late Famine to Late Postfam.	57	+ 2.25	4.70	0.0006
	Prefamine to Postfamine Per.	134	- 0.65	4.46	0.0920
	Prefamine to Late Postfamine	66	+ 1.46	4.37	0.0083
	Early to Late Famine Period	51	- 2.27	4.14	0.0003
	Early to Late Postfamine Per.	103	+ 0.54	3.13	0.0805
Schoolers	Prefamine to Famine Period	43	+ 0.175	1.53	0.4589
	Prefamine to Late Famine	37	+ 0.038	1.25	0.8533
	Famine to Postfamine Period	47	+ 1.63	3.11	0.0008
	Late Famine to Late Postfam.	20	+ 1.59	1.48	0.0001
	Prefamine to Postfamine Per.	47	+ 1.73	3.06	0.0003
	Prefamine to Late Postfamine	17	+ 2.39	1.35	0.0001
	Early to Late Famine Period	5	+ 0.432	0.66	0.2145
	Early to Late Postfamine Per.	36	+ 0.938	1.06	0.0001

* Replicate observations pooled within time periods; analyses not weighted for number of observations.

NOTE: Numbers of subjects in early + late stages of a period do not add to number of subjects in period since the same subject may be represented in both subperiods.

In males, in the absence of famine, there would be no a priori reason to expect a systematic change in RMR across time. In women, there could be effects of progressive changes during pregnancy; in a data set of the present structure these could lead to a trend in longitudinal data. Among schoolers, any growth-related effects would appear as trends in the data. The present preliminary analyses do not control for such possible effects. The adult male data are most easily interpreted. In later and more detailed examinations, expected data trends attributable to physiologic variables (pregnancy, lactation, growth) can be taken into account, perhaps using estimates of these effects derived from the other projects.

Tables 57-59 present the results of paired analysis of the effects of famine on RMR and food intake, expressed in terms of absolute levels and levels per unit body weight. In these analyses, subjects without measures of food

intake, RMR and weight for the specific pair of periods under consideration were excluded; the number of subjects is constant for all analyses within a period pair but numbers (and individuals) differ between sets of paired comparisons. The notable findings for adult men and women are that famine strongly affected food intake and to a much lesser degree affected RMR. Indeed the effect on RMR is not consistently seen in comparisons of prefamine and total famine periods; when the second contrast, famine to postfamine, or the even more striking contrast, late famine to late postfamine, are examined, it is clear that RMR was affected. However, when RMR was expressed in relation to body weight, the effects virtually disappeared. A logical conclusion is that the RMR fell in response to a change in body weight (a reduced mass of metabolically active tissue) rather than as a primary response to altered food intake. Across the 6 week to 3+ month periods during which data were averaged, changes in food intake and in RMR were not parallel in magnitude or even consistent in direction. A much more careful analysis (e.g. time lagging) would be needed to examine the precise relationship among food intake, body weight (lean body mass), and RMR in a definitive manner.

Table 57. Impact of "Famine" on RMR and Food Intake: Kenya Adult Males
(Mean ± Standard Deviation)

Units and Variable		Period	Mean	Difference	P |T|
Prefamine to Famine Period N = 83					
Kcal/day	RMR	Prefamine	1289 ± 181		
		Famine	1263 ± 171	-25.8 ± 182.2	0.0001
	Intake	Prefamine	2187 ± 568		
		Famine	1775 ± 469	-412 ± 563	0.0001
	Avail.	Prefamine	898 ± 565		
	Energy	Famine	511 ± 476	-386 ± 613	0.0001
Kcal/kg	RMR	Prefamine	23.4 ± 3.24		
		Famine	23.5 ± 2.75	+0.15 ± 3.36	0.6855
	Intake	Prefamine	39.6 ± 9.84		
		Famine	33.0 ± 8.49	-6.57 ± 10.99	0.0001
	Avail.	Prefamine	16.2 ± 10.05		
	Energy	Famine	9.48 ± 8.98	-6.72 ± 11.28	0.0001
Prefamine to Postfamine Period N = 88					
Kcal/day	RMR	Prefamine	1281 ± 197		
		Postfamine	1341 ± 192	+60.1 ± 241	0.0217
	Intake	Prefamine	2124 ± 560		
		Postfamine	2024 ± 486	-100.0 ± 559	0.0971
	Avail.	Prefamine	843 ± 582		
	Energy	Postfamine	683 ± 485	-160 ± 601	0.0144
Kcal/kg	RMR	Prefamine	23.2 ± 3.95		
		Postfamine	24.3 ± 3.54	+1.02 ± 4.52	0.0364
	Intake	Prefamine	38.5 ± 10.5		
		Postfamine	36.6 ± 8.91	-1.90 ± 9.73	0.0698
	Avail.	Prefamine	15.3 ± 10.37		
	Energy	Postfamine	12.4 ± 8.87	-2.93 ± 10.53	0.0107
Prefamine to Late Famine Period N = 72					
Kcal/day	RMR	Prefamine	1278 ± 186		
		Late Famine	1259 ± 178	-18.8 ± 192	0.4086
	Intake	Prefamine	2168 ± 580		
		Late Famine	1610 ± 507	-558 ± 558	0.0001
	Avail.	Prefamine	890 ± 577		
	Energy	Late Famine	351 ± 507	-539 ± 580	0.0001
Kcal/kg	RMR	Prefamine	23.4 ± 3.36		
		Late Famine	23.7 ± 2.88	+0.32 ± 3.52	0.4488
	Intake	Prefamine	39.6 ± 10.10		
		Late Famine	30.0 ± 8.67	-9.51 ± 11.19	0.0001
	Avail.	Prefamine	16.2 ± 10.30		
	Energy	Late Famine	6.36 ± 9.44	-9.82 ± 11.05	0.0001

Table 57 Cont'd. Impact of "Famine" on RMR and Food Intake: Kenya Adult Males

Units and Variable		Period	Mean	Difference	P [T]
Prefamine to Late Postfamine Period N = 44					
Kcal/day	RMR	Prefamine	1313 ± 210		
		Late Postfam.	1369 ± 182	+56.3 ± 219	0.0955
	Intake	Prefamine	2093 ± 593		
		Late Postfam.	2024 ± 562	-68.5 ± 642	0.4831
	Avail.	Prefamine	779 ± 583		
	Energy	Late Postfam.	655 ± 557	-125 ± 713	0.2520
Kcal/kg	RMR	Prefamine	23.8 ± 3.68		
		Late Postfam.	24.5 ± 3.43	+0.74 ± 3.96	0.2217
	Intake	Prefamine	37.9 ± 10.20		
		Late Postfam.	36.1 ± 9.38	-1.81 ± 11.16	0.2896
	Avail.	Prefamine	14.1 ± 10.03		
	Energy	Late Postfam.	11.6 ± 9.30	-2.55 ± 12.39	0.1802
Famine to Postfamine Period N = 109					
Kcal/day	RMR	Famine	1268 ± 171		
		Postfamine	1330 ± 176	+61.4 ± 170	0.0003
	Intake	Famine	1793 ± 461		
		Postfamine	2020 ± 442	+227 ± 462	0.0001
	Avail.	Famine	524 ± 463		
	Energy	Postfamine	690 ± 429	+166 ± 474	0.0004
Kcal/kg	RMR	Famine	23.9 ± 2.89		
		Postfamine	24.0 ± 3.01	+0.18 ± 3.46	0.5938
	Intake	Famine	33.8 ± 9.08		
		Postfamine	36.6 ± 8.44	+2.77 ± 9.52	0.0030
	Avail.	Famine	9.97 ± 9.04		
	Energy	Postfamine	12.6 ± 8.07	+2.60 ± 9.19	0.0039
Late Famine to Late Postfamine Period N = 75					
Kcal/day	RMR	Late Famine	1267 ± 190		
		Late Postfam.	1367 ± 200	+99.8 ± 184	0.0001
	Intake	Late Famine	1696 ± 491		
		Late Postfam.	2030 ± 471	+334 ± 485	0.0001
	Avail.	Late Famine	429 ± 508		
	Energy	Late Postfam.	664 ± 463	+235 ± 497	0.0009
Kcal/kg	RMR	Late Famine	24.0 ± 2.71		
		Late Postfam.	24.8 ± 3.29	+0.80 ± 3.60	0.1016
	Intake	Late Famine	32.3 ± 9.97		
		Late Postfam.	36.8 ± 8.66	+4.57 ± 10.42	0.0019
	Avail.	Late Famine	8.30 ± 10.17		
	Energy	Late Postfam.	12.1 ± 8.39	+3.77 ± 9.76	0.0059

Table 57 Cont'd. Impact of "Famine" on RMR and Food Intake: Kenya Adult Males

Units and Variable		Period	Mean	Difference	P |T|
Early Famine to Late Famine Period N = 15					
Kcal/day	RMR	Early Famine	1257 ± 206		
		Late Famine	1285 ± 179	+27.9 ± 172	0.0001
	Intake	Famine	1943 ± 515		
		Late Famine	1707 ± 425	-236 ± 607	0.[illegible]
	Avail.	Early Famine	686 ± 512		
	Energy	Late Famine	423 ± 361	-264 ± 705	0.0009
Kcal/kg	RMR	Early Famine	23.7 ± 3.80		
		Late Famine	23.7 ± 2.15	-0.01 ± 3.92	0.1016
	Intake	Early Famine	36.9 ± 10.89		
		Late Famine	31.3 ± 6.25	-5.58 ± 14.25	0.0019
	Avail.	Early Famine	13.2 ± 9.92		
	Energy	Late Famine	7.61 ± 6.64	-5.57 ± 14.11	0.0059
Early Postfamine to Late Postfamine Period N = 88					
Kcal/day	RMR	Early Postfam	1322 ± 232		
		Late Postfam.	1392 ± 197	+69.9 ± 218	0.0034
	Intake	Early Postfam	1957 ± 463		
		Late Postfam.	2009 ± 436	+52.5 ± 479	0.3064
	Avail.	Early Postfam	635 ± 530		
	Energy	Late Postfam.	618 ± 473	-17.4 ± 539	0.7026
Kcal/kg	RMR	Early Postfam	24.3 ± 4.15		
		Late Postfam.	25.4 ± 3.34	+1.08 ± 4.17	0.0159
	Intake	Early Postfam	36.2 ± 9.70		
		Late Postfam.	36.7 ± 8.04	+0.48 ± 9.29	0.6308
	Avail.	Early Postfam	11.9 ± 10.41		
	Energy	Late Postfam.	11.3 ± 8.55	-0.61 ± 10.05	0.5758

Table 58. Impact of "Famine" on RMR and Food Intake: Kenya Adult Females (Mean ± Standard Deviation)

Units and Variable		Period	Mean	Difference	P |T|
Prefamine to Famine Period N = 117					
Kcal/day	RMR	Prefamine	1190 ± 166		
		Famine	1164 ± 169	-26.0 ± 204	0.1706
	Intake	Prefamine	1728 ± 515		
		Famine	1470 ± 437	-257 ± 495	0.0001
	Avail.	Prefamine	537 ± 526		
	Energy	Famine	306 ± 440	-231 ± 520	0.0001
Kcal/kg	RMR	Prefamine	23.3 ± 3.55		
		Famine	23.4 ± 3.42	+0.13 ± 4.06	0.7343
	Intake	Prefamine	33.8 ± 9.67		
		Famine	29.7 ± 8.81	-4.12 ± 10.28	0.0001
	Avail.	Prefamine	10.5 ± 9.90		
	Energy	Famine	6.21 ± 8.70	-4.25 ± 10.48	0.0001
Prefamine to Postfamine Period N = 128					
Kcal/day	RMR	Prefamine	1183 ± 186		
		Postfamine	1227 ± 186	+43.3 ± 201	0.0163
	Intake	Prefamine	1746 ± 457		
		Postfamine	1764 ± 400	+17.8 ± 527	0.7031
	Avail.	Prefamine	562 ± 476		
	Energy	Postfamine	537 ± 444	-25.5 ± 577	0.6178
Kcal/kg	RMR	Prefamine	23.3 ± 3.85		
		Postfamine	24.4 ± 3.37	+1.11 ± 4.29	0.0042
	Intake	Prefamine	34.4 ± 9.27		
		Postfamine	35.3 ± 8.71	+0.89 ± 11.54	0.3822
	Avail.	Prefamine	11.1 ± 9.41		
	Energy	Postfamine	10.9 ± 9.08	-0.21 ± 11.65	0.8377
Prefamine to Late Famine Period N = 99					
Kcal/day	RMR	Prefamine	1181 ± 162		
		Late Famine	1125 ± 175	-56.6 ± 216	0.0105
	Intake	Prefamine	1727 ± 515		
		Late Famine	1349 ± 452	-378 ± 562	0.0001
	Avail.	Prefamine	545 ± 525		
	Energy	Late Famine	224 ± 454	-321 ± 582	0.0001
Kcal/kg	RMR	Prefamine	23.2 ± 3.47		
		Late Famine	23.1 ± 3.66	-0.11 ± 4.22	0.7794
	Intake	Prefamine	33.9 ± 9.83		
		Late Famine	27.6 ± 9.12	-6.29 ± 11.48	0.0001
	Avail.	Prefamine	10.7 ± 9.92		
	Energy	Late Famine	4.51 ± 9.36	-6.17 ± 11.88	0.0001

Table 58 Cont'd. Impact of "Famine" on RMR and Food Intake: Kenya Adult Females

Units and Variable		Period	Mean	Difference	P |T|
Prefamine to Late Postfamine Period N = 62					
Kcal/day	RMR	Prefamine	1183 ± 192		
		Late Postfam.	1239 ± 161	+55.9 ± 201	0.0325
	Intake	Prefamine	1781 ± 446		
		Late Postfam.	1746 ± 440	-34.8 ± 543	0.6157
	Avail.	Prefamine	597 ± 465		
	Energy	Late Postfam.	506 ± 425	-91 ± 565	0.2110
Kcal/kg	RMR	Prefamine	23.7 ± 4.06		
		Late Postfam.	24.3 ± 3.55	+0.56 ± 4.45	0.3226
	Intake	Prefamine	35.7 ± 8.84		
		Late Postfam.	34.2 ± 8.53	-1.50 ± 10.93	0.2841
	Avail.	Prefamine	12.0 ± 9.26		
	Energy	Late Postfam.	9.92 ± 8.07	-2.06 ± 11.02	0.1460
Famine to Postfamine Period N = 134					
Kcal/day	RMR	Famine	1169 ± 179		
		Postfamine	1215 ± 189	+45.5 ± 206	0.0117
	Intake	Famine	1468 ± 410		
		Postfamine	1779 ± 376	+310 ± 401	0.0001
	Avail.	Famine	299 ± 416		
	Energy	Postfamine	563 ± 412	+265 ± 430	0.0001
Kcal/kg	RMR	Famine	23.4 ± 3.34		
		Postfamine	23.9 ± 3.02	+0.55 ± 4.05	0.1203
	Intake	Famine	29.5 ± 8.45		
		Postfamine	35.2 ± 7.85	+5.79 ± 8.30	0.0001
	Avail.	Famine	6.11 ± 8.29		
	Energy	Postfamine	11.4 ± 8.31	+5.24 ± 8.65	0.0001
Late Famine to Late Postfamine Period N = 57					
Kcal/day	RMR	Late Famine	1160 ± 186		
		Late Postfam.	1225 ± 176	+64.9 ± 199	0.0169
	Intake	Late Famine	1372 ± 475		
		Late Postfam.	1831 ± 460	+458 ± 464	0.0001
	Avail.	Late Famine	212 ± 470		
	Energy	Late Postfam.	606 ± 436	+394 ± 465	0.0001
Kcal/kg	RMR	Late Famine	23.8 ± 3.57		
		Late Postfam.	24.0 ± 3.36	+0.25 ± 4.44	0.6765
	Intake	Late Famine	28.1 ± 9.48		
		Late Postfam.	35.9 ± 8.79	+7.84 ± 10.26	0.0001
	Avail.	Late Famine	4.30 ± 9.63		
	Energy	Late Postfam.	11.9 ± 8.38	+7.60 ± 9.51	0.0001

Table 58 Cont'd. Impact of "Famine" on RMR and Food Intake: Kenya Adult Females

Units and Variable		Period	Mean	Difference	P |T|
Early Famine to Late Famine Period			N = 49		
Kcal/day	RMR	Early Famine	1227 ± 195		
		Late Famine	1118 ± 179	-109 ± 210	0.0007
	Intake	Famine	1574 ± 494		
		Late Famine	1282 ± 369	-292 ± 497	0.0002
	Avail.	Early Famine	347 ± 551		
	Energy	Late Famine	164 ± 401	-183 ± 555	0.0255
Kcal/kg	RMR	Early Famine	23.9 ± 3.37		
		Late Famine	22.7 ± 3.16	-1.15 ± 3.76	0.0377
	Intake	Early Famine	31.4 ± 11.39		
		Late Famine	26.3 ± 8.19	-5.11 ± 10.95	0.0020
	Avail.	Early Famine	7.49 ± 10.89		
	Energy	Late Famine	3.53 ± 8.51	-3.97 ± 11.74	0.0221
Early Postfamine to Late Postfamine Period			N = 100		
Kcal/day	RMR	Early Postfam	1198 ± 185		
		Late Postfam.	1243 ± 160	+45.1 ± 165	0.0073
	Intake	Early Postfam	1745 ± 445		
		Late Postfam.	1805 ± 404	+60.4 ± 465	0.1960
	Avail.	Early Postfam	547 ± 484		
	Energy	Late Postfam.	562 ± 398	+15.4 ± 520	0.7682
Kcal/kg	RMR	Early Postfam	23.7 ± 3.42		
		Late Postfam.	24.4 ± 3.39	+0.73 ± 3.32	0.0300
	Intake	Early Postfam	34.4 ± 8.41		
		Late Postfam.	35.5 ± 8.15	+1.08 ± 9.42	0.2546
	Avail.	Early Postfam	10.8 ± 9.53		
	Energy	Late Postfam.	11.1 ± 7.71	+0.35 ± 10.25	0.7346

Table 59. Impact of "Famine" on RMR and Food Intake: Kenya Schoolers
(Mean ± Standard Devintion)

Units and Variable		Period	Mean	Difference	P |T|
Prefamine to Famine Period N = 42					
Kcal/day	RMR	Prefamine	701 ± 149		
		Famine	712 ± 152	+10.7 ± 159	0.6645
	Intake	Prefamine	1581 ± 390		
		Famine	1339 ± 351	-242 ± 393	0.0003
	Avail.	Prefamine	879 ± 334		
	Energy	Famine	627 ± 344	-252 ± 396	0.0002
Kcal/kg	RMR	Prefamine	34.9 ± 5.78		
		Famine	35.4 ± 6.68	+0.43 ± 8.24	0.7357
	Intake	Prefamine	79.1 ± 17.96		
		Famine	66.5 ± 16.33	-12.5 ± 19.06	0.0001
	Avail.	Prefamine	44.2 ± 16.24		
	Energy	Famine	31.2 ± 16.73	-13.0 ± 18.89	0.0001
Prefamine to Postfamine Period N = 46					
Kcal/day	RMR	Prefamine	722 ± 197		
		Postfamine	810 ± 192	+87.6 ± 189	0.0029
	Intake	Prefamine	1589 ± 560		
		Postfamine	1409 ± 486	-180 ± 310	0.0003
	Avail.	Prefamine	866 ± 582		
	Energy	Postfamine	598 ± 485	-268 ± 365	0.0001
Kcal/kg	RMR	Prefamine	36.9 ± 3.95		
		Postfamine	38.3 ± 3.54	+1.56 ± 8.75	0.2319
	Intake	Prefamine	80.8 ± 10.5		
		Postfamine	67.2 ± 8.91	-13.7 ± 15.58	0.0001
	Avail.	Prefamine	44.1 ± 10.37		
	Energy	Postfamine	28.8 ± 8.87	-15.2 ± 17.28	0.0001
Prefamine to Late Famine Period N = 36					
Kcal/day	RMR	Prefamine	691 ± 186		
		Late Famine	698 ± 178	+7.2 ± 158	0.7865
	Intake	Prefamine	1581 ± 580		
		Late Famine	1267 ± 507	-313 ± 384	0.0001
	Avail.	Prefamine	890 ± 577		
	Energy	Late Famine	570 ± 507	-320 ± 359	0.0001
Kcal/kg	RMR	Prefamine	34.5 ± 3.36		
		Late Famine	34.9 ± 2.88	+0.39 ± 8.02	0.7705
	Intake	Prefamine	79.2 ± 10.10		
		Late Famine	63.4 ± 8.67	-15.9 ± 19.32	0.0001
	Avail.	Prefamine	44.7 ± 10.30		
	Energy	Late Famine	28.5 ± 9.44	-16.3 ± 17.77	0.0001

Table 59. Cont'd. Impact of "Famine" on RMR and Food Intake: Kenya Schoolers.

Units and Variable		Period	Mean	Difference	P |T|
Prefamine to Late Postfamine Period N = 15					
Kcal/day	RMR	Prefamine	677 ± 123		
		Late Postfam.	750 ± 117	+73.2 ± 140	0.0617
	Intake	Prefamine	1479 ± 237		
		Late Postfam.	1484 ± 403	+4.8 ± 355	0.9590
	Avail.	Prefamine	802 ± 171		
	Energy	Late Postfam.	733 ± 402	-68.3 ± 396	0.5143
Kcal/kg	RMR	Prefamine	36.1 ± 5.09		
		Late Postfam.	35.8 ± 6.48	-0.24 ± 7.35	0.8975
	Intake	Prefamine	79.3 ± 13.15		
		Late Postfam.	70.6 ± 17.11	-8.69 ± 13.53	0.0261
	Avail.	Prefamine	43.2 ± 10.54		
	Energy	Late Postfam.	34.8 ± 16.99	-8.44 ± 16.31	0.0648
Famine to Postfamine Period N = 44					
Kcal/day	RMR	Famine	710 ± 127		
		Postfamine	771 ± 134	+60.9 ± 186	0.0358
	Intake	Famine	1347 ± 294		
		Postfamine	1460 ± 337	+113 ± 307	0.0186
	Avail.	Famine	636 ± 271		
	Energy	Postfamine	689 ± 374	+52.2 ± 385	0.3739
Kcal/kg	RMR	Famine	36.1 ± 5.86		
		Postfamine	36.3 ± 5.71	+0.26 ± 8.14	0.8295
	Intake	Famine	68.5 ± 15.09		
		Postfamine	69.4 ± 16.88	+0.90 ± 16.59	0.7178
	Avail.	Famine	32.5 ± 13.87		
	Energy	Postfamine	33.1 ± 17.25	+0.64 ± 18.69	0.8203
Late Famine to Late Postfamine Period N = 18					
Kcal/day	RMR	Late Famine	704 ± 111		
		Late Postfam.	793 ± 85.8	+88.9 ± 151	0.0228
	Intake	Late Famine	1265 ± 337		
		Late Postfam.	1548 ± 348	+282 ± 409	0.0094
	Avail.	Late Famine	561 ± 333		
	Energy	Late Postfam.	754 ± 363	+194 ± 444	0.0821
Kcal/kg	RMR	Late Famine	36.6 ± 5.16		
		Late Postfam.	38.1 ± 5.70	+1.58 ± 7.44	0.3789
	Intake	Late Famine	66.0 ± 18.55		
		Late Postfam.	74.4 ± 17.92	+8.33 ± 22.28	0.1310
	Avail.	Late Famine	29.5 ± 17.96		
	Energy	Late Postfam.	36.2 ± 17.63	+6.75 ± 23.67	0.2430

Table 59. Cont'd. Impact of "Famine" on RMR and Food Intake: Kenya Schoolers.

Units and Variable		Period	Mean	Difference	P |T|
Early Famine to Late Famine Period N = 5 (too low for examination)					
Early Postfamine to Late Postfamine Period N = 31					
Kcal/day	RMR	Early Postfam	721 ± 150		
		Late Postfam.	776 ± 102	+55.6 ± 167	0.0602
	Intake	Early Postfam	1441 ± 274		
		Late Postfam.	1558 ± 357	+117 ± 330	0.0469
	Avail.	Early Postfam	720 ± 300		
	Energy	Late Postfam.	781 ± 368	+61.1 ± 361	0.3309
Kcal/kg	RMR	Early Postfam	36.3 ± 7.89		
		Late Postfam.	37.4 ± 6.16	+1.12 ± 9.35	0.4905
	Intake	Early Postfam	72.5 ± 13.65		
		Late Postfam.	74.5 ± 15.65	+2.02 ± 15.33	0.4465
	Avail.	Early Postfam	36.2 ± 15.11		
	Energy	Late Postfam.	37.1 ± 16.26	+0.91 ± 16.92	0.7566

5. Implications for CRSP Hypothesis

These observations have important relevance for the Nutrition CRSP. They strongly suggest that changes in food intake are not absorbed by compensatory changes in metabolism (energy homeostasis) and hence must be associated with other adaptive mechanisms - change in body tissue mass as demonstrated in the famine period and/or change in activity-demanding behaviors as originally hypothesized in the design of the CRSP.

In the present tables, this hypothesis is illustrated (not tested) in the display of the estimated "available energy" computed as intake-RMR. True energy utilization is greater than the available energy intake during the famine since the loss of tissue mass represents a conversion to energy. Nonetheless, the presented data suggest strongly that some behavioral effects of reduced food intake during the famine period should be found. In drawing inferences such as these, it would be of interest to factor in the likely "release" of energy for functional use during weight loss (going into the famine) and the "removal" of energy, from other functional use, for deposition of tissues during recovery from the famine period. Such phenomena should be taken into account in analyses of the relationship of food intake to other variables, particularly within the Kenya data set.

In any analyses of the "available energy", it will be important to take into consideration the demonstrated within-person variability of RMR. If indeed this is not being driven by food intake or other biological pressures (as is suggested by the Kenya analysis), then it must be considered as "noise" in the measurements and hence a major source of presumably random error in any data analyses. Given that a high within-person variance has always been recognized for food intake, data analyses and interpretations must accept that the random error component of an "available energy" variable based on observed intake-observed RMR will be very high. To the degree that it is possible to average the intake and RMR measurements of individuals across time, this error can be reduced. In follow-up analyses of CRSP data it will be important to further explore the within-person variation to ensure that it is, in major part, methodologic variation as has been assumed in this discussion.

Among schoolers, as previously noted, interpretation of RMR is more difficult because of an expected change over time. (Such trends can be seen in all groups through comparisons of prefamine and postfamine values and from comparisons of early to late postfamine periods.) Nevertheless, two observations are striking. The "available energy" per kg is much greater in the schoolers than in the adults and does not fall to as low a level in the schooler as in the adult. It is of interest to speculate that compensatory behavioral changes during the famine may be quite different between schoolers and adults.

In an analogous manner, care must be taken in the interpretation of data for adult women. The present analyses have not examined pregnant, lactating and nonpregnant-nonlactating women separately. Country project reports (Mexico and Kenya) do provide analyses of RMR in pregnancy and lactation. In general it appears that RMR increases somewhat during pregnancy and returns to prepregnancy levels early in lactation, in agreement with findings recently

reported from Scotland, Holland, the Gambia, the Philippines and Thailand (15). Among nonpregnant-nonlactating women, a further source of within-person variance in RMR is the known effect of the menstrual cycle on RMR; measurements made in the CRSP studies were not controlled for this variable.

Comment is offered also on the appearance of an increased variability in response to food deprivation and recovery (see standard deviations of the differences). This is misleading. If one accepts the estimate of within-person standard deviation of RMR developed for Kenya in Table 49 as a measure of the error of estimation of RMR, then the standard deviations of differences are of the order of magnitude that would be expected for the difference of two measures, each of which carries this error term.

Although it is clear that all three RMR data sets require further examination from a purely methodologic standpoint (to correct remaining errors, see Annex to this chapter), they hold considerable promise to the interested and careful analyst.

6. Technical Methodologic Problems Remaining

a. Outlier deletion rules. For the analysis reported above, data were deleted according to the following rules:

Kenya Data were rejected if the computed concentration of oxygen in expired air was less that 10%. This served as an index of probable data entry errors in the gas volume variable.

Mexico Data were excluded if the record indicated illness on the day of RMR, if the difference between the two measures of respiratory volume was greater than 3.0, or if the difference between the two measures of expired oxygen concentration was greater than 0.80, or if the expired oxygen concentration was greater than 20.0%. In its report, the Mexico project applied different exclusion rules. There are minor differences in data sets analyzed (original 579 males, Mexico analyses, 565; present analyses, 527).

ME analyses were based on initially archived data to which deletion rules described above were applied. In the project reports, analyses may have applied different criteria for exclusion and may have included some data collections not then available to ME. Small discrepancies in analyses may be present.

b. Problems and potential solutions - Egypt RMR data set. There appear to be two sources of error in the Egypt project RMR data set. It is known that humidity affected RMR determinations in Egypt, as in Kenya, since the oxygen concentration in ambient air was measured in the moist state (thereby lowering the concentration of oxygen). The project RMR operator reported that they had experienced difficulty in obtaining stable oxygen concentration measurements until the machine was completely warmed up. They had compensated by frequent readjustments of the O2 detector, between subjects, until the drift in readings stopped. Nevertheless, it was presumed that, for some subjects at least, an error attributable to drift in the O2 meter, as well as an error attributable to humidity, is present. The Kenya project developed a procedure

for correcting the humidity-related error. However, as will be discussed below, application of this procedure can be expected to exacerbate the machine drift error (create an error in the opposite direction).

Sensitivity testing suggested that the humidity-related error in RMR attributable to humidity might be 60-160 kcal/d underestimation. By itself the O2 detector drift (which affects both inspired and expired O2 concentrations) is unlikely to contribute any appreciable error in estimated RMR (perhaps in the order of 1-2 kcal/d). However, application of the Kenya humidity correction procedure to data including the O2 detector drift error did give rise to apparent error (positive deviation). In the absence of internal estimates of the magnitude of the detector drift, estimates of effect on RMR cannot be offered.

In any correction procedure, it would be desirable to isolate and separately address the two error sources. The present report presents such an approach applicable to the Egypt data set. It requires that additional data (original FIO2 values) be entered from the original Beckman MMC tapes. Since this had not been done at the time of this report, the Egypt data set remained uncorrected and analyses were not undertaken. The procedure proposed below involves imputation of humidity and temperature data since these were not recorded at the project site. The project report presents data from a national weather office. No estimate of the error term associated with these imputations has been developed.

The Kenya correction procedures uses information from the Beckman printout (ambient air O2 concentration [FIO2], respiratory volume [VE(BTPS)], volume of O2 used [VO2], volume of CO2 produced [VCO2]) to compute the values for other primary variables used by the Beckman but not printed out (oxygen concentration in expired air [FEO2], carbon dioxide concentration in expired air [FECO2], and respiration volume [VE(STPD)]. The procedure then uses these estimated variables to compute a corrected value for VO2 and then RMR and RQ, using a reference ("corrected") value for FIO2 (0.2093) in place of the moist air concentration originally used.

This proposed correction assumes that

Observed FIO2 = 0.2093 + Humidity Error (HE)
Observed FEO2 = Observed FEO2 (i.e. correct)

If there is a component of machine error, then

Observed FIO2 = 0.2093 + Humidity Error (HE) + Machine Error (ME)
Observed FEO2 = Observed FEO2 + Machine Error (ME)

To separate the two errors, it is proposed that the FIO2 expected for the existing humidity and temperature be estimated (FIO2humid). Following the rationale above,

Observed FIO2 = FIO2humid + ME
then ME = Observed FIO2 - FIO2humid
and HE = 0.2093 - FIO2humid

It follows, therefore, that an appropriate correction procedure for the Egypt data set would be to first estimate the FEO2 values used by the Beckman MMC,

as was done in Kenya and to then correct the FEO2 by adding a machine drift correction (ME identified above). Using the corrected FEO2 and the reference FIO2 (.2093) new VO2 values, RMR and RQ values can be computed.

To implement this correction procedure, the Egypt project will have to enter the original FIO2 values from Beckman MMC tapes (available) and enter additional temperature and humidity data (only quarterly average estimates are now available) or accept an error associated with the imprecise description of temperature and humidity. If the latter approach is taken, it may be that empiric approaches can be developed that approximate the adjustment with equivalent precision and without need to enter FIO2 data. Sensitivity analyses may serve to estimate the potential error terms generated by these alternatives and hence help to guide the selection of strategy.

(It has been reported by the operator of the equipment in Kalama that barometric pressure was entered into the Beckman as 760 to 770 mm Hg rather than 760 mm Hg expected at sea level. This might be taken into account in the correction procedures.)

c. Possible remaining error in Kenya RMR data set. It is known that in both the Egypt and Kenya project data sets there is an error in the estimate of ambient air O2 concentration (FIO2) attributable to the effect of humidity. That is, air was measured moist rather than dry. Correction procedures were developed and applied to the Kenya data sets to adjust for this error, assuming that it was the only error present.

In the case of the Egypt data set, the operator advised that there was also a probable error in the FIO2 estimates attributable to drift in the O2 detector. Such an error would affect also the estimation of expired oxygen concentration (FEO2). A procedure for estimation of both errors was developed. This procedure makes use of information about room temperature and relative humidity to estimate the FIO2 adjusted for water vapor that the machine should have read. The difference between 0.2093 (oxygen concentration of dry air) and this value is the error attributable to humidity per se. The difference between the calculated value and the recorded value of FIO2 serves as an estimate of the machine error, applicable to both FIO2 and FEO2 estimates. Using this information, it is possible to adjust the FEO2 and then correct for both humidity and machine drift effects. This is the procedure developed for application to the Egypt project data set (but not yet applied since FIO2 values had not been entered in computer files).

The stated assumption of the Kenya project was that there was no machine drift or other problem with the O2 detector. If that be the case, and if recorded room temperature and humidity values were correct, calculated moist air FIO2 values and recorded FIO2 values should agree closely. Such a comparison is presented below. In addition "fully corrected" VO2, RMR and RQ values are compared with Kenya-project corrected values to test the impact of any residual, uncorrected, error in the Kenya data set as archived.

For these comparisons, water vapour pressure was computed by the following equations:

Vapor Pressure, Saturated (in contact with water)	VPS	$= 10^{(8.179 - 1766.982/(T + 237.958))}$
Vapor Pressure, Saturated (in contact with air)	VPSA	= VPS x (1 + (0.775 - (0.000313 x T)/100)
Vapor Pressure Actual	VP	= VPSA x Relative Humidity / 100
Expected FIO2	CorrFIO2	= 0.2093 x (BAROM - VP)/BAROM
Machine Error	DFIO2	= FIO2 - CorrFIO2
Corrected FEO2	CorrFEO2	= FEO2 + DFIO2

where T is room temperature in degrees Celsius
BAROM is barometric pressure in mm Hg
FIO2 is the recorded concentration of oxygen in inspired air

The first two equations are taken from ref. 16. For the study of the Kenya data set, barometric pressure was assumed to be 660 mm Hg, the mean barometric pressure at the site of RMR measurements (there was an error in measurements of barometric pressure in part of the data set entered in the computer records). Comparisons generated are shown below. It should be noted that of 760 records in the original data set, 71 were missing data on temperature or humidity or both and hence were excluded from analyses. Other exclusions were based on missing data for weight or other variables used in the analyses. The final data set subjected to analyses comprised 533 records.

Table 60. Effect of "Full Correction" on Parameters of RMR Estimation

Kenya Male Adults

Variable	Mean	SD	Prob |T|
Room Temperature	22.68	1.72	
Relative Humidity	75.34	8.49	
Estimated Vapor Pressure	15.65	1.76	
Observed FIO2	0.2047	0.0018	
Expected FIO2	0.2043	0.0006	
Observed - Expected FIO2	+0.0004	0.0018	0.0001
Observed VO2	192.36	30.23	
Corrected VO2	194.67	30.94	
Observed - Expected VO2	-2.307	11.725	0.0001
Observed RMR	1324.8	207.08	
Corrected RMR	1342.2	210.81	
Observed - Expected RMR	-17.372	70.26	0.0001

The table suggests that there is little bias in the mean FIO2 estimates. On average, there was very little "machine error". However, it is to be noted that there is a variance in the deviation between the observed and estimated FIO2 values. This is the source of the variance in the differences between observed and corrected RMR and VO2 values. To examine the effects of this variance in greater detail, regression analyses were run to compare the observed and expected values of FIO2, the observed and corrected values of RMR, and finally, the relationships between the FAO/WHO/UNU predicted BMR and the two estimates of RMR. (The mean predicted BMR was 1519.19 ± 104.8 kcal/d.) These results are portrayed in Table 61.

Table 61. Regression Analyses of the Impact of "Full Correction" on RMR
Kenya Male Adults

Regression Equation	Rsquare
Expected FIO2 = 0.195 + 0.0451 Observed FIO2	0.022
Corrected RMR = 69.55 + 0.9606 Observed RMR	0.890
Observed RMR = -97.22 + 0.936 FAO/WHO/UNU BMR	0.224
Corrected RMR = -136.75 + 0.973 FAO/WHO/UNU BMR	0.234

It can be concluded that there is a significant discrepancy between the observed and expected FIO2 values. This is particularly apparent from the regression analysis (but note that the very low Rsquare is associated with a relatively tight range of values for both the original and expected FIO2 -- this does not invalidate the original correction procedures). The discrepancy could be attributable to either error in the estimation of FIO2 and FEO2 by the machine (O2 detector drift) or error in the estimation of room temperature and relative humidity by the operator. The former is suggested by the fact that the variance of the original FIO2 is greater than that of the expected

FIO2 (i.e. the original FIO2 may include more sources of variance than that arising from humidity alone). If it is taken as true machine error, then one can conclude that there is an error remaining in the estimated RMR values. When both the original and corrected RMR values are compared to the FAO/WHO/UNU predictions for young adult men, it is found that there is a marginally better agreement with the corrected RMR than with the original. (Since the FAO/WHO/UNU equation used was a linear function of weight with no age variable included, a similar statement would apply to the relationship between RMR and body weight). Nevertheless, the difference in Rsquare values is truly minor and the effect of correction procedures on RMR, while significant, is small at the level of the mean.

While there remains a distinct probability that there is an uncorrected error in the Kenya RMR data set of the same type that is believed to exist in the Egypt data set, to implement a correction would mean the loss of a significant portion of the data set (missing values for temperature and humidity). It would seem to be of limited benefit to implement the further correction. In analysis it will have to be recognized that the presence of this error will constitute a source of "noise" in the data and is likely to attenuate certain of the observed relationships; it is unlikely to generate spurious relationships. The caution, then, is in interpreting negative findings.

F. Anthropometry

1. Variables Collected

Anthropometric variables collected by projects on all target individuals included weight, height, a series of skin fold measures and, for infants, certain other variables. The variables common to all projects are listed below. Measurement intervals were not necessarily identical.

Table 62. Variables Common to All Projects

Variable	Infants	Toddlers	Schoolers	Adults
Height or Length *	+	+	+	+
Weight	+	+	+	+
Skin Folds				
Triceps	+	+	+	+
Biceps	+	+	+	+
Arm Circumference	+	+	+	
Head Circumference	+	+	+	

* Supine length was recorded to age 30 months and standing height after that age.

In the analyses conducted and reported to date, the variables used most frequently have been weight and height. Consideration of these variables is emphasized in the present chapter.

2. Reliability of Measures

Each of the projects used paired observers (for at least a predetermined proportion of measurements) as a method of quality control in the collection of anthropometric data. Although data are recorded for these paired observations, inspection suggests that observers may have checked in the field to ensure that observations agreed, and repeated measurements as necessary to ensure there was agreement, before entering the results. That is not a criticism of the effectiveness of quality control; it is a comment on the futility of any examination of recorded data pairs for the estimation of measurement error. However, in Phase 1 or early in Phase 2, projects implemented controlled protocols to examine the reliability of the weight measurements in toddlers. These data are informative and of use to other investigators. Summaries are presented below.

- A more detailed report of measurement error and day-to-day variation in a series of anthropometric measures in several classes of subjects is reported by the Egypt project (See EFR, chapter 4.II.F). As might be expected skin-fold measures exhibit very low reliability (high measurement error and day-to-day variation) in comparison to weight and height measures. An obvious implication is that skin fold measures and variables derived from skin folds (e.g. body fat, lean body mass) have appreciably lower reliability than weight and hence, when used in bivariate and multivariate analyses, are much more

likely to result in attenuated relationships. This warrants consideration in interpreting analyses.

Table 63. Reliability of Weight Measures in Toddlers - Egypt Project

Source of Variance	Variance	Standard Deviation	Coefficient of Variation
	(kg^2)	(kg)	(%)
Between Measurements, same day	0.0027	0.052	0.46
Between Proximate Days, within-person	0.0479	0.219	1.94
Between Subjects	1.9328	1.390	12.29

Analyses provided by Egypt Project. Design involved 14 subjects weighed by 2 persons on each of 3 alternate days. Global mean weight = 11.31 kg

The Kenya project carried out reliability studies for weight measurements in infants, toddlers and lead females. Table 64 presents the results for toddlers. In the case of infants, the CV of the measurement error was 0.57% and for the within-person day-to-day variation it was 5.68%. Adult female figures were close to these for toddlers (0.07% and 2.33%).

Table 64. Reliability of Weight Measures in Toddlers - Kenya Project

Source of Variance	Variance	Standard Deviation	Coefficient of Variation
	(kg^2)	(kg)	(%)
Between Measurements, same day	0.0011	0.033	0.29
Between Proximate Days, within-person	0.0313	0.177	1.55
Between Subjects	1.0606	1.030	9.03

Analyses provided by Kenya Project. Design involved 12 subjects weighed by 2 persons over 6 or 7 days. Total observations = 72. Mean Weight = 11.40 kg.

Final data sets for adult males from all three projects were screened at Berkeley to determine the proportion of second observations with height lower than the first observation as well as the deviation between two observations (protocols called for one measurement per year (at least two per person; only the first two were used if more than two existed in the file). This served as a very crude index of measurement validity. The results are presented in Table 65. For infants, toddlers, and schoolers, where increases in length measures were expected over time, the data were screened to determine the proportion of measures that were lower than the previous measure (negative deviations between measures). These results are presented in Table 66. No

tests were applied to data for lead females since this would necessitate control for changes associated with pregnancy.

Table 65. Indices of Height Measurement Validity, Adult Males

Country Project	Number of Measurement Pairs	Difference Mean	Difference S.D.	Difference S.E.	Proportion Negative Deviations*
		cm			%
Egypt	37	0.7	2.3	0.35	18.9
Kenya	87	0.0	4.0	0.38	21.8
Mexico	142	0.1	2.1	0.18	48.0

* Proportion with deviation greater than -1.0 cm

Table 66. Indices of Length or Height Measurement Validity in Infants, Toddlers and Schoolers

Country Project	Proportion of Negative Deviations* in Successive Measurements (%)		
	Infants	Toddlers	Schoolers
Egypt	7.6% (132)**	22.5% (726)	5.1% (365)
Kenya	16.9% (666)	17.7% (1037)	5.6% (1129)
Mexico	5.5% (672)	6.1% (520)	1.9% (522)

* Proportion with deviation greater than -0.6 cm
** Number of pairs of measures examined

From these two tables it can be accepted that there were measurement errors in all projects. Since the interval between measurements in infants and children (and hence the expected magnitude of positive deviation) was not controlled, comparisons between projects should not be made. It is clear from Table 66 that correct measurement of length in infants and toddlers is particularly difficult. These crude indices suggest that while height is useful in cross sectional comparisons, measurements may not be sufficiently reliable for use as a variable in longitudinal analyses. Although the measurement error impacts on such calculations as weight-for-height estimates, the error expressed as a percentage of total height is likely to be acceptably low; it is when height changes are of interest that reliability is likely to be a serious problem.

3. Comparison of Population Measures: Variance in Study Populations

The following tables have been based on all data archived by projects without exclusion of subjects not in the longitudinal samples. They are cross-sectional descriptions. There may be outliers remaining in the data. The

present discussion focuses upon two matters of importance in interpretation of analyses: do the anthropometric indices suggest that physical growth and development in the study populations has been influenced by environmental variables (and hence that a search for these variables is justified) and is there substantial variance within the populations (and hence is there a basis for hoping to detect associations between anthropometric variables and environmental influences).

Table 67 presents infant descriptors. Tables 68-70 present summary data (drawn from Project Reports) on the weights and heights of toddlers 18-30 months of age. Information for schoolers is presented in Table 71 and for adults in Table 72.

Table 67. Infant Weights and Lengths at Birth, Age 3 and 6 Months
Sexes Combined

		Weight				Length			
Project (Sex)	Age	N	Mean	Standard Deviation	% of Median*	N	Mean	Standard Deviation	% of Median*
			kg	kg			cm	cm	
Egypt									
(Male)	Birth	45	3.5	.38	107.1	45	50.3	1.59	99.6
	3 m	48	5.8	.76	96.9	48	59.5	2.08	97.4
	6 m	44	6.8	1.13	87.9	44	65.5	3.70	96.6
(Female)	Birth	44	3.4	.45	105.3	44	50.9	2.10	102.1
	3 m	39	5.3	.71	98.2	39	58.8	2.64	98.7
	6 m	35	6.4	1.10	88.8	35	64.5	1.66	97.8
Kenya									
(Male)	Birth	66	3.3	0.44	101.0	44	49.7	2.53	98.4
	3 m	70	5.9	0.75	98.6	70	58.2	2.23	95.2
	6 m	60	7.1	0.88	90.5	60	62.7	2.74	92.5
(Female)	Birth	56	3.2	0.43	99.1	40	48.6	2.51	97.5
	3 m	51	5.5	0.94	101.9	51	57.1	2.66	95.9
	6 m	50	6.8	1.09	94.3	50	61.7	2.68	93.6
Mexico									
(Male)	Birth	47	3.2	0.35	97.9	37	50.0	2.23	99.0
	3 m	51	5.8	0.84	96.9	51	58.8	2.36	96.2
	6 m	54	7.2	1.14	91.8	54	64.6	2.74	95.3
(Female)	Birth	44	3.1	0.45	96.0	41	48.9	2.10	98.1
	3 m	54	5.2	0.62	96.3	54	65.6	2.12	95.0
	6 m	42	6.5	0.92	90.2	41	62.2	2.21	94.3

SOURCE: Derived by ME from Cross Sectional Data Sets.

* Mean value for individual measures expressed as % of the median weight or height for a reference North American population (NCHS)(17)

Table 68. Weights and Lengths of Toddlers by Age - Egypt Project

Age in Months	Males N	Males Mean	Males Standard Deviation	Males % of Median	Females N	Females Mean	Females Standard Deviation	Females % of Median
				A. WEIGHT (kg)				
18	53	10.25	1.17	89.4	43	9.58	1.63	88.5
19	37	10.17	1.19	87.4	27	9.86	1.15	89.6
20	38	10.87	1.60	91.7	37	9.83	1.44	87.8
21	34	10.80	1.57	89.7	39	9.99	1.29	87.9
22	32	11.15	1.24	91.2	38	10.27	1.42	88.9
23	41	11.20	1.61	90.3	37	10.38	1.48	88.5
24	31	11.46	1.64	92.9	32	10.83	1.07	91.8
25	36	11.53	1.22	91.9	34	11.09	1.22	92.3
26	38	11.89	1.42	93.3	39	11.21	1.22	91.7
27	28	12.23	1.96	94.5	33	11.41	1.58	91.8
28	44	12.26	1.28	93.3	34	11.61	1.27	91.9
29	44	12.58	1.76	94.4	43	11.99	1.36	93.4
				B. LENGTH (cm)*				
18	50	76.8	3.10	93.2	41	75.0	3.70	92.7
19	35	77.2	3.17	92.6	21	77.0	2.78	94.0
20	37	78.8	3.69	93.5	32	77.5	4.11	93.5
21	31	78.7	3.85	92.5	37	77.9	3.46	93.0
22	33	80.0	4.09	93.0	35	77.9	4.05	92.0
23	42	80.3	3.66	92.5	36	79.1	3.41	92.4
24	30	80.8	3.61	94.4	30	79.3	2.94	93.9
25	35	82.0	3.33	94.9	34	80.2	3.93	94.0
26	35	81.7	3.68	93.6	39	80.8	3.82	93.7
27	26	82.6	4.02	93.8	32	81.3	3.84	93.4
28	42	83.7	3.32	94.2	31	81.8	3.32	93.1
29	42	84.2	3.50	93.9	41	83.0	3.59	93.6

AGE 18 m = 18.0 to 18.99 Source: Computed by ME from Cross Sectional Data Sets.

* Note that reference standard is discontinuous at 24 months when the data base changes from supine length to standing height. Project data are based on supine length at all ages.

Table 69. Weights and Lengths of Toddlers by Age - Kenya Project

Age in Months	Males N	Males Mean	Males Standard Deviation	Males Percent Median	Females N	Females Mean	Females Standard Deviation	Females Percent Median
				A. WEIGHT (kg)				
18	40	9.41	1.00	82.1	51	9.25	1.20	85.4
19	45	9.55	1.05	81.9	44	9.13	1.11	82.9
20	46	9.78	1.03	82.5	53	9.26	1.15	82.8
21	50	10.10	1.03	83.9	52	9.60	1.26	84.4
22	49	10.18	0.99	83.3	50	9.49	1.05	82.2
23	50	10.46	1.06	84.3	62	9.80	1.37	83.6
24	53	10.51	1.16	85.2	54	9.96	0.90	84.4
25	49	10.66	1.12	85.0	55	10.32	1.36	85.9
26	51	10.86	1.10	85.2	48	10.40	1.22	85.1
27	51	10.83	1.14	83.7	55	10.17	1.09	81.8
28	48	11.21	1.11	85.3	42	11.05	1.30	87.5
29	50	11.35	1.12	85.2	57	10.74	1.31	83.7
30	43	11.73	1.21	86.8	49	11.12	1.12	85.4
				B. LENGTH (cm)*				
18	40	75.7	3.49	91.9	50	75.0	3.01	92.7
19	45	76.3	3.26	91.6	43	75.7	2.81	92.4
20	46	76.9	3.33	91.3	53	75.8	2.75	91.5
21	50	77.9	3.53	91.5	51	77.0	2.93	91.9
22	49	78.2	3.59	90.9	49	77.4	2.99	91.4
23	50	79.7	3.59	91.8	62	77.6	2.93	90.6
24	53	79.8	3.61	93.2	54	78.7	2.83	93.2
25	49	80.4	3.68	93.0	54	79.4	3.15	93.0
26	51	81.1	3.57	93.0	48	80.8	2.98	93.7
27	51	81.3	3.64	92.3	55	80.2	3.28	92.1
28	48	82.2	3.55	92.5	42	81.5	2.96	92.8
29	50	83.4	3.53	93.0	57	82.2	2.96	92.7
30	42	84.0	3.63	92.9	49	83.1	2.88	92.9

AGE 18 m = 18.0 to 18.99 Source: Computed by ME from Cross Sectional Data Sets.

* Note that reference standard is discontinuous at 24 months when the data base changes from supine length to standing height. Project data are based on supine length at all ages.

Table 70. Weights and Lengths of Toddlers by Age - Mexico Project

Age in Months	Males N	Males Mean	Males Standard Deviation	Males Percent Median	Females N	Females Mean	Females Standard Deviation	Females Percent Median
				A. WEIGH (kg)				
18	19	9.44	1.08	82.3	13	9.55	1.23	88.2
19	27	9.78	0.88	83.9	29	9.69	1.33	88.0
20	29	10.09	1.21	85.2	33	10.02	1.28	89.5
21	37	10.04	1.02	83.4	41	9.86	1.30	86.7
22	39	9.94	1.16	81.3	38	10.39	1.24	90.0
23	42	10.41	1.07	83.9	44	10.74	1.85	91.6
24	48	10.57	0.89	85.6	53	10.69	1.29	90.6
25	46	10.80	1.16	86.1	47	10.62	1.34	88.4
26	53	10.79	1.20	84.7	52	10.87	1.41	88.9
27	51	11.24	1.16	86.9	54	11.24	1.62	90.4
28	52	11.16	1.17	85.0	47	11.20	1.42	88.6
29	46	11.48	1.24	86.1	48	11.36	1.40	88.5
30	48	11.56	1.16	85.5	40	11.62	1.37	89.2
				B. LENGTH (cm)*				
18	14	74.83	4.25	90.8	8	75.32	3.83	93.1
19	12	76.33	3.64	91.6	11	76.20	3.17	93.0
20	12	76.44	2.46	90.7	9	76.49	2.08	92.3
21	15	77.77	3.37	91.4	14	74.11	3.18	88.4
22	17	77.24	3.31	89.8	12	78.67	3.42	92.9
23	11	80.04	4.54	92.2	9	76.00	3.56	88.8
24	19	80.11	3.71	93.6	12	80.30	4.02	95.0
25	11	80.19	2.75	92.8	13	79.37	4.09	93.0
26	18	79.77	3.37	91.4	9	79.51	3.18	92.2
27	13	82.02	2.64	93.1	8	81.19	5.47	93.3
28	11	81.53	2.27	91.8	14	82.34	4.49	93.7
29	12	83.11	3.36	92.7	8	80.94	4.70	91.3
30	18	83.16	3.60	92.0	10	85.24	3.13	95.3

AGE 18 m = 18.0 to 18.99 Source: Computed by ME from Cross Sectional Data Sets.

* Note that reference standard is discontinuous at 24 months when the data base changes from supine length to standing height. Project data are based on supine length at all ages.

Table 71. Schooler Weights and Heights
(At Entry)

Project (Sex)	Mean Age	Weight N	Weight Mean	Weight Standard Deviation	Weight % of Median*	Height N	Height Mean	Height Standard Deviation	Height % of Median*
Egypt									
(Male)	7.7	53	23.7	7.6	97.0	53	120.5	9.7	96.2
(Female)	7.9	49	21.9	3.1	89.1	50	117.9	6.2	93.6
Kenya									
(Male)	7.0	93	19.2	2.7	84.0	91	113.2	5.7	93.0
(Female)	7.1	76	19.3	2.8	87.5	75	112.3	6.5	92.8
Mexico									
(Male)	7.7	93	21.8	3.0	89.2	89	118.1	6.2	94.3
(Female)	7.8	89	20.3	3.0	83.5	82	115.8	6.3	92.3

Source: Computed by ME from Cross Sectional Data Sets.

* Median weight or height is taken as median value for the average age.

Table 72. Adult Weights and Heights
(Females Not Categorized by Physiologic State)

Project (Sex)	Mean Age	Weight N	Weight Mean	Weight Standard Deviation	Height N	Height Mean	Height Standard Deviation
	yrs		kg	kg		cm	cm
Egypt							
(Male)	35.3	203	71.6	36.8	91	167.3	6.29
(Female)	29.0	259	64.4	11.5	107	154.4	11.8
Kenya							
(Male)	36.8	247	56.1	6.4	245	165.7	6.4
(Female)	31.4	288	51.7	5.9	283	154.6	5.7
Mexico							
(Male)	37.2	211	66.0	9.8	211	166.0	5.8
(Female)	31.6	268	58.3	9.2	208	152.8	5.4

Source: Computed by ME from Cross Sectional Data Sets.

In many regards, the anthropometric data of Kenya and Mexico are remarkably similar. In both cases, the data suggest that growth rates have been depressed, presumably by environmental factors operating early in life. For both study populations, older children and adults tend to be thin, suggesting continuing environmental constraint. Birth weights in both of these project populations are similar and low. By 6 months of age, body weight and length lag behind reference standards. The depression of infant growth is perhaps not as much as might be suggested in the direct comparison of observed values and NCHS reference standards. The reference weights are drawn largely or entirely from formula-fed infants. There are now available a number of studies establishing that during or after the first few months of life, breast fed infants exhibit lower body weights than do formula fed infants in industrialized countries (19,20). If the observed infant weights were referred to breast fed reference populations, the apparent deviation would be reduced. Nevertheless, such a comparison would not alter the general conclusion: in Mexico and Kenya, the populations under study exhibit the characteristics that have been associated with an adverse environment during early growth and development and a continuing constrained environment.

In the case of Egypt, the picture is somewhat different. Although birth weight is somewhat higher than in Mexico and Kenya, by 6 months there is again the suggestion of faltering. By 18 months of age, growth faltering is as great in Egypt as in Mexico and Kenya, but weights are greater. Whatever may have been the factors that influenced early growth and physical development, in Egypt the continuing constraint on weight gain appears to have been relieved. Comparison of Egyptian weights for height with international standards (not presented in tables) suggest that the Egyptian children are very close to expected levels.

Among schoolers and adults, heights are generally comparable across projects (perhaps somewhat taller in Egypt). However, Egyptians are appreciably heavier than Kenyans or Mexicans, again suggesting that conditions affecting weight gain and retention differ across projects.

The variabilities of height are surprisingly similar across projects (see Standard Deviations presented in Tables 67-72), except in Kenyan females who seem more homogeneous. In general, weights appear to have similar standard deviations in Kenya and Mexico and a slightly higher variability in Egypt. This may suggest that both overweight and underweight (expressed in terms of weight-for-height) are found in Kalama while in Kenya and Mexico, people are more similar in their degree of thinness. In turn, this may be suggestive of a more uniform constraint in food intake within the Kenya and Mexico populations; while in Egypt, either the energy intake available for deposition as body tissues is more variable or factors other than food intake may be primary determinants of achieved body size.

Such a conclusion seems surprising given that the Kenya study population underwent a period of major food constraint. It might be expected that this would have resulted in either or both of a more deviant body length and height or at least an increased variability in comparison to the other two projects. It is surprising that this is not readily apparent in the tabulated weights and heights or their variances. To explore this phenomenon a little further,

weights and heights of toddlers at increasing ages were examined within three-month age cohorts. In this way, it is possible to examine the effects of the period of food deprivation on toddlers entering this period at different ages. The result of this examination, shown graphically in Figures 8 and 9, was not dramatic. There were no grossly apparent impacts of food constraint on either the pattern of mean weight across age or in the variability of weight at a given age. This finding should be contrasted with the paired analyses of weights in schoolers and in adult men and women (see RMR, section VI.E). In those groups, weight difference patterns strongly suggested effects of food constraint. The absence of readily apparent effects in toddlers may be a physiological phenomenon (body growth and composition are protected from shortfalls of intake with accommodation coming in other sectors of energy utilization) or it may be a social-cultural phenomenon in which the young are protected at the expense of the older. While a study of the processes will be of considerable interest, the data continue to suggest that using anthropometry as an outcome measure may be fraught with problems. As a minimum, much more subtle longitudinal parameters of growth and physical development may have to be used in analyses.

4. Implications for Analysis and Interpretation

Examination of anthropometric data leaves a distinct impression that whatever environmental constraints to physical growth may have operated, they occurred early, probably before 18 months of age, and were not reversed even in a population in which gross food supply appears to have been adequate to sustain the development of mild obesity (Egyptian population). It may be, then, that the most interesting period for examination of effects of environmental factors on physical growth and development is in the one-year age span between the target infants and target toddlers.

This difference may be exploited in inter-project analyses. It seems clear that Kalama presents an adverse environment and that this influences growth in height to almost as great a degree as do the environments in Kenya and Mexico. If food intake *per se* is not as constraining, then comparisons among the countries may serve to dissociate the effects of the general environment from those of food.

Although the Standard Deviation is not a fully adequate descriptor of the distribution of heights and weights, the data presented do seem sufficient to suggest that the spread in weights and heights is less than had been anticipated in the original design. The intent had been to undertake a stratified sampling of the populations to increase the variance within the study populations. The standard deviations achieved are similar to those seen in well nourished populations, but not much greater. Thus, in toddlers, the CV of weights seen in the projects fall in the range of 10-13% (somewhat lower range in Mexico). In reference US population data, the reported CV's are 11-13% for 2 to 2.5 year olds (17). Among schoolers, study population weight CV's are 14-15% in Kenya and Mexico and much greater for males (but not females) in Egypt. In US population data, the observed CV's are 17-22% (17,18). A generally similar comparison holds for lengths and heights. In the study populations, the CV's of toddler lengths were about 4-4.5%; in the US, they were in the same range (17). For schoolers, the values were 5-5.7% in Kenya

FIGURE 8

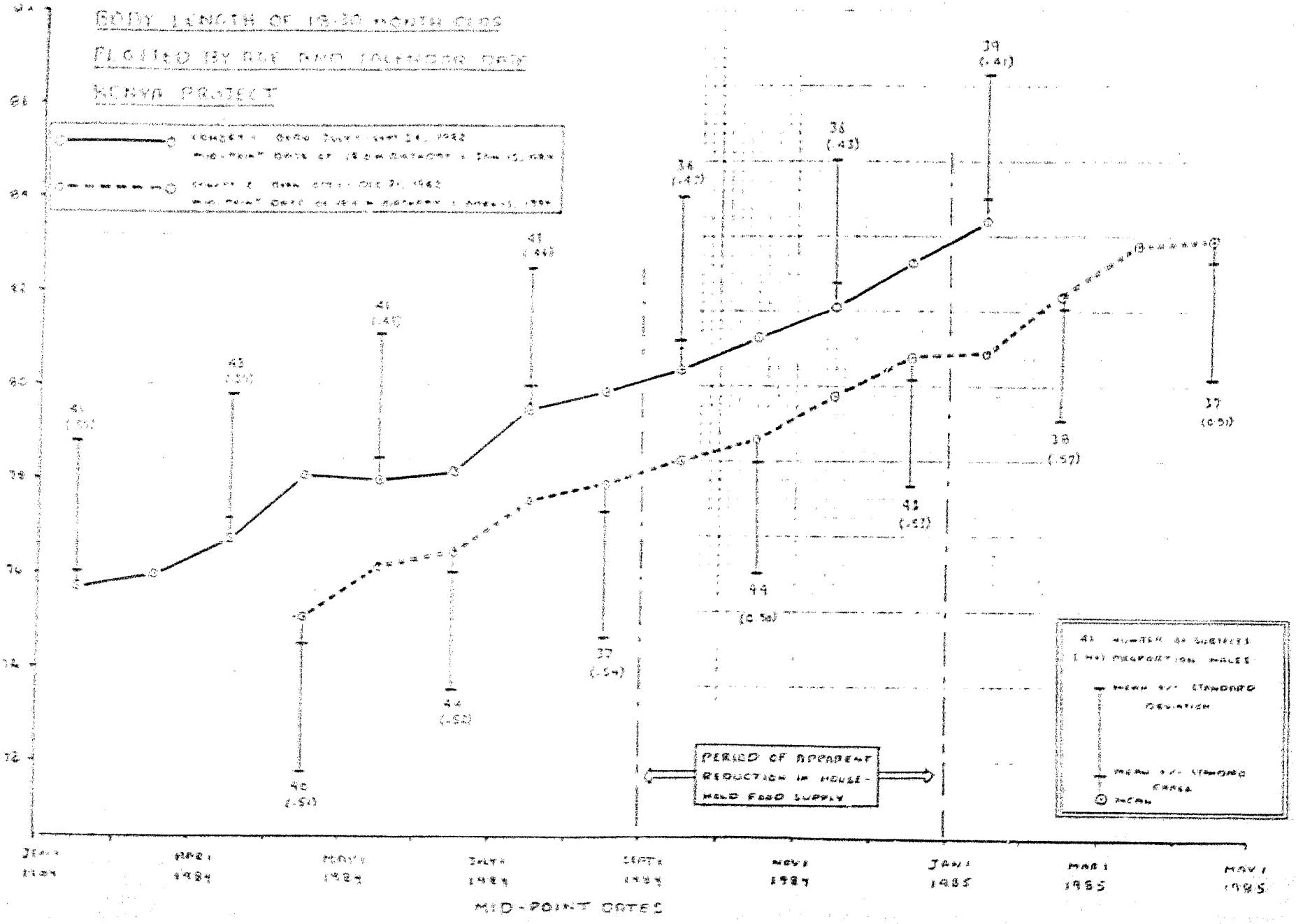
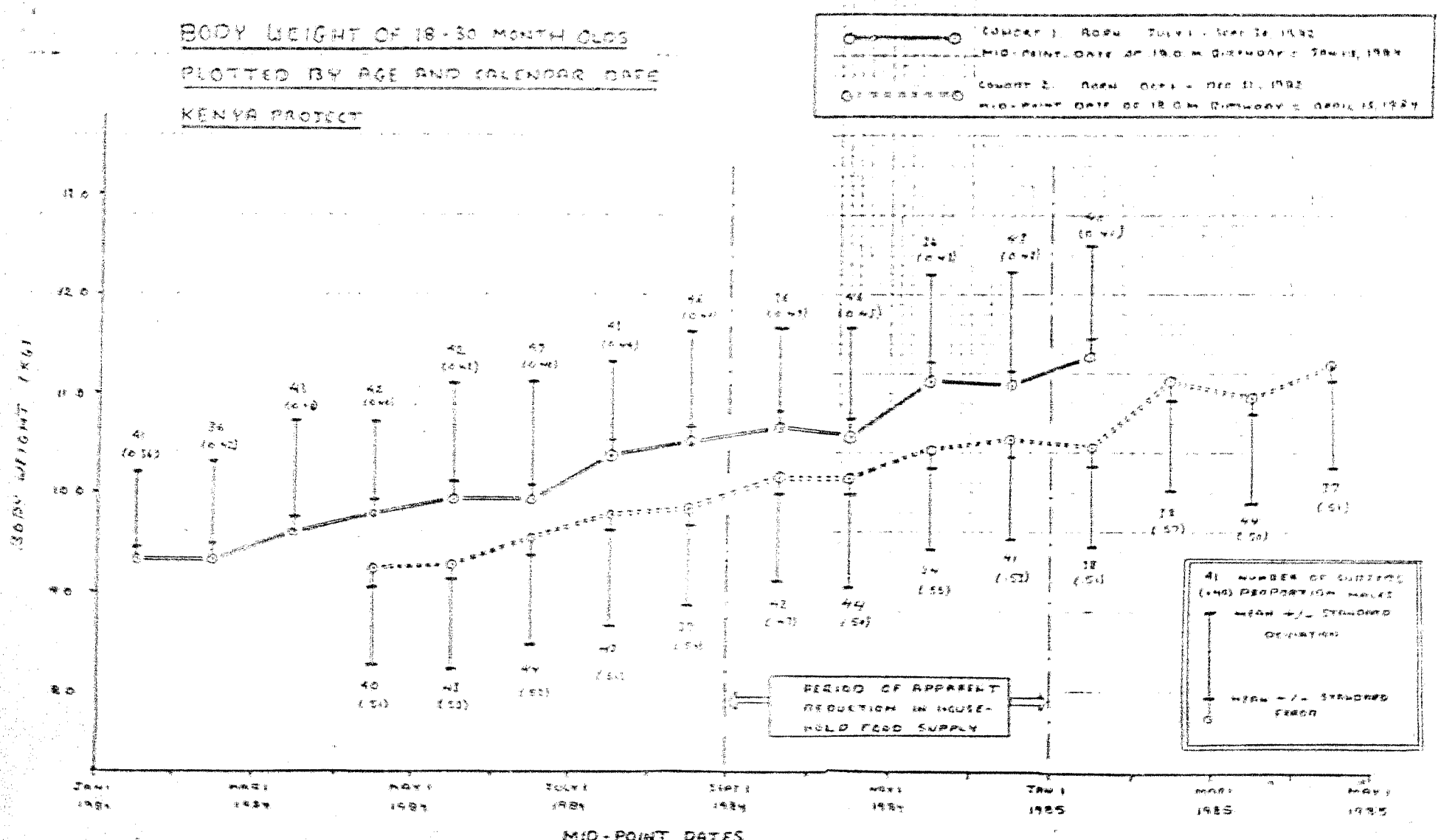

BODY LENGTH OF 18-30 MONTH GIRLS
PLOTTED BY AGE AND INTERVIEW DATE
KENYA PROJECT
PERIOD OF APPARENT REDUCTION IN HOUSEHOLD FOOD SUPPLY
MEAN +/- STANDARD DEVIATION
MEAN +/- STANDARD ERROR
MEAN
LENGTH (CM)
MID-POINT DATES
JAN 1984
APR 1984
MAY 1984
JULY 1984
SEPT 1984
NOV 1984
JAN 1985
MAR 1985
MAY 1985

FIGURE 9


BODY WEIGHT OF 18-30 MONTH GOATS
PLOTTED BY AGE AND CALENDAR DATE
KENYA PROJECT
BODY WEIGHT (KG)
MID-POINT DATES
PERIOD OF APPARENT REDUCTION IN HOUSEHOLD FEED SUPPLY
MEAN +/- STANDARD ERROR

and Mexico and 5.2% in Egyptian females and 8% in males; in the US, the values were 3.8-4.5% (17). It would appear that the variabilities within the study populations are similar to those seen in well-nourished populations (weight in older children and adults in Egypt is an exception).

These observations have important implications for analyses of the present data sets. The general picture is that there is not an excessive variation (in comparison to reference populations). Whatever factors constrain growth and development in these study populations seem to operate with relative uniformity across all sectors and individuals within the study populations. Identification of factors that associate with "growth failure" may be difficult because of a limited range of effect. There will be need for caution against false negative conclusions in analyses conducted within single projects.

Conversely, the Egypt project maintains an interesting divergence from the pattern of Kenya and Mexico. In Egypt, there is the suggestion that constraining factors do operate as in Kenya and Mexico (see discussion of height data), but there is also suggestion that variability may be greater among older children in this project than in the other two or than in the US population. This at least suggests that certain interproject comparative analyses should be of interest.

G. Morbidity and Mortality

1. Definitions

Terms that need definition and clarification to improve their appropriate use in the CRSP include: rate, prevalence, incidence, and duration of disease.

a. Prevalence and incidence. A rate is a measure of the rapidity of change associated with a phenomenon, typically defined in terms of units of time. A rate is an instantaneous measure. A useful way of evaluating risk of disease is an average rate, which results from "averaging" instantaneous rates giving,

$$average\ rate = \frac{\delta y}{\int_{t}^{t+\delta t} y(t)dt}.$$

where, change in a continuous variable $y = \delta y$ and change in time $x = \delta x$.

This technical definition leads to the fundamental difference between a rate and other measures of mortality and morbidity. That is, a rate explicitly takes into account the amount of time at risk by the population under study. If time is not part of the calculation, then the quantity is not a rate. Many quantities computed from data that are not rates are, however, called "rates" by tradition. The "prevalence rate", for example, is almost always the proportion of persons ill at a given time, and the term "prevalence proportion" is not common but is correct. In this report, only the term "prevalence" is used. There is a clear distinction between this and "incidence" which is the rate of new episodes in a defined period of time. Prevalence is the product of incidence and duration provided that either all illness or specific illness is considered.

b. Duration and severity. As a descriptor of a disease process in a particular individual, duration, like "severity", may be a useful variable; this is distinct from the consideration of measures of risk. The duration of a disease is, however, a poor substitute for a measure of risk.[1] There are, as well, clear difficulties to be considered in interpreting duration. A number of intervening variables can affect duration of an illness. Type and effectiveness of therapy is an obvious example; it affects both duration and

---

[1] A simple example illustrates the problem with interpreting average duration of a disease. Suppose that community A and community B each has 100 members, and that in A, 10 individuals are ill for 10 days each, while in community B, two individuals are ill for 50 days each. Both communities have the same number of person-days of illness, namely 100 person-days. If the average duration is related to the number of persons ill, then for community A, the mean is 10 days and for community B, it is 50 days. It would not be meaningful to conclude that community A is "better off" than community B. Alternatively, if the mean is related to the population of each community, then the average duration is one day in both. Again, it is not useful to view these two communities as having equivalent morbidity experience.

severity. Therapeutic regimens vary culturally and access is likely to vary with SES, among other factors. The occurrence of other illnesses can seriously confound the measure. It is clearly misleading to consider the sum of durations of specific, potentially overlapping, illnesses as a measure of the duration of total illness. Average duration is related to incidence of illness, but its magnitude depends on intervening variables such as the prevalence of illness. This lack of a direct relationship with incidence (risk) makes duration a potentially misleading measure of morbidity. Caution is required in interpreting the values reported.

2. Individual Morbidity as a Variable.

a. General considerations. All projects collected morbidity information for target individuals through a weekly interview procedure. (Table 72) The procedures made provisions for the collection of descriptions of morbidity which, through an interpretive classification, permitted the definition of diseases. Although local disease terminology was used initially, all projects have provided recoding procedures adopting international disease codes. The procedures employed also permitted the tracking of morbidity episodes so that durations of episodes could be estimated. This last step required that weekly records be joined to give continuing information about the same individual. Each project responded to this need by preparing monthly (28 day) morbidity summaries. In data analyses, obtaining estimates of incidence or duration continues to require the linking of sequential monthly records when a disease continues beyond the end of the period covered by the monthly summary. A major challenge existed in the development of procedures and project-specific conventions for dealing with missing data (non-administration of weekly instruments), and hence for estimating both the duration of diseases that started or ended during the period (usually done by a probe recall on the occasion of the next visit), and for ascertaining whether or not any short duration diseases occurred in the interval.

Projects have archived morbidity data in differing formats, which provide differing possibilities for analysis. Only the Egypt project has specifically distinguished between days not observed and days observed with no morbidity, in the summary records archived. This becomes a major issue when attempting to generate analytical variables relating to "wellness", or incidence or prevalence data.

Only the Mexico project has attempted to archive the details of signs or symptoms actually recorded in the field and used to generate diagnosis codes. For Egypt and Kenya, derived "summaries" have been archived. Future researchers who wish to regenerate morbidity variables using different operational definitions of morbidity, or wishing other detailed information, will need to consult original (paper) field forms filed only at the level of the individual projects.

Some very important issues exist with reference to any interproject comparisons of morbidity incidence, prevalence or duration. The first is a basic difference in the field procedures, and hence the nature of data collected, between Mexico and the other two projects.

Table 73
Planned Morbidity Protocol

Subject	Weekly Morbidity Recall[a]	Procedure following illness report (Days 1,3,7,14 and weekly until Recovery): Food Intake	Qualitative Activity Index[b]	Body Weight	Task Assignment[c]
		BASIC DATA SET[d]			
Non-target household members	x				
Lead Male	x		x		x
Lead Female	x		x		x
Toddler	x[e]	24-hour recall by care-giver	x	x	
Infant	x				
		BASIC SUPPLEMENT[d]			
Schooler	x				
		COMMON DATA SET[d]			
Lead Male		Qualitative[f]			
Lead Female		Qualitative			
Schooler		Qualitative	x		
Infant		24-hour recall by care-giver	x	x	

[a] Structured questions: Is anyone sick today? Has anyone been sick during the past week?

[b] Qualitative Activity index: 1=no change; 2=moderate reduction; 3=confined to bed; 4=critically ill.

[c] Added variables after Qualitative Activity Index, as appropriate to project, for example: Did you work outside of home (employment) since illness occurred or last visit? 1=no change 2=less hours/vigor 3=unable to work; Did you cook meals? if 2 or 3: Did anyone also do/help with your work because you were ill? Who?

[d] The Basic Data Set was a minimum designated by ME, to be supplemented with prescribed data on schoolers if projects chose to include this target group. The Common Data Set was agreed upon by two or three projects with SCB concurrence.

[e] Added variable for toddler whose diagnosis is diarrhea on day of report, and 1,3,7...days: No. of bowel movements in past 24-hours; stool characteristics.

f Qualitative Food Intake: 1=no change; 2=decreased; 3=caloric liquids, no solids; 4=water, clear fluids only; 5=no intake.

The second is a potential or real difference in the operational criteria by which subjective judgments about the presence or absence of symptoms, and indeed what constituted a symptom, were taken, and hence the nature and exact meaning of data recorded. Both of these classes of difference yield potential biases in reporting across projects. The biases would differ for the two sources. A third concern is variation in illness intervention. These are discussed below.

b. Morbidity protocol difference - Mexico project. The intended CRSP protocol for the collection of morbidity data (Table 72) required that during weekly visits the interviewers ask two questions: "Is anyone sick today?", and "Has anyone been sick during the past week?" The protocols then provided for the solicitation of details about the identified illnesses. This approach was implemented in Egypt and Kenya throughout the study. In Mexico it was implemented only in the last few months of the project. Throughout the data collection period, the Mexico project interviewers asked, "Is anyone ill today?" and then probed for details of the index illness, but did not ask about illness occurring during the week unless it had been present on the previous visit and was being followed for the evolution of symptoms, duration, etc. As was demonstrated by the Mexico project (see MFR, chapter 22), this resulted in a bias of underestimation of incidence and prevalence in the reporting of short term illness. It may be illustrated by a simple calculation: with an interval of 7 days and an illness duration of 2 days, the likelihood of that illness being seen on the day of the morbidity visit is only 2/7 or .29. It would be expected that some 70% of such illnesses would be unreported with the Mexico project primary methodology. As duration of the illness increases, the proportional under-reporting would decrease. Direct comparisons of the two methods have been offered by the Mexico project for the period when both were in place. A crude adjustment of data has been undertaken in comparisons presented below.

The Mexico project cautions that illness is further underreported in the months of December and January because field staff did not conduct morbidity interviews during the holidays.

c. Illness intervention. The procedure to be followed upon identification of an illness or injury always presents a problem in field research. It is unethical to withhold treatment that it is within a field team's resources to provide, but intervention, particularly with therapies that would ordinarily not be available to the community, may distort the phenomena under study. Projects took decisions based on their separate circumstances, which may have affected detection, duration, and severity of illness.

In Kenya, records were reviewed by trained nurses, and the presence of serious illness was confirmed by a visit from a nurse or physician. These professionals gave first aid, started treatment, or referred the patient to a hospital, providing transportation if necessary. Suspected malaria was treated with chloroquine after taking a blood smear. All subjects having blood hemoglobin values below 9 g/dl were treated with ferrous sulfate. Minor ailments were treated symptomatically with aspirin, decongestants, and antacids. The project expressed concern that the study population became increasingly dependent on project staff for health care. A cursory

examination of morbidity data suggests that morbidity may have been less in the second year of Phase II than in the first year. The effects of food shortage and health intervention may be confounded.

The plan to be followed in Mexico on diagnosis of illness by physicians is not specified in the MFR. The plan was to follow a "decision tree" for each symptom, in order to standardize treatment. The project staffed a few small outreach clinics and supplied them with common pharmaceuticals; they also assisted in community vaccination programs and record-keeping. The report states that "households turned to this [responsible] staff person for help in medical...emergencies...One of the most valued and usual ways that assistance was given was transportation of a seriously ill household member to Mexico City and provision of medical treatment at the INN hospital."

In Egypt, morbidity data were gathered by a cadre of young physicians who also conducted the physical examinations. The project decided to have these professionals provide to their assigned households medical care for "primary care problems" and refer complicated cases to appropriate facilities. Some families used both CRSP-provided and other local health care services. The project developed standardized treatment protocols for commonly-encountered conditions (documented in the Morbidity Field Manual).

d. Differences in operational criteria of illness. Such differences could result from the background training of the interviewers, from community perceptions of illness and hence subject responses to the interviewers, or from the operational criteria developed and standardized within the individual project. It would be difficult to standardize these dimensions across projects at the design stage. It is difficult to determine in retrospect whether there were differences in data between projects attributable to these effects.

In the Egypt project, morbidity information was collected by physicians. In Kenya, the primary data were collected by trained interviewers who had no health training as a background. Coding of illnesses (interpretation of recorded symptoms) was done by a physician. In Mexico, the primary data were collected by a physician and community worker who visited the household together. Coding field data into illness categories was done by the physician.

It is generally recognized that the prior training of the interviewer can affect the data recorded, and indeed that a physician examining a patient may render a different judgment than if he/she were examining only recorded symptoms (perceptions of "illness" and "wellness" are conditioned by experience). Regardless of who actually conducts the field interview, it is quite possible that the recording of an "illness" is contextual. That is, "ill health" may be marked as a departure from "normal health" (the background picture of the community) rather than from "good health" (conceptually a diagnosed situation based on objective criteria). This distinction is implicit in the approach taken in the interviewing. The spontaneous answers to the questions, "Has anyone been ill?" or, "Is anyone ill today?" will be influenced by perceptions of illness in the community. It follows that these perceptions, and the degree of interviewer probing of a negative answer, will influence reported morbidity.

There are suggestions that these factors operated across projects to give differences in the meanings of the morbidity variables, particularly between Egypt and Mexico on the one hand, and Kenya on the other. The Kenya project may have had a more comprehensive system of detecting and reporting signs and symptoms of what may mark illnesses. According to their report, "If, during the probe inspection, an enumerator found a household member to be sick, even though illness had been denied on the probe, this was reported..." and, "If a probe was negative for a given [target individual], yet illness was observed, an individual morbidity report was filled out..." Augmentation of the record by direct observation would have been limited to those individuals present when the interview occurred. These differences might have influenced the Kenya finding of more morbidity and longer duration of illness episodes (see below) than did either of the other projects. Consideration must be given to these potential sources of reporting differences in any cross-project comparisons. It has not been possible to develop objective procedures by which "reporting bias" can be assessed across projects.

e. Estimation of incidence, duration, and severity. Given the differences in protocols, classification approaches, and interviewer backgrounds, direct comparisons of morbidity across projects are tenuous. Nevertheless, in this report, an attempt is made to present crude comparisons for selected classes of morbidity. These are: diarrhea, and upper and lower respiratory tract illness in infants, toddlers, and schoolers in longitudinal windows, as defined in section VI.B. The number of days of morbidity observation within these windows is shown for children in Table 74.

The interproject comparisons involved the identification of episodes of illness and estimation of their durations. The approaches taken differed somewhat between projects because the nature of the archived data sets differed.

The basic approach to calculation of average incidence was as follows:

$$\text{Average Incidence per 1000 person days} = \frac{\text{\# of Illness Episodes in window}}{\text{estimated days observed in window}} \times 1000$$

ME has calculated average incidence from the archived data on morbidity for the Egypt and Kenya projects. Occurrences of morbidity were reported by the Mexico Project, summarized by calendar month and by age. That project also reported estimated days observed in summary form. ME calculated "average incidence" using those data bases. Because of differences in the data sets, specific approaches differed between projects as described below.

For toddlers and schoolers, the expected number of observed days of observation would be 365 (1 year). For infants it was 183 days (6 months). The actual period of observation would reflect the periods when morbidity data were not collected (see Table 74). A person had to be free of symptoms for two days before a reoccurrence of symptoms of the same illness category was counted as a new illness episode. Other illnesses could overlap.

TABLE 74

Estimated Number Of Days Observed Within Longitudinal Windows*

Target Type	N	Mean	SD	Minimum	Maximum
		EGYPT PROJECT			
Infants	54	122	33	35	168
Toddlers	83	276	61	49	357
Schoolers	41	249	79	42	350
		KENYA PROJECT **			
Infants	93	183	-	183	183
Toddlers	91	365	-	365	365
Schoolers	129	365	-	365	365
		MEXICO PROJECT			
Infants	69	139	24	29	179
Toddlers	55	306	24	226	340
Schoolers	109	305	33	145	340

* Length of longitudinal windows:
Infants - 183 days
Toddlers, Schoolers - 365 days

** Missing observations could not be identified in the archived Kenya project data. In consultation with project staff, it was judged that relatively few morbidity recalls were missed. For purpose of calculations, the duration of observation was set as equal to the length of the window.

An example of the approach to determination of the number of illness episodes within the longitudinal window as carried out for Egypt and Kenya is given in Table 75. In the calender-month analysis utilized by all projects for investigation of seasonal morbidity patterns, a continuous illness which spans two months shows up as two "Illness Episodes".

Table 75

Identification of Episodes of Morbidity: ME Approach

Sample for Subject 209342 in the Kenya Data Set for the Window Age 18 months (16 May, 1984) to 30 months (16 May, 1985)

SYMPTOM CODE	STARTDATE	ENDDATE	EPISODE
501	15MAY84	18MAY84	
501	20MAY84	23MAY84	1
501	04APR84	16APR84	2
501	10JUL84	17JUL84	
507	18JUL84	28JUL84	3
507	01NOV84	10NOV84	
502	08NOV84	21NOV84	4
501	05MAY85	10MAY85	5
501	13MAY85	20MAY85	6

This subject was taken as having 6 upper respiratory illness episodes within the longitudinal window.

Some differences in aggregation and classification for interproject comparison purposes are accepted as unavoidable because of project differences in methodology. Approaches used by ME to improve comparability are outlined below.

Specific Approaches, Egypt Project. The number of illness episodes within the longitudinal windows was derived from the file "Illness Episodes". The following symptom codes were collapsed into the three illness categories:

Diarrhea	702 = diarrhea
Upper Respiratory	501 = common cold 503 = tonsilopharyngitis 504 = croup/laryngotracheitis
Lower Respiratory	505 = bronchitis 506 = asthma/wheezing 507 = pneumonia 508 = bronchiolitis

"Estimated days observed" was derived from the Weekly Morbidity Recall file. According to the Egypt Project, absence of data in records indicates that the subject was not observed. Such records have subsequently been eliminated. The remaining records within the longitudinal windows have been counted and the variable "estimated days observed" has been calculated as:

estimated days observed = # of records in longitudinal window x 7

The Recall file gives the information DATE1ST = beginning of observation, and DATE2ND = end of observation for each recall visit. Theoretically an exact count of the days observed would be possible (see example in Table 75). However, since there is often an overlap of 1 to 2 days from the end of one observation week and the beginning of the next, the time and effort needed to eliminate the overlapping days from the count can not be justified.

One toddler has been excluded from the analysis because of errors in the data collection.

Specific Approaches, Kenya Project: The number of illness episodes within the longitudinal windows was derived from the Morbidity Episodes File (data recording forms #313 and #314). The following symptom codes were collapsed into the three illness categories:

Diarrhea	702 = diarrhea
Upper Respiratory	501 = common cold
	502 = tonsilopharyngitis
	503 = croup/laryngotracheitis
	507 = acute upper resp. tract infections
Lower Respiratory	504 = bronchitis
	505 = asthma/wheezing
	506 = pneumonia
	508 = bronchiolitis

The Kenya Household Morbidity File (data recording form #315), unlike the Egypt Weekly Morbidity Recall File, was found to be unsuitable for the calculation of "Estimated days observed". If a record for a household in a 28-day-period exists, it is not known which members of the household were actually present. Nor is it known whether the household was observed for the whole 28-day-period or just a part of it. In addition, the file has missing data for the months from July 1984 to October 1984 for all households with household ID from 1000 to 1999. However, illness episodes for individual members of those households in that time period have been reported in the Morbidity Episodes File. These factors prevented direct estimation of number of days observed. For the present calculations, the assumption was made that no observations were missing as long as the month was included. While this may overestimate the period of observation, it is thought that very few morbidity recalls were uncollected; the resultant underestimation of average incidence per 1000 person days is likely to be very small.

Specific Approaches, Mexico Project: The average incidences for diarrhea and upper and lower respiratory tract infections have been calculated by using the variables "Number of diarrhea episodes", "Number of upper respiratory illnesses", "Number of runny nose episodes", "Number of lower respiratory illnesses" and "Estimated days observed in age-month" from the file "Mexico Method Morbidity, summarized by age-month". The variables "Number of upper

respiratory illnesses" and "Number of runny nose episodes" were added and the total was used for calculation of the average incidence of upper respiratory tract illness. This approach does not preclude the possibility of overlapping disease episodes.

According to the Mexico Project, the following symptom codes have been collapsed into the four illness categories (diarrhea, UR, LR, runny nose) in the Mexico Morbidity Summary files. Because the Mexico symptom codes vary substantially from the two other projects, the frequencies of symptoms among all subjects during the whole study period (recorded with the "Mexico Method") are shown in Table 76. For interproject comparability, ME has included the category "Runny Nose" with "Upper Respiratory" in the tables that follow.

Chapter 22 of the MFR gives an explanation of the calculation of "Estimated days observed":

> Estimated days observed (calender month basis) was derived as follows: In cases where either 4 or 5 visits were scheduled in a month, estimated days observed were calculated as:
>
> $$\frac{\text{Number of Interviews Completed in the Month}}{\text{Number of Interviews Scheduled}} \times \text{Number of Days in the Month}$$
>
> with the result rounded to the integer.
>
> In cases where 3 or fewer visits were scheduled, the number of days observed were estimated to be:
>
> Number of visits completed in the month X 7

As noted above, the Mexico morbidity methodology can be expected to underestimate the prevalence and incidence of diseases having a duration of less than 7 days. For purposes of this report, an adjusted incidence rate has been calculated as

$$\text{Adjusted Incidence} = \text{Reported Incidence by Mexico Protocol} \times \frac{\text{Estimated Average Duration}}{7}$$

Table 76
Illness Symptom Codes and Frequency of Reporting:
Mexico Project

Illness Category	Symptom	Reported Frequency
Diarrhea	07 = nausea	9
	17 = diarrhea	390
	20 = vomiting	59
	62 = dysentery	1
	64 = tenesmus	0
	68 = dehydration	0
	90 = hyperperistalsis	0
Upper Respiratory	03 = pharyngotonsillitis, inflamed throat	167
	11 = sore throat	135
	15 = sore throat	10
	27 = dysphonia	42
	31 = nasal congestion, stuffy nose	10
	33 = phlegm	0
	36 = garraspera	13
	37 = nasal bleeding	15
	60 = odynophagia, painful swallowing	0
	63 = reddened pharynx	0
	66 = reddened tonsils	0
	67 = purulent tonsils	0
	75 = dry cough	1
	76 = productive cough	2
	78 = purulent tonsils	0
	89 = difficulty in swallowing	0
	95 = enlarged tonsils	0
Runny Nose	01 = flu, common cold	1379
	73 = clear nasal discharge	0
	74 = mucous nasal discharge	0
Lower Respiratory	13 = cough	886
	42 = chest pain	39
	57 = bronchitis	9
	71 = rhonchi	6
	79 = abnormal lung sounds	2
	80 = lung stertorous	0

As is demonstrated in Table 77, the Mexico protocol yields an overestimate of the duration of illness (seen in all classes except adults). This is the direct result of under-detection of short episodes of illness. It would be inappropriate to use the Mexico protocol data base to estimate duration for purposes of the above adjustment. Data obtained using the CRSP protocol applied in the last few months of the Mexico project would be preferable, but the constraint is in the limited volume of data available (implied wide confidence intervals on the estimate of mean duration). Expected durations used in calculation were judgmental based on a review of durations observed with the CRSP protocol in Mexico (Table 77) and the Egypt project data base (Table 78). The judgmental reference durations for diarrhea, upper respiratory infection, and lower respiratory infection, in days, were: schoolers 5.0, 4.2, 5.5; toddlers 4.3, 5.0, 6.2; infants, 5.5, 5.2, 6.2. These durations were applied in adjusting the Mexican data in Table 79.

Table 77.

Estimated Duration of Specified Classes of Illness, Mexico Project: Comparison of Data Collected by Two Protocols *
(days per episode)

Class of Target	Mexico Project Protocol **			CRSP Protocol ***		
	Diarrhea	Upper Resp	Lower Resp	Diarrhea	Upper Resp	Lower Resp
Adults	5.3	6.3	6.7	5.9	7.1	6.4
Schoolers	6.6	5.4	5.9	4.0	4.0	5.2
Toddlers	5.4	6.2	6.2	4.1	4.0	6.1
Infants	6.4	4.9	6.5	5.6	5.1	6.3

Data Source: Analysis of archived files, Berkeley

*Definitions of diseases set by Mexico Project; these were categorical variables in analyzed data sets

**Protocol applied throughout the project

*** Protocol concurrently applied from September 1985 through May 1986; actual data base much smaller than implied by dates since households were progressively withdrawing during this period.

Table 78

Mean Duration of Illness Episodes in Children
Egypt and Kenya Projects[a]
(days per episode)

Target	EGYPT Project				KENYA Project[b]			
	NL[c]	NS[d]	NE[e]	Mean Duration	NL[c]	NS[d]	NE[e]	Mean Duration
A. Diarrhea								
Infants	54	32	58	5.5	93	27	36	7.0
Toddlers	83	62	208	4.4	91	61	124	6.1
Schoolers	41	5	10	7.0	129	8	10	4.8
B. Upper Respiratory Tract Illness								
Infants	54	21	33	5.2	93	91	275	18.9
Toddlers	83	49	87	5.9	91	90	547	16.1
Schoolers	41	10	16	4.4	129	103	234	14.2
C. Lower Respiratory Tract Illness								
Infants	54	29	58	6.2	93	12	13	15.2
Toddlers	83	40	85	6.4	91	19	29	11.2
Schoolers	41	6	7	5.7	129	17	10	7.1

[a] The mean durations of the illness episodes were calculated using data from the longitudinal subjects as described earlier. If an illness episode overlapped the boundaries of the longitudinal window, the duration was not truncated. For example, if a child became ill 4 days before the longitudinal window began and continued to be ill for the first 10 days of the longitudinal window, the illness episode was counted as 14 days rather than 10.

[b] The data set for the Kenya schoolers contained illness episodes with negative duration or missing duration (because of missing startdate or enddate). Those values have been excluded prior to the calculation of the mean durations.

[c] Number of all longitudinal subjects

[d] Number of longitudinal subjects with at least one diarrhea episode within the longitudinal window

[e] Total number of episodes within longitudinal windows

Table 79
Average Incidence of Specified Disease Categories Among Longitudinal Infants and Children in Egypt, Kenya, and Mexico
(Episodes per 1000 person days)

Target Type	Sex	EGYPT PROJECT N	EGYPT PROJECT Average Incidence	KENYA PROJECT N	KENYA PROJECT Average Incidence	MEXICO PROJECT N	MEXICO PROJECT Average Incidence	MEXICO PROJECT Avg. Incid. (adjusted)[a]
				A. Diarrhea				
Infants	M	30	9.9	50	2.6	34	3.6	4.5
	F	24	7.4	43	1.5	35	4.0	5.1
Toddlers	M	41	9.4	42	2.9	29	1.8	2.9
	F	42	8.1	49	4.4	26	3.0	4.9
Schoolers	M	22	0.68	72	0.27	55	0.11	0.15
	F	19	0.99	57	0.14	54	0.19	0.27
				B. Upper Respiratory Tract Illness				
Infants	M	30	5.7	50	15.4	34	6.0	8.1
	F	24	4.6	43	16.9	35	4.8	6.5
Toddlers	M	41	3.8	42	16.1	29	4.3	6.0
	F	42	3.8	49	16.7	26	3.7	5.2
Schoolers	M	22	2.2	72	6.6	55	1.1	1.9
	F	19	0.71	57	4.6	54	1.6	2.6
				C. Lower Respiratory Tract Illness				
Infants	M	30	8.8	50	1.0	34	4.0	4.5
	F	24	8.3	43	0.49	35	2.7	3.0
Toddlers	M	41	3.4	42	0.58	29	3.1	3.6
	F	42	3.8	49	1.1	26	2.4	2.7
Schoolers	M	22	0.96	72	0.61	55	0.47	0.60
	F	19	0.16	57	0.43	54	0.93	1.2

[a] Adjusted to take into consideration the underreporting of short term illnesses in Mexico; see text.

f. Classification of severity. Severity of disease experience also was recognized by all projects to be an important variable, and one that is difficult to quantify. Criteria that entered into a classification of severity were agreed upon, in principle, by all projects, but in the end were treated differently. Illnesses were to be classified by diagnosis as "mild" or "severe" (e.g. common cold is "mild", pneumonia is "severe"), taking into account such factors as age, degree of febrile response, presence of complications, etc. Duration has different clinical significance depending on the disease condition (e.g. some skin diseases have long duration but little metabolic impact). The degree of incapacity was deemed a significant criterion and the protocol (Table 73) included a qualitative assessment of this. Finally, change in food intake is accepted to be an important measure but one that is difficult to deal with analytically as a component of a severity classification when testing hypotheses where both food intake and morbidity severity are variables. In the event, all projects collected qualitative food intake information during illness episodes; none carried out the protocol that called for quantitative records for sick toddlers and infants.

Projects have used these variables in different ways analytically (see project reports). Sometimes the component measures have been used as single variables; sometimes they have been collapsed into an index of severity. In this report, project severity codes were aggregated for all symptoms and all subjects archived in the morbidity files. (See Table 80.)

g. Interproject comparison of morbidity of target children. A striking difference between the Kenya project morbidity data on the one hand and the Egypt and Mexico data on the other hand is seen in Tables 77 and 78. The reported duration of illnesses is much longer in the Kenya project than in either of the others. This could be due to major differences in the nature of illnesses in the different sites (as well as differences in the groupings of illnesses reported in the tables); however, it seems more likely that there were true differences in the operational criteria of illnesses in Kenya. That is, it seems probable that the Kenya project reported symptoms that were accepted as a part of the background in the Egypt and Mexico projects.

There were also differences between projects in the incidence of specified diseases among children. Diarrhea (Table 79) was least frequent in Kenya; incidence was slightly higher in Mexico and several-fold higher in Egypt. The highest incidence reported, for Egypt male infants, is 9.9 episodes/1000 person-days or 3.6 episodes per year. Infants and toddlers in the other two projects were reported to experience one or two episodes per year. By contrast, upper respiratory illness is much more common in Kenya children, two to three times that reported in Mexico and Egypt. In the category lower respiratory illness, Kenya incidence is quite low, Mexico is intermediate, and Egypt is high, with incidence in infants matching that for diarrhea in both Mexico and Egypt data sets. In interpreting these comparisons, it is important to consider the reported durations. Thus, in Kenya, although the incidence of lower respiratory infection is very low, the reported duration is very high. If prevalence is calculated as incidence x duration, one can estimate that in Kenya, toddlers (average of both sexes) are ill with lower respiratory

infection about (0.85 x 11.2) x 100 /1000 = 0.95% of the time. The comparable figures for Egypt and Mexico would be 2.3 and 2.0%. If similar calculations are carried out for upper respiratory infection in toddlers, the results for Kenya, Egypt and Mexico are 26.4, 2.2 and 2.8%. Upper respiratory disease, like diarrhea, appears to show substantial differences across projects with Kenya being extremely different from Egypt and Mexico.

Table 80
Comparison of Frequencies of Application of Severity Codes
by Target Type
(% of reported morbidity episodes)

	EGYPT		KENYA		MEXICO*			
					MEXICO Method		CRSP Method	
Target Type	mild	severe	mild	severe	mild	severe	mild	severe
Lead Males	45.7	54.3	93.7**	6.3**	96.4	3.6	100	0
Lead Females	51.9	48.1	95.5	4.5	95.4	4.6	94.2	5.8
Schoolers	54.2	45.8	94.7**	5.3**	90.8	9.2	100	0
Toddlers	42.5	57.5	94.0	6.0	96.9	3.1	100	0
Infants	29.3	70.7	95.6	4.4	100	0	100	0

* For Mexico, the variable "Degree of Incapacity" was classified into the 2 categories:

- mild: 1=no change
  2=some incapacity
- severe: 3=bedridden

** For Kenya lead males and schoolers the frequency of the severity code was calculated using uncleaned data sets.

Severity classification differs widely among projects, with Kenya and Mexico reporting a much smaller proportion of illnesses classified as severe than Egypt. The Egypt data base codes about half of all illness in target individuals as severe, except infants, where 70.7% is severe. Again it is of interest to do some calculations which might address the incidence of illnesses that become severe. A crude estimate is obtained by multiplying the incidence of illness by the proportion classified as severe. Table 80 does not present data categorized by type of illness, therefore the calculations are tenuous at most. Nevertheless, they serve to emphasize the point made in the introduction to this chapter - interpretation of morbidity data and choice of mode of expression must be done with care. For upper respiratory infection in toddlers, the estimated incidences of "severe" infection for Kenya, Egypt

and Mexico would be 0.98, 2.2 and 0.17 episodes per 1000 days. The difference between projects is much smaller than would have appeared from Table 80 and might well reflect differences in the project definitions of illness and wellness as discussed earlier. That is, if indeed the Kenya project categorized as "ill" upper respiratory conditions that the other projects accepted as "background", the differences in incidence, duration and severity might be explained. This certainly warrants further exploration in later analyses.

As discussed above, these differences may relate to the different background preparations of the projects' interviewers, and hence what they judged to be "illness". It is not possible to say which is "correct". It is absolutely clear that if the definitions of what constitutes an illness differ as much as is implied by the data presented, analyses of relationships between food intake and morbidity will be asking quite different questions across the projects. In the same sense, when morbidity is included in analyses as an explanatory variable, it will have very different meanings across projects. This has very major implications for the cross-project interpretations of analyses involving morbidity. Coupled with the methodological differences and associated underreporting of short term illnesses in Mexico, it would appear that cross-project analytical comparisons remain difficult unless a standardized definition of morbidity can be imposed on all data sets. As a minimum, this would seem to require examination of the raw data that are currently available only in hard copy form at the project level.

3. Food Intake During Illness

In the original research plan (Table 73), episodes of illness in target individuals were to invoke additional study of food intake, body weight, and activity. These measures were intended to answer important questions about the complex interaction between disease, nutritional state, and household function. All projects determined that detailed study of sick persons not to be feasible under field conditions. As noted above, projects did obtain qualitative ratings of food intake and degree of physical disability.

Lack of detailed information on activity, task assignment, and body weight eliminates the possibility of estimating the energy cost of illness. Perhaps more importantly this lack also eliminates the possibility of estimating the extent to which illness of parents affects the welfare of the household through decreased productivity and burden shifting from adults to children (affecting energy needs, school attendance, and the like).

Lack of quantitative information collected purposefully during and following illness means that the cycle of morbidity-recovery and the specific role of food intake in this cycle cannot be addressed in the CRSP data sets. However, within the overall data set it is possible to classify food intake days by the presence or absence of reported morbidity. By such an approach, two questions can be asked: (1) was there a systematic difference in food intake during identified illness; and (2) is there a bias in the food intake data base consequent to over- or underrepresentation of illness days. Both of these questions are examined here.

a. Food intake during morbidity.

Mexico Project. For the period during which the "NCRSP Method" of morbidity data collection was in place (and the final food intake methodology), those days with recorded illness were identified for all toddlers. Twenty-two toddlers were found who had intakes recorded on days recorded as ill. For each of these 22 subjects, all intakes on ill days were pooled. For each subject, all intakes on days not recorded as being ill were pooled. The ill day and non-ill day intakes were then compared with results as shown in Table 81. There is no demonstrable difference in intake between days "reported ill" and "not reported ill".

Table 81.

Comparison of Reported Food Intakes of Toddlers on Days Reported Ill and Days not Reported Ill: Mexico

Measure	Mean	Standard Error	P |T|
Intake on Days Reported Ill	1036	78.44	
Intake on Days Not Reported Ill	976	57.05	
Difference (Not Ill - Ill)	-60.7	64.83	0.3601

Care must be taken in interpreting this result. Specifically, it cannot be concluded with certainty that days not reported ill are indeed well days. From the structure of the primary files, it is not certain that the absence of a morbidity notation marks a well day or a day that was not covered by the questionnaire. This uncertainty is unlikely to have masked any major effect of identified morbidity in relation to the general background condition of the toddlers.

Kenya Project. The Kenya project report presents evidence from qualitative questioning that food intake of all targets combined is frequently decreased in the presence of even mild respiratory disease, febrile episodes and diarrhea (KFR, Table 8.13). Quantitative intake by toddlers is shown to be significantly decreased during fever and digestive illness (by about 150 kcal/d) but not during afebrile respiratory and other illness (KFR, Table 8.15). These numbers agree well with the results reported by Martorell et al (21), who found that intake decreased 160 kcal/d for Guatemalan children with diarrhea, but there was no significant decrease in energy intake during respiratory illnesses.

The contrast with the apparent strong association between morbidity and intake in the Kenya data and the absence of an association in Mexico is unexplained.

Egypt Project. Similar data are not now available from the Egypt project but can be generated. This comparison will be of interest in view of the recorded severity of disease in Egypt.

b. Bias of food intake data sets due to illness.

The question of bias due to illness has been investigated for target toddlers. Records have not been examined for other categories of target individuals. Confirmation of the Mexico and Egypt projects' belief that records were not obtained when the lead female was ill is a priority, because of the significance this may have to an understanding of the transactions in deprived households.

Egypt Project. The question of bias due to illness in the entire toddler data set was examined and reported by the project (EFR, Table 5.16) The basis was as follows: for each toddler, define a pseudo year of 388 days, consisting of 12 pseudo-months of 30 days each and windows of 14 days on each side. First, classify each day as sick (from the morbidity episode file), not sick (from the recall file) or don't know (days not covered by either of the two previous files); discard "don't know" days. For a crude analysis; a 2x2 contingency was created for the entire data set (days):

	Well	Sick
No food record	34980	4496
Food recorded	1742	231

By Chi Square analysis, there was no significant difference, suggesting that there was no bias in sampling days for the whole study group (chi.square = 0.19, p=0.66). However, this analysis did not take the individual toddler into account. The project carried out a more sophisticated log-linear analysis (not included in the final report) which confirmed that sickness did not encourage or discourage reporting of food intake information on toddlers.

Mexico Project. During the period when Mexico project followed the CRSP morbidity protocol, 61 toddlers were studied. These toddlers were reported ill 5.9 ± 7.6% of the time; of the food intake record days, 7.5 ± 12.9% coincided with recorded morbidity. This correspondence indicates that the data set as a whole is not biased, as is true in Egypt. However, for analyses relating intake of an individual to an outcome variable, the problem of bias remains. Among these 61 cases, in 15 cases, illness days were overrepresented in food intake records and in 24, illness days were underrepresented; of the 23 records where illness days were equally represented, 21 had no illness days reported at all. This more complex problem requires further study in relation particularly to longitudinal analyses.

Kenya Project. The morbidity file for this project contains much overlapping illness and the Kenya project has examined illness categories separately. To answer the question posed will require extensive reformatting of the file. From an earlier project report (year V), ME calculated about

30,000 ill days total for toddlers, or about 300 d/child/year. However, these are not different days in all cases.

ME has examined a subsample of 16 toddlers. Some 34 ± 25% of food record days coincided with reported morbidity.

4. Household Morbidity as a Variable

Household morbidity includes illness of both target and non-target members. Probing for illness that was not spontaneously reported was limited to target individuals in the Kenya project, and is likely to have been so in the others; thus, household morbidity is likely to be biased by number of target individuals in a household. Because of differences in methodology between projects (and within projects between target and non-target), the morbidity statistic of greatest reliability is prevalence. All projects gathered weekly information which, in the cases of Kenya and Egypt, is period prevalence (illness during the week), and in Mexico, point prevalence (illness on the day of interview). These prevalence figures are not quite comparable, as the Mexico project underreports illnesses of short duration. (See above.)

To relate household morbidity to household food energy intake (and other variables) it is necessary to aggregate individual morbidity into an operationally useful term that takes account of household demography. The readily apparent issues are two-fold. First, since incidence and prevalence of morbidity seem to be age-group specific, dividing total morbidity by total number of people would seem to be very misleading. An age-based morbidity expectation might be used in weighting the household demography (e.g. expressing adult men as a "toddler unit" based on the community age-specific morbidity rates). The second problem is more difficult to address; since the illnesses of concern are indeed contagious, the composition of the household (age and density) can be expected to influence the probability of illness in the household. The occurrence of an index case can be expected to affect the probability of illness for all others in the household, etc. There does not seem to be an easy way of taking this into account, yet it seems clear that this could effect a bias between large and small households, a bias which might then associate with other household variables that move with household size.

A first approach is to develop a morbidity unit based on the morbidity experienced by the toddler in a given project. If, for example, the mean point or period prevalence for toddlers is 15 of 50 observations, and the mean for lead males is 3 of 50, then lead male morbidity is 0.2 toddler units. (A refinement would examine the data by household size/composition; e.g. in household with x children under age y, lead male mean is z, etc.) Household morbidity experience could then be related to the average expectancy to provide a numeric ranking.

A second approach would allow comparison within a household over time, providing assessment of seasonal trends but not comparing households. It would have the advantage of eliminating the need for a denominator for household morbidity. The approach would be as follows: i) generate aggregate

household morbidity by month; ii) generate the household's median morbidity score; iii) categorize each month's morbidity as above or below the household's median level; iv) generate analogous variables for household total food intake yielding binary classification of above or below the median for each month.

Examine, by calendar month

		Household Intake	
		Below Mean	Above Mean
Household Morbidity	Below Mean		
	Above Mean		

Chi Square analysis should tell us whether there is any association between deviation in household intake and deviation in household morbidity, using the household as its own control and in the above analysis, controlling for month of the year.

The more elaborate analysis would involve building a table as above for each household across months and then dealing with the series of such tables developed for the community.

5. Chronic Parasite Burden

The study populations are exposed to infectious agents that form a chronic disease burden and contribute to some acute illness episodes as well. Major infectious agents were known to differ among locations and include schistosomiasis in Egypt, malaria in Kenya, and amoebiasis in Mexico. Intestinal parasites are common in all areas. All projects examined fecal samples to characterize the population; urine specimens were examined in Egypt and Mexico, and blood smears in Kenya. The methods employed for this element were not suitable for research on endemic parasitism; CRSP resources would not have permitted such research. It was not intended that these studies be used as an analytical variable, but contribute to an understanding of the health milieu.

Egypt project has not reported its findings in detail in the EFR. A previous report (22) states that 85% of households had one or more individual infected with parasites, based on community surveys. Parasitic infection was greater during the winter than summer. In the second survey, 60.5% of a sample of 642 persons had one or more parasites. Types were: Entamoeba histolytica, 36%; Hymenolepsis nana, 17%; Giardia lamblia, 7%; Ascaris lumbricoides, 10%; and others. Schistosoma mansonii was found in 8.5% of the population and Schistosoma haematobium in 3.4%. Infected individuals were said to have been treated.

Malaria prevalence was seasonal in Kenya, reaching a level of 7.9% in the rainy season. Hookworm was very prevalent in a first survey, found in all age groups except infants. Percent of positive specimens was: toddlers, 6; schoolers, 36; male adults, 48; and female adults, 55. In a second survey (after treatment presumably was offered, according to project plan), percentage of positive fecal specimens was much lower: 16 in male and 22 in female adults. About 2 to 3% of fecal specimens revealed ascaris infection, except in schoolers where it was 6% and in infants who were not infected.

The Mexico project reported heavy infection with Entamoeba histolytica, Giardia, and various worms. The highest values, among surveys recorded in the MFR, are as follows:

%	Infant	Toddler	Schooler	Lead Male	Lead Female NPNL	Lead Female P	Lead Female L
Amoeba	19	18	43	27	39	63	48
Giardia	2	18	13	10	3	9	4
Int. Worms*	0	13	36	14	16	23	21

* Sum of all types reported; specimens may have contained more than one type. Evidence of bacterial infection was found in about 12% of urine specimens of Mexican women at all reproductive stages.

6. Mortality

Project plans called for recording of deaths in study households, including the proximate cause of death. There was difficulty in learning the cause of death in Egypt; respondents often did not know or would not say. In Mexico, a death certificate is required for burial, but the project expressed some reservations about the specificity of the cause of death. Kenya data appear to be medically reliable.

The first year of life was hazardous (Table 82). In Egypt, ten deaths occurred among target infants; five additional deaths occurred in the second six months, bringing the death rate to about 140/1000 live births. This number is large enough to warrant an exploration of associated characteristics. Infant mortality also was high in Mexico, but as the MFR states that "none was in the longitudinal sample", analysis would be restricted to household characteristics. In Kenya, there were 2 stillbirths and three perinatal deaths; there were no deaths among target infants.

Deaths of infants and toddlers in the Mexico project were mainly due to infectious diseases. One schooler drowned and two teenagers died of trauma. There was no consistent pattern of disease among male adults.

Two of the three male adults and the schooler who died in Kenya had liver problems (hepatoma, hepatitis). The other adult death was due to laryngeal cancer. Exposure to an environmental agent (aflatoxin contamination in food and/or other) is reported to be likely.

These results have clear policy relevance in identifying age groups and categories of risk but are difficult to treat analytically.

Table 82

Mortality During CRSP Phase II
in Study Households

	Egypt	Kenya	Mexico*
Stillbirth/Perinatal/Neonatal	8	5	3
Infant	2**		7
Toddler	4	1	3
Schoolers			1
Older Children			2
Male Adults	1	3	18
Female Adults	1		

* In study households but not target subjects; "infant" includes up to one year, toddler age 1-3 years.
** Target infants, through 6 months; there were 5 additional deaths between 6 to 12 months of age.

H. Psychology

1. General Comments.

Selection and organization of the cognitive and behavioral variables presented difficulties from the start. Firstly, the nature of this area of inquiry is more evidently affected by cultural factors than any other the CRSP attempted. Secondly, neither the US nor host-country PIs, nor ME, were expert in this field. It is not surprising then, even if disappointing, that the project final reports and data archives are least complete in this area.

Principal responsibility for planning in the psychology domain was vested in able senior investigators from US institutions. All projects were fortunate in finding very well-qualified host-country investigators who participated in the field work as full partners, adapting instruments to local conditions and evaluating outcomes. Still, this was an area in which intra- and interproject communication proved most difficult to sustain.

The difficulty inherent in cross-cultural application of standardized structured tests, most of which have been developed in Western cultures, was foreseen. While these were included in the core design (see Table 83), it was agreed the projects would be free to use culturally appropriate subscales and/or different numbers of component tests, based on experience gained in Phase I.

Predicting which parameters will have greatest power to detect the influence of an independent variable is always somewhat chancy. For this reason, projects were enjoined to archive data for each test item, rather than aggregated scales.[1] For unstructured assessments, the problem is more severe; decisions already taken at field level, about classification of behaviors, if wrong are irretrievable. It is anticipated that cross-project analysis will be feasible when projects have first identified the components (of intelligence, development, behavior, etc.) that effectively differentiate among the population in their own settings. Full program-analysis of these data is likely to be last in emerging from future work.

---

[1] At this writing, the Mexico archive is not complete in filing Brazelton test items.

Table 83

Summary of Planned Measures of Psychological Function
Basic and Common Data Sets[a]

Target Subject	Timing/Frequency	Method/Test
	BASIC DATA SET	
Lead Male and Female	Once	WAIS subscales[b] all to include: Block design Digit span Ravens progressive matrices
Infant at Birth	1-7 days, once	Brazelton Neonatal Assessment Scale (BNAS)
Infant at Birth	At 3, 6 months[c]	Observations of caretaker and infant including physical contact, talking to infant, response to infant vocalization, amount of vocalization.
	At 6 months	Bayley Motor Scale[b, f]
	At 6 months	Bayley Infant Behavior Record
Toddler	At 18, 24, 30 mos	Bayley Mental Scale[b]
	At 18, 24, 30 mos[d]	Bayley Motor Scale[b]
	At 18, 24, 30 mos	Observation of caretaker and and child including physical contact, carrying, adult vocalization to child, child vocalization, caretaker response to child.
Schooler	Entry and Exit	WISC-R subscales[b] all to include: Block design Digit span Ravens progressive matrices
	Twice during school year	Playground and classroom observation

Table 83 (cont.)

Psychological Function (cont.)

Target Subject	Timing/Frequency	Method/Test
	COMMON DATA SET	
Infant	Birth-6 months (timing to be described in more detail by projects)	Dimensions added to caretaker-child observations (# of persons in home, location, caretaker, active involvement by caretaker and infant, response to infant distress) (Egypt and Mexico)
	6 months	Fagan test of infant novelty preference (Egypt, Kenya)
Toddler	At 18, 24, 30 mos[e]	Toy play dimensions
Toddler	At 18, 24, 30 mos	Additional home observation dimension (location, # of persons in home, caretaker, response to distress) (Egypt, Kenya)
Toddler	At 18, 24, 30 mos[g]	Uzgiris-Hunt Scales including object performance, means/ends, verbal imitation, schemes (Kenya, Egypt

[a] This table gives the final plan approved by SCB and notes some deviations from agreed on schedules. Details of variables archived etc. will be found in Appendix E.
[b] Individual projects use culturally appropriate subscales and/or different numbers of components of tests. Projects to report actual variables recorded and weighting factors and groupings for the construction of a composite rating/score for each test.
[c] Kenya frequency every other month.
[d] Kenya and Egypt at 30 months only.
[e] Kenya tested at 30 months only.
[f] Mexico added Bayley Mental Scale at 6 months.
[g] Egypt did test at 18 and 24 months.

2. Assessment of Tests and Procedures.

ME sought expert advice on the tests and procedures finally selected by the projects as presented in the projects' final reports. The consultant, Dr. Elliot Turiel of the University of California, Berkeley, reported as follows:

> My overall assessment of the instruments and procedures used in the projects on cognition and behavior is that, for the most part, the methods used were appropriate, sound, and applied in correct and careful ways. Some of the assessments, however, are more valid and reliable measures than others. In the following, I will provide my estimation of the strengths and weaknesses of the different measures in each project and draw comparisons between their use in the three projects.
>
> It is fair to say that, in principle, all the measures are appropriate to the research hypothesis that food intake affects cognitive development. It is less clear that all the measures are sufficiently developed to adequately accomplish the task. The measures can be classified into two categories. One would include the intelligence tests (e.g., the Bayley Scales and the WISC) and the measures of cognitive development (the Uzgiris-Hunt Scales). The second category includes the measures of social interaction and emotions (e.g., observations of mother-child interactions, rearing variables, playground and classroom observations, self-drawing task). The measures in the first category are better developed and more clearly tied to the research hypotheses than those in the second category.
>
> The measures of cognitive development and tests of intelligence are among the best available assessments of intellectual functioning. If the concerns of the study are with cognitive development (i.e., changes in cognitive functioning with increasing age), the Uzgiris-Hunt Scales would be the most appropriate. These scales measure conceptual attainments of infants and toddlers (such as object permanence) that are developmentally ordered. Accordingly, they are particularly appropriate for assessing types, rates, and levels of cognitive development in the early years. I note that the Uzgiris-Hunt Scales were used in Kenya and Egypt, but not in Mexico. In addition, comparable assessments of cognitive development (the Uzgiris-Hunt Scales are only for the early years) were not used with the school age children. This is probably because such assessments would have been time-consuming and expensive to administer.
>
> Another way to assess cognitive functioning is through standardized tests of intelligence. These are not as powerful as measures of cognitive development because they do not provide direct information regarding age-related changes in the different types of concepts mastered by children in their development nor provide indicators of rates of development. Furthermore, intelligence tests are more likely to be influenced by cultural factors, and less easy to adapt to different cultures, than direct assessments of cognitive development. Nevertheless, the better standardized tests of intelligence are adequate and appropriate for the projects. The Bayley Scales are probably the

best available tests for use with infants and toddlers. Since they were used in each of the projects, the data should be comparable. Similarly, the Wechsler Intelligence Scale for Children (WISC) is a suitable measure that was used across projects (note that I have no information for methods with school age children in Egypt). Other less comprehensive intelligence tests were administered to the school age children. The Raven Progressive Matrices Task (Kenya and Mexico) and the Attention Span Test (Mexico) are less comprehensive than the WISC, but still useful. The Peabody Picture Vocabulary Test (Mexico) and the word recognition test (Kenya) are limited and most bound to specific information known by children. However, the entire battery of intelligence tests (including the sub-scales of the WISC) can be used as a package to examine particular patterns of scores on different aspects of motor, verbal, and conceptual skills. (These considerations apply to the adult samples, since the measures were similar to those used with children.)

The quality control procedures and measures in each project were good (especially in Mexico and Egypt). Testers were carefully trained and monitored to ensure that they knew how to administer the tests and be sensitive to the cultural settings. The reported interrater and test-retest reliabilities are more than adequate.

The social and emotional measures used in these projects are more difficult to evaluate because they were not based on standard methods nor clearly associated with a set of conceptual or theoretical formulations about the phenomena under investigation. Specifically, I am referring to the observations of infant-mother interactions and of caregiving activities (Kenya and Egypt), the mother-infant interaction observations (Mexico), structured toy play (with preschoolers in Mexico), and the playground and classroom observations (Kenya and Mexico). I am not questioning the potential validity and value of these types of observations of social interactions for developing explanations of children's social competencies or for uncovering the role of social interactions in development. In some respects, these measures could constitute creative aspects of the research. It is unclear, however, how these measures will bear on the main hypothesis of the research. That is, it is unclear as to how these findings will be used to examine the relation of food intake to development.

The main reason that these measures may not be well suited for answering the question of how food intake affects development is that there are no formulations of those aspects of development being measured by the procedures. The material I reviewed does not specify what counts for what. How will it be determined that particular patterns of behavior and/or social interactions (reflecting social/emotional characteristics) represent negative or positive features associated with food intake and nutritional status?

The observational measures, especially those with infants and toddlers, actually serve two purposes in the research. In addition to measuring social/emotional characteristics, the observations were meant to provide

an assessment of environmental influences on development. The difficult problem of estimating the extent of influences on cognitive development from environmental factors other than food intake is one for which there are no easy answers. (I am sure the researchers are well aware of all this.) I wish to note, however, that there is need for varied assessments of environmental influences on cognition. The child-rearing environment, as measured in the Kenya study, represents only one type of influence on cognitive development. Actually, the mother-infant interaction observations in the Mexico project go further in that they also focus on the reciprocal mother-infant interactions. Still, there are interactions with other persons, exploratory behaviors with the physical environment, and other such influences that can affect development.

In conclusion, let me reiterate that the work is solid, that as a package the measures are valid and appropriate, and that there are sufficient commonalities among the projects.

3. First Look at Opportunities.

Project final reports are least forthcoming as regards adult and schooler categories, and neither Egypt nor Kenya project report infant parameters. There is less concern about the adult structural assessments than the others, because intelligence is a state variable expected to have explanatory power, but not seen as an outcome of current food intake. Maternal/caregiver data are, however, central to hypothesis-testing. Infant data are a key element in assessing reproductive competency.

Table 84 displays some statistics for the archived cross-sectional samples of adults and schoolers. Cursory examination suggests that age and sex are important considerations in future analyses, as is formal schooling. The range of values recorded for "Block design" differs so markedly between projects as to indicate that the test reported under this title has different components. These kinds of problems will need to be sorted out as a first step in analysis.

Table 84

Cognitive Scores of Adults and Schoolchildren (7-9 years)
(CRSP Cross-sectional Samples, Mean and Range)

	N*	Digit Span Forward	Digit Span Backward	Block Design	Raven's Matrices
Adult Males					
Egypt	81-82	4.6 (0-7)	2.8 (0-5)	4.3 (0-11)	18.1 (7-33)
Kenya	245-304	4.0 (0-10)	3.6 (0-8)	6.8 (0-32)	12.5 (0-32)
Mexico	199-200	4.0 (0-10)	3.5 (0-9)	13.0 (0-46)	16.7 (7-33)
Adult Females					
Egypt	140-143	4.0 (0-6)	1.8 (0-3)	2.9 (1-10)	15.1 (5-35)
Kenya	201-309	3.7 (1-9)	2.9 (0-9)	5.9 (0-25)	13.0 (6-30)
Mexico	244	3.4 (0-8)	2.7 (0-6)	10.0 (0-36)	14.4 (2-29)
Male Schoolers					
Egypt	99 (56)	4.2 (0-7)	2.4 (0-6)	6.2 (1-11)	14.8 (9-25)
Kenya	245 (79)	2.8 (0-8)	1.7 (0-5)	5.1 (0-31)	12.5 (0-32)
Mexico	146 (90)	2.3 (0-8)	2.2 (0-6)	6.9 (0-27)	13.0 (6-24)
Female Schoolers					
Egypt	103 (64)	3.9 (0-8)	1.5 (0-6)	5.2 (1-11)	13.1 (7-23)
Kenya	201 (65)	2.8 (0-8)	1.8 (0-6)	5.0 (0-19)	13.0 (6-30)
Mexico	139 (84)	3.4 (0-8)	2.0 (0-5)	6.6 (0-38)	12.7 (4-22)

* N = number of observations. Adults were tested once. There were repeated tests of children; number in parentheses is number of subjects.

Results of one component (orientation) of the Brazelton Neonatal Assessment Scale are presented in Table 85.

Table 85

Comparison of the Scores Recorded for the Orientation Components of the Brazelton Neonatal Assessment Scale in Egypt and Kenya Projects and the Aggregated Lester Orientation Score Reported by Mexico Project (Data based on variables in ME archive)

Brazelton Score	1	2	3	4	5	6	7	8
Inanimate, visual								
Egypt			4.3	7.6	5.4	31.5	44.6	6.5
Kenya	10.2	12.7	16.1	18.6	14.4	18.6	9.3	
Inanimate, auditory								
Egypt			2.2	2.2	2.2	28.3	61.1	1.1
Kenya	5.1	4.2	16.1	38.1	28.0	6.8	1.7	
Animate, Visual								
Egypt			2.2	1.1	1.1	29.3	62.0	4.3
Kenya	7.6	16.1	19.5	14.4	16.9	13.6	11.0	0.8
Animate, auditory								
Egypt			1.1	1.1	1.1	35.9	60.9	
Kenya	3.4	2.5	18.8	36.4	28.0	7.8	2.5	0.8
Animate, Visual and Auditory								
Egypt			4.3	4.3	2.2	39.1	44.6	5.4
Kenya	9.3	11.0	28.0	16.1	11.9	13.8	7.8	2.5
Lester Aggreg. Score	1-1.9	2-2.9	3-3.9	4-4.9	5-5.9	6-6.9	7-7.9	8
Mexico	9.3	18.1	22.9	9.8	24.6	9.8	3.2	

A quick inspection of this table suggests either that the three populations of neonates are unlike developmentally, or that the test administrators did not use uniform methodology or criteria for scoring. Orientation scores in the Egypt project were clustered at the upper end of the scale (scores 6 and 7). Kenya data describe a normal distribution curve with midpoint scores of 3 or 4 on an 8-point scale. The aggregated Lester BNAS orientation scores of the Mexico project have a bimodal distribution with peaks at scores 3 and 5. ME looks forward to project psychologists' future explication of the divergence of these measures, prior to any attempt at cross-project comparisons.

I. Reproduction

1. General Comments

The protocol designed to address reproductive function called for study of the lead female during the last six months of pregnancy and of the mother and infant for six months after delivery. The hypotheses to be tested linked maternal intake during pregnancy to the outcome of pregnancy in terms of infant endowment at birth and maternal health, and during lactation to infant growth, health, and development and maternal condition.

Measurements to be made during pregnancy and lactation were added to the protocol for all lead women (i.e. monthly food intake, weekly morbidity, quarterly anthropometry, etc.). Most added measurements were scheduled for the fifth and eighth month of pregnancy, at or near delivery, and the first, third, and sixth month of lactation. These were clinical examination and obstetrical history; RMR; hemoglobin; and serum ferritin, albumin, and IgG. Measures specific to the lactation period were breast milk IgA and IgG and infant measurements as follows: anthropometric measures at birth and monthly; Dubowitz score at birth (for gestational age); Brazelton test within the first week (for cognitive assessment); weekly record of morbidity and supplementary feeding; and Bayley Motor Scale and Infant Behavior Record at six months of age. Observations of infant and maternal/caretaker behavior were included as appropriate to each site. Data sets are most incomplete with respect to the immune factors and most complete with respect to anthropometric measures (except Egypt) and food intake (except Mexico).

Projects entered subjects at different times in relation to conception. Plans generally called for enrolling households where a pregnancy would be expected based on age of youngest child (see part VI.A., above). Women's reproductive history was to be obtained at entry and a monthly questionnaire inquired about conception. In the regular reproductive cycle in the project sites, many women fail to menstruate between pregnancies; projects offered but did not require, a pregnancy test. In some cases, measurements were begun as soon as the household was entered; in others, measurements began only after conception. Thus, there are variable, but rather small, numbers of women for whom data were recorded both before pregnancy and during the first trimester, in addition to the CRSP protocol period (second and third trimesters, first six months of lactation).

According to the MF criterion which is based on the presence of food intake measures during the CRSP protocol period, the longitudinal samples have an N of 62 in Egypt, 93 in Kenya, and, in Mexico, 69, but which falls to 26 if limited to the last food intake method. Depending on the variable of interest, longitudinal N's are larger or smaller. In the Kenya sample, 94% have complete anthropometric measures; in the Mexico population, about 72% have at least three-fourths of those measures, but fewer than 50% of the Egypt sample has as many as one-half of the scheduled measures (see section VI.B.) The cross-sectional samples are substantially larger, but cross-sectional data are of limited value (unless N's are very large) because many topics of interest in research on pregnancy and lactation require information about changes in status at key times or over time.

Projects have defined periods of pregnancy and lactation for their report purposes as shown in Table 86. Kenya and Mexico projects define the date of conception as birth minus 280 days (except for abortions); Egypt project uses birth as 0 time for all calculations.

Table 86

Definition of Reproductive Periods
in Project Final Reports

	Egypt	Kenya	Mexico
Prepregnancy		0-90 to 0+60d	0+60d
Pregnancy			
Trim1		0 to 75 -105d	0-92d
Trim2	3-6mo	75-105d to 195-205d	93-186d
Trim3	6-9mo	195-205d to term	187-280d
Early Preg	Birth-190+30d(3mo)		
Mid-Preg	Birth- 95+15d(6mo)		137+15d
Late Preg	Birth-15+15d(9mo)	Birth-15 to 30d	250-280d
Post-partum		Birth to 9d	
Lactation(time 0=birth)			
1st per	0-3 mo	0-45d	30+15 - 91+21d
2nd per	3-6 mo	46-135d	92+21 - 183+21d
3rd per		136-180d	184+21 - 254+21d
1st mo	30+15d		
6th mo	180+30d		

Some data in the final reports are tabulated by month or aggregate months. Note that lactation periods in the MFR are exceptional in including a third entry covering the sixth through eighth month, a period outside CRSP protocol.

Mexico lactation periods one and two correspond roughly to Egypt periods designated 0-3 and 3-6 months; Kenya lactation periods do not coincide exactly with any other project, being 0-1.5 months, 1.5 through 4.5 months, and 4.5 through 6 months. These differences present some difficulty in comparing project reports directly, but equivalent periods can be generated from the data archived for future analyses.

2. Maternal Food Intake

Data on food intake as reported by projects are displayed in Table 87 and may be compared with values derived by ME from these data sets (Table 25, section D., above) (but defining the target population differently. See section VI.B.) Mexico project reports food intake in two places (chapters 20 and 31). Values shown in Table 87 are from MFR Tables 20.6 and 20.7, where physiological state is clearly identified as either pregnant or lactating or both. Figures in MFR Table 31.1 are much higher (2636, 2300, and 2425 kcal/d in trimesters 1, 2, and 3, and 2700 kcal/d in the first three months of lactation); the pregnancy sample includes some women who are lactating (personal communication, L. Allen). Physiological state cannot explain the discrepancy in reported values, given that the N's are approximately the same in both Tables and that a difference is shown even in early lactation. Resolution of this question will require reanalysis of the data set.

What immediately stands out is that energy intake is greatest in Mexico, even in comparison with the lower reported figures; it is at least 200-300 kcal/d higher than intake by Egyptian women and 600-900 kcal/d higher than in Kenya. The difference is even greater when body weight is taken into account; the calculated Mexico figures of 42-48 kcal/kg/d are much higher than is usually encountered in poor populations.

A second striking point is that intake does not vary with reproductive phase in Egypt, but is higher during lactation than pregnancy in both Kenya and Mexico (by about 300-400 kcal/d). There is no evidence to indicate that food intake increases during pregnancy in any project population; in fact, it may be less.

The evidence for decreased intake during pregnancy and increased intake during lactation appears strongest in Kenya, but the data are confounded with the effect of the drought-associated food shortage. The KFR (Tables 23.7 and 23.8) documents differences in food intake across pregnancy by birth cohort. The KFR gives the dates of the early food shortage (EFS) period as July through September, 1984, and the late food shortage (LFS) period as October through December, 1984. Birth cohorts would have experienced food shortage at different reproductive stages, as shown in Table 88. Except for the April-June 1985 cohort, which has N of 9-10, each cohort has 20-30 cases.

Table 87

Maternal Food Intake During Pregnancy and Lactation as Reported by CRSP Projects ($\bar{x} \pm SD$)

Period	Egypt[a] per day	Egypt[a] per kg	Kenya[b] per day	Kenya[b] per kg	Mexico per day[c]	Mexico per kg
			Energy, kcal			
Prepregnancy			1851±587	37.9±12.4	2371±833	40.9±10.8
Pregnancy						
Trim1			1590±492	32.2±10.0	2139±844	41.4±15.2
Trim2	2001±475	32.5±9.1	1482±492	28.3± 9.1	2208±851	42.1±16.3
Trim3	1972±445	30.4±8.5	1351±475	24.6± 8.5	2230±876	42.8±16.7
Lactation						
0-3 mos	2120±401	35.1±9.0			2476±933	46.6±17.2
3-6 mos	2097±340	34.4±7.6			2573±925	48.1±17.4
0-1.5 months			1830±729	35.5±13.9		
1.5-4.5 months			1791±448	35.5± 9.6		
4.5-6 months			1771±529	35.0±10.3		
			Protein, g			
Prepregnancy			56.3±23.1	1.2±0.5	68.6±27.4	1.2±0.3
Pregnancy						
Trim1			45.9±17.3	0.9±0.4	58.6±25.1	1.2±0.4
Trim2	67.1±18.7	(1.1)	41.6±15.0	0.8±0.3	60.6±24.6	1.2±0.4
Trim3	65.6±20.4	(1.0)	37.9±15.4	0.7±0.3	64.1±28.3	1.2±0.4
Lactation						
0-3 mos	73.7±20.8	(1.2)			68.1±27.4	1.3±0.4
3-6 mos	71.8±16.1	(1.2)			72.3±29.2	1.4±0.4
0-1.5 months			51.0±23.1	1.0±0.4		
1.5-4.5 months			53.7±17.1	1.1±0.3		
4.5-6 months			53.3±19.7	1.1±0.4		

[a] EFR Table 5.31. N per day = 80, per kg = 66-79.
[b] KFR Tables 23.1 and 23.2. N is the same/d and /kg for respective periods shown = 40, 80, 108, 117, 118, 124, 122.
[c] MFR Tables 20.6 and 20.7. N for the respective periods = 27, 59, 87, 102, 106. "Prepregnancy" entry is the values recorded for 98 women who were neither pregnant nor lactating, not a longitudinal pregnancy sample.

Table 88

Coincidence of Reproductive Phase
and Food Shortage in Kenya

Birth Date	Pregnancy			Lactation		
	Trim1	Trim2	Trim3	0-1.5m	1.5-4.5m	4.5-6m
Apr-Jun 84					EFS	LFS
Jul-Sep 84			EFS	LFS	LFS	
Oct-Dec 84		EFS	LFS	LFS		
Jan-Mar 85	EFS	LFS				
Apr-Jun 85	LFS					

A rough calculation from cross-sectional data given in the KFR indicates that intake was at least 300 kcal/d lower, on the average, in food shortage periods than at other times. (Data not reported here.) The data also show that, in equivalent calendar periods, energy intake was not appreciably different between the first and second trimesters of pregnancy and that the differential in intake between lactation and pregnancy was about the same (300-400 kcal/d) in periods of comparable food stringency. The exact dates during which food shortage was experienced by a given woman need to be coupled with more precise estimation of pregnancy/lactation stage for complete analysis of this issue, using the longitudinal data set.

The proximate composition of diets was very similar in Kenya and Mexico; the percentage of calories derived from protein were 11.9 and 11.6, respectively, and in both cases 8% of protein was from animal sources. About 13% of energy was derived from fat in all three projects, but the Egyptian diet had somewhat more protein - 13.4% of energy - and much more was of animal origin (40%).

Future analysis will depend on accurate identification of reproductive phase, including knowledge of whether or not women included in the projects' categories "pre-pregnancy" and "pregnant trimester 1" were lactating. If they were lactating until they discovered themselves to be pregnant, and promptly weaned the infant, food intake may be inflated in these periods due to lactation demand.

The Mexico project reports that the great majority of women continued consumption of pulque during pregnancy and lactation. The "alcohol" variable will need to be examined in relation to pregnancy outcome.

3. Maternal Anthropometric Variables

Variables in common across projects are: body weight during pregnancy and lactation; height when not pregnant; head circumference; mid upper arm circumference; triceps skinfold; and biceps skinfold. Projects have reported these data in different formats, but a uniform set can be constructed from the archived data for part of the longitudinal sample. Values for some measures are presented, as reported by the projects, in Table 89.

Women in the Mexico sample were the shortest, on average, and in Kenya, the lightest in weight; the Egypt population was a bit taller and significantly heavier. The ranges of height and weight overlap sufficiently that comparable subsamples could be generated by stratification. The degree of fatness and its change in the course of pregnancy and lactation have been assessed indirectly by measures of skinfold thickness. Egypt project has tabulated the values measured at two arm sites; measurements at other body sites were ruled out because of cultural constraints. Both Kenya and Mexico projects recorded subscapular skinfold thickness, and from the three measures they have calculated percentage body fat using the Durnin and Womersley formula. (23) The validity of this formula in application to pregnant women has not been tested, but the calculations (KFK Tables 23.3 and 23.5, MFK Table 31.1) indicate the Kenyan women to have had about 24-26% body fat and Mexican women, about 34-36%. A simpler indicator of fatness, the Body Mass Index (Wt_{kg}/Ht_{m}^{2}), was calculated by the Egypt and Kenya projects. BMI in Egypt was about 28 and in Kenya, 21. ME's rough calculation yields an estimate of 23 for the Mexico sample.

Body weight gain during pregnancy was the same in Egypt and Mexico, about 7.4 kg for the second and third trimesters; Kenyan women gained less, 6.3 kg. On the average, Egypt women gained a negligible amount during six months of lactation; Mexico and Kenya populations lost almost 1 kg during that period. There is substantial variation in all these figures, more so during lactation (CVs 300% +) than pregnancy (CVs 50%), and ranges overlap between projects.

Table 89

Maternal Height and Weight During Pregnancy and Lactation as Reported by CRSP Projects ($\bar{X}$ ± SD) (N)

	Egypt[a]	Kenya[b]	Mexico[c]
Height, cm	155.4±5.3(117)	154.3±5.7(90)	152.5±5.0(130)
Weight, kg			
Prepregnancy		50.7±7.6(90)	54.4±7.2 (45)
Trim1 or 3rd mo	61.6±10.4(96)	50.5±7.7(90)	54.5±7.2 (29)
Trim2 or 6th mo	63.2± 9.8(63)	52.6±7.1(111)	59.0±7.0 (68)
Trim3 or 9th mo	67.9±10.1(88)	55.6±6.9(121)	62.0±8.2 (88)
Post-partum		51.6±6.4(115)	
Lactation			
mo 1	63.7±10.8(61)	51.6±6.5(123)	57.1±7.6 (111)
mo 3	61.3±10.1(61)	51.2±6.9(98)	57.3±8.0 (115)
mo 6	63.5±11.9(73)	50.7±7.4(105)	56.4±8.6 (116)
Weight change, kg			
Pregnancy			
Trim1			0.18±2.1 (18)
Trim2	3.7± 2.9(45)		4.48±2.6 (20)
Trim3	3.7± 2.5(46)	3.1±1.8(111)	2.86±1.8 (52)
2 & 3	7.4± 3.9(73)	6.3±3.4(111)	
1 - 3		6.4±3.4(61)	7.66±3.8 (34)
Lactation			
0-2 mo		-0.9	
1-3 mo	0.27± 2.1(30)		-0.04±2.2 (99)
3-6 mo	0.01± 3.0(40)		-0.71±1.9 (105)
1-6 mo	0.18± 3.5(38)		-0.92±3.4 (99)
Body Mass Index[d]			
Pregnant			
3rd mo	25.6± 4.1(95)	21.2±2.6(90)	[23]
9th mo		23.5±2.5(121)	
Lactating			
6th mo	26.1± 4.8(70)	[21]	[24]

[a] EFR Tables 5.36, 5.38
[b] KFR Tables 23.3, 23.4, 23.5, 23.6
[c] MFR Table 31.1
[d] Wt_{kg}/Ht_{m}^{2}. Figures in brackets calculated by ME from mean ht and mean wt in table.

4. Resting Metabolic Rate

RMR (called REE, for resting energy expenditure, in the KFR) was measured by all projects, but the Egypt results are not yet validated. (See section VI.E. for full discussion of this variable.) Different methods were used in Kenya and Mexico, but the tabulated values are remarkably comparable in the two populations (Table 90). Because of the difference in body fatness between Mexican and Kenyan women, the RMR per kg lean tissue would appear to be higher in the Mexico sample. An apparent (but not real) difference in this direction would be anticipated based on the methodology, in which the Mexico method gives slightly higher values. (See section VI.E.) Within a project, there is no significant difference in RMR per kg between reproductive states. Thus, RMR per day would vary according to changing body weight during these cycles.

5. Maternal Morbidity

Morbidity was recorded by the methods discussed in section VI.G.,and the results are subject to the same caveats. There was no maternal mortality. Findings from the repeated clinical examinations have yet to be analyzed fully.

Egypt project reports extensively on the type and duration of illness among 73 pregnant and lactating women. According to Egypt criteria, women had very slightly more illness during pregnancy than lactation; they were ill 4.3±10.1% of the time in the second trimester of pregnancy and 4.4±10.1% in the third, 3.7±8.0% in the first three months of lactation and 2.7±5.3% in the next three-month period. Unacceptably low hemoglobin values were reported in 23, 57, and 65% of women in the first, second, and third trimesters, respectively.

Kenya reports there was no difference in acute morbidity between pregnant/lactating women and the population of head females. Morning sickness was frequent during the first trimester and the eighth and ninth months of pregnancy. About 2-4% of the women were positive for malaria. Hypochromic, microcytic anemia was the most common form; mean hemoglobin concentration was 12.4±1.7g/dl in nonpregnant, nonlactating women and 10.8±1.6g/dl in the pregnant sample. Macrocytosis, characteristic of folic acid deficiency, was found in 6-8% of women. Ten percent of pregnant women had 1+ goiter and 6% had 2+ goiter (thyroid enlargement) probably due to iodine deficiency.

Mexico project reports lactating women to be ill about 2-8% of the time observed. Some 10-12% of samples revealed bacteriuria during pregnancy, about 3% had Giardia in the feces, and about 20% had low and 20% high counts of Entamoeba. Unacceptable hemoglobin levels were reported in 28% of nonpregnant, nonlactating women and about 17-21% during pregnancy and lactation.

Table 90

Resting Metabolic Rate of Pregnant and Lactating Women
as Reported by Kenya and Mexico Projects
(kcal, $\bar{X}\pm SD$)

RMR	Kenya[a] per day	Kenya[a] per kg	Mexico[b] per day	Mexico[b] per kg
Nonpreg, non-lact			1229±234	22.4±4.1(33)
Prepreg	1114±158	23.8±3.0(21)		
Pregnancy				
Trim1	1116±158	23.1±4.3(19)	1258±198	24.1±3.8(39)
Trim2	1230±201	23.5±3.9(63)	1281±153	23.4±3.8(43)
Trim3	1280±180	23.7±3.2(74)	1356±172	23.3±3.4(36)
Lactation				
0-1.5 mo	1209±183	21.0±3.8(75)		
1.5-4.5 mo	1185±192	23.3±4.0(92)		
4.5-6 mo	1151±152	23.6±4.3(15)		
0-3 mo			1280±206	22.9±3.2(38)
3-6 mo			1257±228	24.1±4.7(32)

a MRC Table 23.13; (N) is number of subjects.
b MRC Table 25.2; (N) is number of observations from 35 subjects, the cohort followed from prepregnancy through lactation.

6. Infant Parameters

Anthropometric, morbidity, and cognitive variables are discussed fully in earlier sections of this report (sections VI.F. G.,and J., Tables 67,78-80, and 85). Much remains to be done before meaningful cross-project analyses of outcomes can be mounted.

Supplementary feeding of infants is a topic of special concern to all projects. Food supplements are the only component of infant food intake for which quantification may be sufficient to allow estimation of energy. There was no feasible way of measuring breast milk composition and intake with sufficient frequency to provide a reliable estimate of infant energy intake from that source.

Egypt reports that 100% of infants were fed sugar in water during the entire first 6 months of life, and a third to a half were also given rice water and sugary concoctions. Except for these items, 92% received no supplement in the first month, 73% in the third month, and 54% in the sixth

month. Thus the supplementary energy source was devoid of any other nutritional value for over half of the infants even as late as the sixth month. The energy value of supplements has not been estimated to date. Dairy products were the first nutritious foods introduced.

Kenya reports that about 40% of mothers gave formula, cow's milk, or other foods during the first month; supplementation increased to 90% by the third month, and 97% thereafter. Sugar-water was given during the first several months. Other common foods were local staples (porridge, etc.).

Mexico reports that foods were introduced gradually beginning about three months of age. Sweetened herbal teas were "often the only supplemental food given...in the first few months of life", and sugar constituted 12% of the energy derived from food supplements. Dairy products provided 18% of non-breast milk calories. Infants commonly received ingredients of pasta and bean dishes.

Given these descriptions of supplementary foods, precise identification and quantification will present some difficulty to and between projects.

J. Social and Economic Measures

Early discussions within the CRSP focused on social and economic outcomes of mild-to-moderate malnutrition. It was judged, however, that variations in social and economic performance would be difficult to measure and analyze during the relatively short period of time available to the CRSP. It was decided that socioeconomic information would provide background data against which the relationships between intake and physical and psychological function of the individual could be examined. Nutritional intake is usually found to be closely related to household income and assets. The CRSP seeks to examine the effects of nutritional intake on function independently of relative poverty or wealth.

1. Description of Methods

a. Scaling. Information on household social and economic status was collected in order to be able to assign scores for social and economic status to each household. To do this required collection of information on housing, occupation and education of lead adults, and information on assets such as land, animals, vehicles, and tools. Each project undertook to collect that set of data which they believed to be most appropriate to permit the construction of indexes or scales. In each project's report will be found a description of methods developed to provide quantitative socioeconomic scales. Because each of the projects approached the problem in a different way, it is not easy to make comparisons and each must be interpreted separately.

Egypt:

The Egypt project's early ethnographic work which focused on family and social structure provided the basis for designing instruments to rate households. Concentrating on instruments that would be objective while culture-specific, investigators examined household demographic structure, housing characteristics and assets of the household, occupation and formal education of target males.

Two scores were constructed: (1) a demographic index score rating occupations and formal education of target males by years of schooling; and (2) an economic index derived from household and farm assets.

For each component, a score of 1,2,or 3 was assigned. EFR Table 4.33 gives the breakdown of core households showing that most households lie in the lower intermediate range. When broken down according to farm and non-farm households, it appears (EFR Table 4.39) that agricultural households are rated lower socioeconomically than non-farm households.

The Egyptian measures of social and economic status locate most households in the "lower intermediate" class. The combination of indicators has a remarkable consistency of measurement, apparently acting as an internal reliability check as well.

Kenya:

This project was the only one to develop a separate "emic" or within-culture scale where rankings were done by community leaders. Toward the end of the field project, sub-chiefs and senior indigenous fieldworkers assisted in developing hierarchical criteria of social and material success among the Embu. It is striking, but not surprising, that the locally developed scale ranks some criteria significantly lower--occupation, for example--than would most "etic" or outside-culture expectations.

In addition, an extensive survey instrument was used quarterly throughout the field period, asking questions about over thirty variables felt to be related to SES. The correlation of the two rankings was 0.48.

Additionally, in the final months of the study a second questionnaire was devised to resolve some apparent inconsistencies between land ownership and land use, and their relationship to agricultural income.

The SES scale presented in KFR Table 13.3 reports on the original survey questionnaire-based ("Form 631") variables, after selecting 30+ original variables on that questionnaire. The basic pattern would appear to be similar to the situation in Egypt: a population rather clustered toward the lower/middle end of the scale, a small group who appear to be significantly deprived, and small populations of increasing wealth and property.

Further work integrating the 'emic' scale with analysis of variables from the household survey will be undertaken in the future and should provide interesting and valuable results.

Mexico:

Mexico has presented one combined scale of SES, in a single table (MFR Table 19.3). This combines, through rank ordering within each scale, two scales: CASA, a scale based on household construction items and number of rooms and beds, and MSL, a scale using market price value of material possessions that require cash purchase. The CASA scale is reported to have a normal distribution and the MSL to be skewed due to the presence of a few families with many possessions. Distribution of items that make up the scales is not given.

The combined scale is called SES, and is further distinguished as SES1 (= entry) and SES2 (= exit), since the base questionnaires were administered at entry and at exit from the study. As might be expected with SES conceived as a state variable, the entry and exit variables correlate well.

The Mexico project differs from the other two projects in that it did not use education, occupation, or non-material, prestige-related items. The project reports that occupation was not a useful variable, as almost all list the same two occupations (farmer and laborer.) Some degree of detail in a less aggregated form might be useful, and this level of detail might make comparisons easier. It is, however, striking that the Egypt project has chosen

one discriminatory scale based on occupation alone, while the scale for the Mexico project looks only at material possessions.

b. Some related inferences about the study populations themselves.

Egypt:

In comparison with the overall Kalama village population, the study population is found to be more homogeneous on socioeconomic considerations. It is also clear that the degree of dependence on agriculture runs through both the village as a whole, and the project study population.

The report creates a strong impression of an increasingly sophisticated, urbane and urbanized community, where farming was once survival and is now coming to be more and more an "extra". The term "village" is probably misleading; it appears to be more an agricultural community becoming a suburb of an enormous city. Many of its members have lived and worked elsewhere, many in other countries. Egypt itself has been in the center of Mediterranean life, and the Nile valley the center of Egyptian life for so many millennia that a Nile valley settlement would be, in many respects, more urban than rural.

Kenya:

This report will be discussed in more detail, below. However, it appears to be dealing with a population of some substantial heterogeneity in terms of land use/ownership and some other forms of wealth. This may be due to the varying agricultural ecology of the area, it may be due to the changing nature of the cash economy, or it may be due, as in the case of Egypt, to increasing variation in dependence on agriculture and on whether farming is for cash or food crops.

The Embu population would appear to be the most local, and least urbanized, of the three CRSP communities. There is little indication of migration of workers to the provincial capital, for example. Agriculture would appear to be an extremely important source of food and livelihood here.

Mexico:

In spite of early concerns that ultimate ethnic origins of the Solis Valley villages might make the contemporary populations too diverse to be readily comparable, this population would appear to be substantially homogeneous, at least as measured. One might wonder if other measures of SES such as class and prestige, or a finer classification of occupation might have revealed greater diversity among these village populations. It can be argued that material items of culture are more easily transported and give the appearance of greater homogeneity than may be the case.

The ingenious use of a question of what individuals would do with an extra 5,000 pesos gives an indication that few people regard improvement in agricultural implements as worthwhile.

This site appears to be relatively sophisticated, rural yet open to the outside world, affected by and drawn to the enormous Mexico City metropolitan area. Workers are reported to commute to work in that city, and some women go there to give birth.

There is much speculation about the influence of truly enormous cities, whose populations are on a scale not previously known in human history. Cairo and Mexico City are among the world's very largest cities, and it is likely that even remote proximity to such an urban mass may have its effects on populations that might otherwise be termed rural. Nairobi, of course, is more distant in miles from the Embu site, but is considerably smaller than either of the other two capital cities. Enormous cities have enormous centripetal effects, economically and perhaps culturally as well; the length of their shadow may be being hinted at in some of these CRSP results.

c. Some reflections on methodological point of view across the projects. In reviewing the SES sections of the three reports, one sees a mixture of quantitative/survey and descriptive/ ethnographic methods. In some cases--Kenya's use of internal and external SES scales, for example--both methods were used somewhat independently, with interesting if predictably divergent results. In other cases, survey questionnaire methods were the only method used, at least as reported.

While early discussions on what was to become the Nutrition CRSP were favorable to descriptive data and to qualitative methodology in many study areas, later reviews favored more quantitative approaches. Ethnographic surveys were very important in preliminary work in the villages, providing investigators with information, sometimes serendipitously, about the community and its values that could not have been foreseen in any a priori hypothetical formulations.

Ethnographic research also permits a substantial degree of insight into how the society views itself. This information can be important for policy purposes.

At the end of a long, complex study, investigators often wish that they could begin again to do the study properly; this is the well-known "had I known then what I know now" phenomenon. Ethnographic field notes, painful as they are to accumulate, are more susceptible to meaningful re-analysis than are most questionnaire protocols. Survey and demographic analysis are less likely to determine internal values and structures, and harder to re-evaluate once the investigator has a better notion of what the internal rankings may be.

At any rate, each project appears to have used some combination of methods, by design or by necessity. Ideally, the Kenyan 'emic' scale would have been developed first, not second; but it was developed, and is worthwhile. In fairness, the counter-argument would be that the 'etic' perspectives, developed principally in relation to American investigators' culture, would be more likely to provide comparability across the three cultures studied. To develop three separate 'emic' sets would arguably make comparability much more difficult, at least initially. Perhaps the best way would have been to develop both, and see, empirically, what transpired.

d. Comparability of SES scales. Given the above, a key question is whether or not the scales are essentially comparable in any meaningful way. With some reservations, it would appear that the answer is affirmative or could at least easily be made to become affirmative.

Specifically, two of the three Egyptian sub-scales--farm wealth and home wealth--are quite similar in conception to the MGL and CASA sub-scales of the Mexico project. Once analyzed and integrated, the two 'etic' questionnaires of the Kenya project would probably provide items for two comparable sub-scales, and would probably be roughly comparable with the combined farm/home material possessions and wealth scales of the other two projects.

The task of comparing scales across projects would take some time and a reasonable knowledge of the field situations, but it would appear feasible.

2. Problems and Possibilities in Comparability Analysis

At the MF level, concern about analysis is principally involved with the cross-project efforts, and therefore with the need to see which of the results can be compared with reasonable confidence across two or more projects. Much of the value of the CRSP, for the funding agency and for policy formulation, is thought to lie in results that can be generalized to other cultures.

a. Some of the difficulties. First, there are relatively few variables that have only one meaning in all cultures. Worse, perhaps, is the fact that some variables may appear to be the same, but are not: an automobile will have a different weight, as will electrical appliances, depending on local value, price, and availability, as well as subtler cultural values. Thus the task of comparing across cultures requires sensitivity and considerable foresight.

Second, the ease and validity of response may vary considerably across cultures. Some may exaggerate cash income to an interviewer, another culture may underestimate; some may fear retributive jealousy if they cite material possessions, others may fear loss of prestige if they do not.

Third, the apparent mix of methods in developing scales may make analysis across projects very difficult.

b. Some possibilities. First, there are some cases where individual scale items are the same across two or more projects. Motor vehicles, window construction, and some appliances can be found across projects. Additionally, possession of types of animals, of units of land, and of some types of income also exist across projects.

Even a very moderate understanding of cultural relativism would advise against taking these indications too literally and too immediately, of course. Still, something can certainly be made of these comparable scalar elements, provided the comparisons are done with some sensitivity.

What most definitely is lacking is a reliable indicator of relative real income across sites: detailed person/hour/compensation equivalencies, and comparability of value of diets raised by the household to the cost of a comparable diet in a cash economy.

Second, it is often helpful, in dealing with two different measures of the same variable, to compare the two extreme populations on the two scales. The logic behind this is classic: if both scales really are measuring the same variable, albeit differently, one would expect there to be less ambiguity of classification at the extremes.

Other methods involve examining distributions by percentiles, by measures of skewness, and by graphing. Percentiles can be useful in making summary measures for comparisons across groups. Care must be taken in these comparisons, however, because of the differences in scoring methods.

3. An Approach to Analysis: Kenya as an Example

The Kenya Final Report was selected for some additional analysis here for three reasons:

- It came in the first, providing more time for reflection and analysis;

- It is the only one that lists all the elements used in the different scales (the "questionnaire 631" SES variables);

- The Food Composition data base was developed in Berkeley, so simple cross-tabulations and correlation analytical files could be rapidly created and employed.

a. Relationship among the SES variables. In what follows, the reader will need to be aware of two choices made in terms of the Tables which follow.

First, the analyses were based on the initial, or entry, household interview--as opposed to the average of all the quarterly SES household interviews, or on the equivalent ratings on all interviews. This was done to permit comparability across projects, to avoid an artificial smoothing of the data, and to avoid factoring in any drought period effects that might create a downturn in family fortune. Second, Spearman correlations are presented since many of these measures are of an ordinal (not interval) nature.

Table 91 gives a fairly representative view of some interesting, and also occasionally puzzling, aspects of the SES variables among themselves. They do not, in general, appear to correlate very strongly, with one or two fairly predictable exceptions.

A second aspect of the SES variables, which it has not been possible to develop here, is that the correlation matrices give subtle indications that there may be some disparate groups, or types, of wealthy persons.

Table 91

Spearman r's of Selected Kenya Final Report SES Variables from Form 631.[a]

Land/Cash/Cattle Relationships.	r	p
1. Land-cash		
Land Ownership and Land Use	0.61	0.0001
Land Ownership and Cattle Ownership	0.40	0.0001
Land Ownership and Sheep and Goat Ownership	0.19	0.0001
Land Ownership and Ownership of "Other Animals"	0.19	0.0001
Land Ownership and Income/Cash at Last Pay-out	0.17	0.0001
Land Use and Income/Cash at Last Pay-out	0.14	0.001
2. Animal Ownership		
Cattle Ownership and Sheep and Goats	0.37	0.0001
Cattle Ownership and "Other Animals"	0.34	0.0001
Sheep and Goats and "Other Animals"	0.33	0.0001

[a] Variables collapsed as follows: Land ownership - none, <8 acres, >8 acres; Land use - <2 acres, 2-5 acres, >5 acres; Sheep and Goat ownership - yes, no; "Other Animal" ownership - yes, no; Income/Cash - none, <2 shillings/mo, >2 shillings/mo.

The reality may be something as simple as different paths to wealth in a traditional society in transition, but further analysis may uncover a pattern wherein one type of wealth leads to a generally better nutritional status, another type of wealth does not.

Coming to particulars, land use and land ownership shows one of the highest associations among the SES variables (r=0.61). However, given the inherent association that might be expected between the two, the Spearman's r reported could as plausibly be considered low.

Land owned and cattle owned shows a weaker relationship (r=0.40), although still one of the strongest among the reported variables. Land owned and sheep and goats owned shows a very much reduced relationship (r=0.19). Land owned and "other" animals (chickens, cats and dogs, and sometimes rabbits) owned shows the same measure of association as for sheep and goats.

The association between land owned and income (defined as cash received from last pay-out) invites some further investigation into the pattern of compensation for land use, and into the real nature of land owned as wealth.

A further related question is raised by the slightly lower association between cash received at pay-out and land cultivated.

Similar relationships, although weaker, exist among the manufactured consumer durables listed in Form 631. (Data not presented here.)

b. Selected SES variables and toddler food intake. Table 9C displays a pilot analysis of SES variables and food intake of one target group, toddlers. Mean energy intake over the study for each toddler was correlated with several individual items from the Kenya SES field form (#631). All SES measures for each toddler were used, giving a total of 611 observations for 131 toddlers. The selection of categories for the SES variables was judgmental; the criteria were admittedly subjective. The intent was to explore potential major relationships, and to identify relationships which appeared to be of unusual strength, weakness, or consistency.

The immediately striking point is the relatively slight nature of the measures of correlation (a maximum correlation coefficient of 0.2) between food intake of the toddlers and the individual SES variables. The small size of the correlations, even though occasionally statistically significant, indicate generally weak relationships. It would be hard to argue that toddler food availability is a proxy for poverty/wealth in the Embu case, at least. It would be of interest to re-do this type of analysis using a composite SES score such as has been developed by Kenya Project. If these initial observations are confirmed by more in-depth analyses, and across all targets, then there will be considerable support for the view that measures of SES and measures of food intake are not readily interchangeable when studying the problem of malnutrition. It is fair to note, however, that the Mexico project reports a correlation between their SES score and intake of animal foods, a topic in need of further study.

In sum, there seems to be much useful information to be learned from the present data set. It is not likely to be simply attained, and there do not appear to be any outstandingly clear directions of influence. This may be partly explained by heterogeneity among the populations, by populations in the midst of change, or by the active presence of major intervening variables, among other possibilities. It would be important to find out which it is.

Table 92

Spearman's r's of Selected Kenya SES Variables and Toddler Energy Intake

Items Compared	r	p
Intake, Kcal/d and:		
Cattle Owned	0.02	0.69
Sheep and Goats Owned	0.09	0.02
"Other Animals" Owned	0.12	0.003
Land Owned	-0.07	0.08
Land Used	0.03	0.50
Household Head Education Level	0.12	0.004
Head/Wife Participate in Harambee	-0.05	0.20
Attend Meetings	-0.02	0.66
Children's School Expenses Met	0.13	0.001
Cash Business Owned	-0.05	0.21
Hired Labor	0.15	0.0002
Household Condition Compared with 3 Mo Ago	-0.02	0.68
Household Head Employed off Shamba	0.20	0.0001
Income/Cash Last Pay-out	0.002	0.96
Family Possessions and Structures (Assets Owned)	0.08	0.05
House Construction: cement floor	0.02	0.66
milled window frames	-0.001	0.98
glassed windows	0.09	0.03
painted window/door	-0.09	0.02

VII. DATA MANAGEMENT AND ARCHIVING

The central role of the data management process has been recognized by designers of the CRSP since its inception. Because of the importance of careful monitoring of the data collection to facilitate inter-project comparisons, contractual responsibility for assuring the quality and consistency of the data on a continuing basis was placed in the office of the Program Director. According to the inter-university contracts, each project's data management methods were to be consistent with inter-project agreements to ensure the quality and comparability of data collected and of the subsequent analyses. As stated in the Grant between the University of California and USAID: "Data collected in each project will be compared to generate conclusions about intake and function. Thus similar data from each project must be collected and described in a consistent form, to facilitate comparison. The Program Director is responsible for assuring and monitoring the consistency of data collection on a continuing basis."

A. Design of the Data Management Plan

The original data management plan of the CRSP was formulated early in 1982 by the Technical Advisory Group on Statistical Design and Data Management. Its report represented the consensus reached by representatives of each of the projects and consultants; its recommendations served as the basis of data management planning. Recognizing that the goal of the CRSP is the "completion of three highly comparable investigations in three different country settings," the TAG report stressed the need for continuing inter-project collaboration to develop comparable protocols, coding procedures, quality control, and data management. Noting that "the goal of the data management system is to permit and facilitate the comprehensive intra- and inter-project analyses," the TAG report offered a set of principles:

(1) Major concern should be quality control in data collection and management;
(2) All data transfers and changes must be fully documented;
(3) There must be redundancy in data bases to prevent loss of data;
(4) As data are rearranged on files for analytical purposes they must be "demonstrably identical", i.e. any edit procedure which modifies data in one set must modify the data similarly in the others;
(5) At all steps there must be adequate protection of the data and the interests of participating investigators;
(6) The data management system must provide for the monitoring of both data collection and transfer of collaborative data.

The original plan laid out by the TAG on Statistical Design and Data Management was for data to be collected and coded in the field, reviewed by field staff supervisors before data entry and transmitted on a regular and timely basis (within a three-month period from field to US project)) to the country institution's data entry center, preferably within the host country. Once entered onto computer files, the files and hard copy (i.e. survey instruments) would be sent to the collaborating institution in the US with copies of both computer files and questionnaires retained in the host country for later analyses. At the host country institution, some editing and quality control checks would be done, with feedback to the field if necessary. Once

received at the participating US institution, additional checks would be carried out with returns to the overseas institutions if necessary. The final stage would be the shipment of data on computer files to the data center of the ME in Berkeley.

The basic scheme of data management as described by the TAG on Statistical Design and Data Management was elaborated in a meeting of the SCB and data managers in June, 1982. A subsequent meeting of data managers representing US and overseas projects was held in December, 1982 to develop more specific data management methods. The ME data management staff, under the direction of the Deputy Program Administrator, provided assistance in the development and implementation of these plans.

ME's role, according to original agreements was to monitor the operation of the quality control system, to offer advisory support as requested and feasible and to intervene directly if a country project's data management system was not operating properly. To assist the projects and provide the monitoring function, ME provided time-tables and methods of data transfer.

For the efficient organization and handling of data in assembling the archive and for cross-project analysis by the CRSP and by others, it was agreed that a uniform and flexible data management system should be adopted across the three projects. Because of the availability of IBM equipment within every project, the Statistical Analysis System, SAS[1], was chosen as the most appropriate and readily available data management system for the CRSP. Use of this system by all projects would make it possible to standardize data files across projects and to merge files easily across time and among subjects.

B. Definition of the Basic Files and Variables

A comprehensive Data Management Manual describing the data management process for the CRSP was first completed in September, 1983 and was later modified (May, 1984) to reflect decisions of the project scientists concerning the data collection. To assist the ultimate users of the data base, procedures were developed for documenting data, mapping variables from forms to the data base, and for recording changes when errors were detected. This manual was the primary document describing data management protocol for the CRSP.

Underlying the data management principles was the agreement that there would be a core data set collected by each of the three projects to ensure uniformity in data forms, coding systems, and the structure of organized files. To assist all future users of CRSP data, the data archives of each project, including computer files and documentation, were to be made as uniform as possible. Because it was anticipated that many users would want to select files by area across projects, (e.g. morbidity, psychology), whenever feasible, data would be stored in similarly formatted files across projects. Thus, the Data Management Manual contains sections detailing the structure of each of 25 basic files, and the basic variables to be included in each file.

---

[1]SAS Institute Inc., Cary, North Carolina.

Basic variables are those that were selected by the SCB as crucial to the hypotheses of the CRSP, and that were to be collected by all projects.

When major changes took place in the research design in May and June 1984, protocols and variables were simplified to conform with these design changes. The portion of the Data Management Manual defining the basic variable was prepared as a separate document in September, 1984, under the title, "Basic Data Set". The Basic Data Set reflected the final sets of Basic Variables, in most cases a subset of the original ones. At the May, 1984 SCB meeting, the original hypothesis about school age children was omitted, but collection of schooler variables was continued by all projects - the schooler variables are considered Supplementary Basic Variables. In addition, Common Variables were defined as those being collected by at least two of the projects.

The 1984 Data Management Manual and Basic Data Set became the basis for which the remainder of work in the CRSP was carried out. In the second year of Phase II, additional changes were made according to the wishes of investigators and with approval of the MF.

The variables defined in the Data Management Manual provide the structure for the CRSP data files. Although no project was able to archive the Basic Files exactly as specified by the manual, all projects used these files as guidelines for the collection and construction of their specific files. To facilitate cross-project comparisons, each project provided "maps" that specified which project variables corresponded to the Basic Variables. These maps are archived as part of the CRSP Data Documentation. At the close of the data archive, November 30, 1987, Kenya Project had archived a complete set of maps, but those received from Mexico and Egypt Projects were incomplete for some files - for these files, it would be necessary to consult the field manuals and forms in order to determine the location of the Basic Variables.

To further facilitate this cross-reference, a Master Index of Basic Variables has been developed by the MF data management staff to provide a summary of the Basic Files and Variables that were archived by each project (Appendix F). The Master Index is complete for all files and variables and provides a comprehensive overview of the status of the Basic Variables for all three projects. Although this index has been reviewed by project personnel, the exact interpretation of equivalence was based to some extent on the judgment of MF data management personnel - in particular, in those cases when mapping variables had not been supplied by a project. Researchers wishing to compare variables across projects would be well advised to examine carefully the relevant documentation before concluding variables are equivalent. The Master Index should be utilized as a guide to possible cross-project comparisons, not as an indicator of exact equivalence.

C. Project data management

Each project's data management staff consisted of a data manager and programmers in the US institution and a data manager and data entry personnel in the host country institution or field site. Working with the data management staff in each project, a faculty member in statistics or

biostatistics served as senior statistician and provided statistical expertise to investigators and data management staff.

Data management activities included all stages from field to final archive: data entry and transfer at field site, data cleaning, editing, documenting and some simple analysis. These activities began with the processing of preliminary data in Phase I and continued through Phase II. Since decisions about design and protocol development were the responsibility of the PI's of each project, it was recommended that data managers participate in the design of data capture instruments to assure that coding or the direct entry of data could take place efficiently. In the Kenya project, this was the case; the project data manager did help to design the survey instruments which facilitated later work with data files. It was also recommended that data managers work closely with project statisticians to carry out analyses for project co-investigators.

It was agreed by all participating institutions and investigators that copies of all data tapes for each project would be retained by the host institutions overseas and by each of the collaborating US institutions. For each project, completed survey instruments were to be kept at the overseas institution with copies sent to the participating US universities for that project. Once entered, computer files were to be copied and shared within each project and with the ME in Berkeley. As data files were received at each participating institution, quality control checking was to be carried out in order to locate out-of-range values, to monitor missing data, and to confirm documentation by individual and household identifiers. When errors were detected, questions were to be returned to the field where the field supervisors traced the sources of the difficulties. Error correction forms were to be employed for the recording of any changes made once data were entered on the computer.

As work proceeded into Phase II, some modifications in the data management process were adopted as appropriate to each project's conditions at the field sites and host-country institutions. The following are the general procedures followed:

1. Egypt Project

The Egypt Project had two Data Management Units (DMUs) - one on site at the Egypt Nutrition Institute in Cairo and one at the University of Kansas Medical Center in the United States. The initial intent was for all data to be entered, verified, and corrected at the Cairo DMU and transferred on tapes to the Kansas DMU. In fact, due to delays in establishing the Cairo DMU, the Kansas DMU entered all data collected through July, 1984. Data for the remainder of Phase II were entered in Cairo according to the original plan. Two exceptions were the coding and entry of time allocation data and child caregiving data by the Kansas DMU for all these data that were collected.

Responsibility for final verification of the data and organization of the data into SAS files resided with the Kansas DMU. The staff at Kansas followed the specifications from the Data Management Manual and the Basic Variables List with great care, and all Basic Data archived prior to July, 1987, were in SAS

format and used the Basic Variable names. During this time, Egypt was the only project to achieve conformation with the specified format.

Since archiving the data in Basic File format required substantial reformatting of the field data, the Kansas DMC realized the necessity of also archiving the field data. In several cases, these field data were converted to a "raw SAS" format - SAS files that contained all the original variables. For the remaining files, the data were archived exactly as received from the field (not converted to raw SAS). Thus, most of the Egypt Project Basic data exist in two formats: as SAS Basic Files and as either raw SAS or raw data files.

Data flow from Kansas to the other US sites and to ME proceeded on a regular basis throughout Phase II. However, unanticipated delays in the field and in the data entry in Cairo substantially affected data flow from Egypt to Kansas, so that substantial portions of the data arrived toward the end of the time that had been allocated for data management activities in Kansas. Since the Kansas DMC was methodically following the protocol specified in the Data Management Manual, there was insufficient time to properly prepare the files. At the time the Kansas DMC was closed (June 30, 1987), a substantial portion of the data were still in Kansas in raw form (not converted to SAS Basic variables). With no personnel to perform this conversion, much of the remaining data were archived as received from Egypt (as raw tape or diskette files). Most of these raw files were archived in October and November, 1987, with minimal documentation. In particular, new data were not clearly identified from data which had already been archived in a different format, nor could corrected data be distinguished from original data. Therefore, researchers using these data will be required to merge files in multiple formats and scan carefully for duplicate records. At the time of the close of the archive (November 30, 1987), a number of files had not been received by ME. The majority of these outstanding files were received during December, 1987, and January, 1988, and were added to the archive. However, it was not possible to perform the usual ME verification procedures on these files, and they were archived and documented exactly as received.

2. Kenya Project

The majority of the data for the Kenya Project was entered directly onto computer-coded forms in the field and then key entered and verified in Nairobi by a private computer firm. All data were range-checked, and possible errors were returned to the field site for correction. Once corrections had been entered, the data were written on tapes and sent to UCLA.

The Kenya method of assigning every field form a form number and entering this number on each data record simplified the data management procedures. The original field forms could be used as code books in most cases, eliminating the need to further document files as they were archived.

Kenya Project chose not to use SAS as the data management language, and thus archived all files in raw format. A complete set of mapping documents were also archived to provide a cross-reference with the Basic Data Set.

Kenya was the first project to complete its data archive. All field data had been received in Berkeley by mid-1986, with additional corrected and re-organized files received by the close of the archive.

3. Mexico Project

The Mexico Project data were entered in Mexico, primarily by keypunch professionals at the Mexican Public Health Department in Mexico City, although some early data were entered at the field site using microcomputers. The Nutrition Institute's computer staff and IBM/Mexico also helped in preparation of the files. After data entry, listings of the data were sent to the field headquarters for manual verification against the original forms. Corrections were indicated on the listings and returned to Mexico City. This methodology provided a data set that conformed almost exactly with the field forms. However, since range checks were not performed at the time, out-of-range values that were recorded (incorrectly) on the field forms would not be identified.

The corrected data were then sent to the University of Connecticut, where they were either forwarded to ME as received, or in some cases, reformatted into SAS files. Extensive error checking was performed by the University of Connecticut data management staff on the dietary variables. Virtually none of the Mexico Project archived files are in the format specified by the Basic Data Set, although correspondence between the Mexico variables and the Basic Variables can be determined in most cases from the mapping documents provided by the University of Connecticut.

Numerous delays in setting up the data management procedures in Mexico resulted in corresponding delays in data flow from Mexico to the University of Connecticut. Initial attempts to convert the raw files into SAS Basic Data Set format were abandoned in favor of archiving the data in a timely manner. By the close of the archive in November, 1987, all data sent to the University of Connecticut had been received in Berkeley.

4. Comments on Project Data Management

Although the critical role that data management would play in the proper completion of the CRSP was recognized by all project personnel, the magnitude and scope of the work to be performed was not uniformly appreciated. The mere size of the data sets precluded most types of manual verification - for example, the dietary data exceeded 200,000 records for each project. Due to unforeseen complications and, in come cases, to lack of experience in handling such large quantities of data, the time and staff needed for data management were uniformly underestimated.

The initial problems arose during the process of data entry. In some cases, the forms were not designed for direct data entry, and the data required transfer from the original forms to entry forms - clearly a time-consuming and error-prone process. Furthermore, finding the personnel in the host countries to key enter the quantity of data required extensive searching, and in some cases, use of multiple methods of data entry - again, a time-consuming process. All projects eventually overcame these problems and found efficient

means of entering their data, but at the expense of substantial delays in the data flow from the host countries to the United States.

An additional set of problems arose when trying to execute the proper methods of data validation. All projects agreed to range check their data in the host country, in order to detect unreasonable or impossible values and to correct these errors prior to archiving. However, the magnitude of the task of investigating every suspicious data point was beyond the capabilities of any of the field data management staffs. In reality, all three of the project data sets archived in Berkeley have some out-of-range values. Much cleaning remains to be done, particularly of files that were not used during investigation of the initial hypotheses. Future researchers must be prepared to perform their own range checks, and in some cases, return to the original field forms for data verification, when analyzing CRSP data sets. The number of out-of-range values and logical errors is remarkably small, however, and attests to the substantial efforts of both the host country and US data management teams in detecting and correcting errors. These efforts also had a cost in time delays, however, and contributed to the delays in data flow.

Thus, due to the magnitude of both the data entry and data verification tasks, the data management process took substantially longer than was initially predicted. In retrospect, however, the time taken for this task was not unreasonable, and the initial time estimates appear unrealistic. Setbacks are inevitable in a project of the size of CRSP, and insufficient time was allowed for these occurrences. The data management teams uniformly showed great dedication to their tasks, and the quantity and quality of the resulting data files are remarkable.

D. ME Data Management

At Berkeley a data manager and two to four programmers comprised the staff which prepared the data management organization and monitored and archived the data received from the projects.

ME data management staff carried out the following activities: received data files in SAS and raw form from the three projects, stored data on tape and disk for use on the Berkeley computer; monitored the quality of data received, computed simple descriptive statistics by all major variables with files linked by ID numbers; analyzed specific sets of data in response to questions raised at EEP and SCB meetings; prepared the analytical framework of research hypotheses; developed model tables; and communicated with project statisticians and data managers. During Phase II, substantial time was spent in providing guidelines to ensure that projects could prepare and format data in a uniform manner. In the end, much of this effort was to no avail since the projects were generally unable to conform to the specified formats. In retrospect, the requirement to reformat the data into SAS files was perhaps unrealistic. Data flow would have proceeded more expeditiously if projects had been encouraged from the beginning to archive field data as received from the host country. ME data management staff might have invested their time better in fully understanding the field data collection methods, and then developing more extensive cross-project comparisons.

Much of the work during the final year centered around the preparation of the final data archive, since the majority of the data arrived during this time. The responsibility for gathering the data and corresponding documentation for each of the three projects was assigned to a specific programmer who prepared the data reports and files for the archive. The MF Data Manager had the final responsibility of ensuring uniformity of the archive format.

Frequent communications between the ME data management staff, the US institutions' data managers and each of the projects' staffs in the US and overseas insured that there was continuing close coordination of efforts. Meetings of data managers representing the three projects and the ME data management staff encouraged the closer cooperation among projects and the opportunity for sharing technical advice.

E. The Final Archive

The final preparation of the data archive of the three project's data was an explicit responsibility of the Management Entity as stated in the USAID Grant. The three projects' data files and documentation were monitored on a continuing basis during Phase II and assembled at Berkeley at the conclusion of the CRSP.

A partial list of the files comprising the final data archive from each of the projects is shown in Tables 93 through 95. These tables give the names and number of records for the files which contain Basic Variables. The file names are in alphabetic order and correspond to the names used in the Master Index of Basic Variables (Appendix E). A complete list of all archived files is available as part of the Data Archive Documentation (not included as part of this report). This documentation includes detailed information on the name and tape location of every file in the archive. Users wishing to access the archived files will need to obtain a copy of the Data Archive Documentation (for further information about the archive format, see Table 96).

Table 93

EGYPT BASIC FILES *

DESCRIPTION OF FILE	NUMBER OF RECORDS
ANTHROPOMETRY	
Targets (includes pg, lact, infant)	9320
Non-Targets	1112
CHILDCARE/SANITATION	2178
ENTRY/EXIT/CHANGE	230
FOOD INTAKE	
Individual nutrients	15901
Breastfed toddlers	2799
Missing meals for men	4763
Individual food items	396262
Nutrient data base	992
LAB ASSESSMENT/IMMUNOLOGY	
Biological	532
Hematology/urine	3961
Immunology	829
Parasitology	1809
Water microbiology	229
LACTATION/INFANT FEEDING	685
MEDICAL HISTORY/PHYSICAL EXAM	
Medical history	1327
Physical exam	49
Medical history/physical exam	5902
MORBIDITY	
Weekly recall	82672
Illness episodes	3566
PHYSIOLOGICAL STATE OF FEMALE	5150
PREGNANCY OUTCOME	
Pregnancy outcome	124
Dubowitz	110
PREGNANCY MONTHLY SURVEY	220
PSYCHOLOGY, INFANTS	
Infant cognitive	32
Infant behavioral	246
Brazelton	101

Table 93 (cont.)

EGYPT BASIC FILES

PSYCHOLOGY, TODDLERS	
Toddler cognitive	293
Toddler behavioral	2586
PSYCHOLOGY, SCHOOLERS	
Schooler cognitive	202
Schooler classroom	543
Child behavior ratings	137
PSYCHOLOGY, ADULTS	226
REPRODUCTIVE HISTORY	334
RESTING METABOLIC RATE	798
SANITATION/HYGIENE	
HH sanitation/hygiene	933
Toddler HH sani/hyg	580
Toddler indiv sani/hyg	575
SES	
Household SES - scores	312
Community SES	9
Individual SES	2188

* Egypt Project data have been archived in multiple formats: as SAS Basic Files, SAS raw files, and raw files based on the field forms. The final Egypt Project archive consists of approximately 300 data files and 150 programs. The files and record counts shown in this table represent primarily the SAS Basic Files that were archived prior to June, 1987, and may exclude some records belonging to these files but which were received later in differing formats.

Table 94

KENYA BASIC FILES *

DESCRIPTION OF FILE	NUMBER OF RECORDS
AGRICULTURAL CROP HH QUESTIONNAIRE (671)**	6041
ANTHROPOMETRY (211)	29153
CAREGIVING ACTIVITIES (641)	30969
CENSUS UPDATE (612)	11193
CLINICAL EXAM SUMMARY (323)	2301
DEPRESSION, ALCOHOL CONSUMPTION (951)	499
DISABILITY AND CHRONIC DISEASE (301)	727
DROUGHT QUESTIONNAIRE (663)	169
ENTRY/EXIT/CHANGE (611)	1062
FOOD INTAKE	
Individual nutrients	39192
Household energy ratio	8421
Nutrient data base	180
Standard recipes	164
Recipe nutrient data base	42000
Meal preparation summary (114)	42226
Consumer unit (115)	37153
Target/non-target foods (116)	27723
LABORATORY	
Physiological samples, Embu (341)	2067
Immunology and HH water contamination (342)	1025
LACTATION QUESTIONNAIRE (421)	867
MARKET SURVEY, SUMMARY (681)	131
MORBIDITY	
Individual four weekly report (313,314)	11694
Household four weekly report (315)	7211
MORTALITY NOTIFICATION (331)	4
PREGNANCY	
Pregnancy survey (412)	3028
Pregnancy outcome (413)	138

Table 94 (cont.)

KENYA BASIC FILES

PSYCHOLOGY, INFANTS	
Cognitive infant, summary (511)	276
Infant interaction (512)	5947
Infant BNAS (Brazelton) (513)	132
PSYCHOLOGY, TODDLERS	
Cognitive toddler, summary (521)	867
Toddler interaction (522)	10404
Cognitive toddler, summary, 30 months (523)	218
PSYCHOLOGY, SCHOOLERS	
Cognitive schooler, summary (531)	447
Schooler classroom observations (532)	516
Schooler playground observations (533)	1465
School attendance/performance (534)	211
PSYCHOLOGY, ADULTS	
Cognitive adult, summary (541)	622
Adult literacy test (542)	408
REPRODUCTIVE HISTORY (411)	462
RESTING METABOLIC RATE (711)	4288
SANITATION AND HYGIENE (621,622)	1714
SES	
Socioeconomic (631)	1373
Household economic questionnaire (662)	338
Subchief SES	273
TIME ALLOCATION (661)	12465
WEATHER DATA (651)	566

---

* Kenya Project files were archived as raw files, reflecting the format of the field forms. The final archive consists of 62 data files and 34 programs. An additional 96 data files were also archived by the data management staff at UCLA. These files represent reorganized raw files and some analytic files. The record counts shown in the above table are for the original raw field records.

** Numbers in parentheses refer to the field form number.

Table 95

MEXICO BASIC FILES *

DESCRIPTION OF FILE	NUMBER OF RECORDS
ACTIVITY	
Women	56276
Men	6786
Family	13846
Observations on women, toddlers	65076
Harvest	259
ANTHROPOMETRY	
Targets	6751
Non-Targets	699
BLOOD ANALYSIS	
Hematology	537
Immunology	536
Biochemistry	1755
CENSUS: BASAL	7134
CLIMATE	1151
CLINICAL HISTORIES	
Women (13 sections)	10621
Pregnant/lactating (14 sections)	33003
Adult Men (14 sections)	9796
Pediatric (14 sections)	39334
Subsequent pediatric history (5 sections)	850
FECAL ANALYSIS	2618
FOOD INTAKE	
Individual nutrient intake	20736
Diet recall	309104
Nutrient data base	520
Weaning dates	326
FOOD PRICES	7195
MILK	
Crematocrit	253
Milk collection	273

Table 95 (cont.)

MEXICO BASIC FILES

MORBIDITY**	
Mexico method	36522
Additional morbidity data	5872
Summary by calendar month, Mexico (analytic)	33165
Summary by calendar month, all (analytic)	5850
Summary by age month, Mexico method (analytic)	6320
MORTALITY	603
PHYSICAL EXAMINATION	
Adults	17841
Women with newborn infants	6038
FNODUCTIVITY	4345
PSYCHOLOGY, INFANTS	
Brazelton	61
Infants, 3 months	95
Infants, 6 months	99
Bayley motor, 6 months	97
Bayley mental, 6 months	97
IBR, 6 months	97
Mother/child ratings	97
PSYCHOLOGY, TODDLERS	
Bayley motor (18, 24, 30 months)	307
Bayley mental (18, 24, 30 months)	307
IBR (18, 24, 30 months)	307
Bayley language (24, 30 months)	213
Bayley scores (18, 24, 30 months)	315
PSYCHOLOGY, SCHOOLERS	
Teacher Evaluations	188
Intraversion/Extraversion	188
Drawings	133
Classroom	213
Playground	270
Cognitive	292
PSYCHOLOGY, ADULTS (COGNITIVE)	452
RESTING METABOLIC RATE	
RMR	1930
RMR (analytic)	1821

Table 95 (cont.)

MEXICO BASIC FILES

SANITATION/HYGIENE	
Sanitation/hygiene	8920
Five sanitation measures	235
SES/MIGRATION	
SES	545
Migration	543
SES scores	331
SOCIOCULTURAL	588
TARGET INDIVIDUALS	1068
URINE ANALYSES	9085
WEANING DATES	326
WEEKLY HOUSEHOLD FOOD USE	86494

---

* The Mexico Project archive consists primarily of raw data files, reflecting the format of the field forms. Several files have also been put into SAS and reformatted as analytic files - these files are indicated above with "analytic" in parentheses. The complete archive consists of approximately 100 files and 120 programs.

** Morbidity data were collected using two methods: Mexico method recorded illnesses that were present on the day of the interview; beginning in September, 1985, additional data were collected to reflect illness not present on the day of the interview. See the Morbidity section of this report and the Mexico Project final report for more details.

Table 96

CRSP ARCHIVE FORMAT

There are three media for the archive: hard copy, tape, and microfiche. The archive is organized around the project files, as opposed to the Basic Files. That is, the data are in the archive in the same format as received from the projects, and there is generally a one for one correspondence between the documentation (on microfiche and hard copy) and the data files. The master indexes and project maps provide the cross-reference between the project files and the Basic Files.

The complete Data Archive (files and programs that have been archived in Berkeley) is on magnetic tape. The Data Archive Documentation is primarily on microfiche, although several critical documents are also supplied as hard (paper) copy. Following is a list of the types of information that are included in the archive:

Hard copy:

Detailed documentation on the contents of the tape archive (filenames, number of records, record length, etc.)

Master indexes (as developed at Berkeley) - appear as Appendix E of the ME final report.

A detailed index of the contents of the microfiche.

Tape:

All "final" files sent to Berkeley. When possible, similar data have been combined, e.g., if multiple files of the same format were received, the records were combined into a single file. There has been no attempt to reorganize any of the data files or combine records that were received in multiple formats.

In general, the tape files contain the following categories of data:

Basic data files.

Project specific data files.

Analytic files (including the longitudinal target ID files developed at Berkeley, plus any other files developed at Berkeley or by the projects that might be useful to other researchers).

SAS programs (including the programs written at Berkeley to generate the Data Reports and any programs sent by project data managers or statisticians).

Table 96 (cont.)

CRSP ARCHIVE FORMAT

Microfiche:

All manuals and field forms.

Data Reports (run at Berkeley) for each Basic File archived by the projects.

"Maps" from Basic Variables to project variables (as developed by projects).

Range check information (as sent by projects).

Other documents and computer listings that might be useful to other researchers.

All items shown under "Hard copy" above.

Thirteen copies of the complete archive were distributed in January, 1988:

USAID - 2 copies

Mexico Project - 2 copies (University of Connecticut, and INN, Mexico City)

Egypt Project - 4 copies (Purdue, University of Arizona, University of Kansas, and Nutrition Institute, Cairo)

Kenya Project - 2 copies (UCLA and University of Nairobi)

Management Entity - 3 copies (Master copy and 2 working copies, one in Berkeley and one in Toronto)

It was agreed that access to the files by non-CRSP researchers would be restricted for one year after the end of the project (until January 31, 1989). Enquiries should be directed to the Office of Nutrition, USAID.

VIII. TESTING CRSP HYPOTHESES: REVIEW OF THE PROJECT ANALYSES

A. Introduction

The goal of the Nutrition CRSP program was to test the generic hypothesis that food intake related, in a causative manner, to functional outcomes that were believed to have import for the development of individuals, households, and communities. A key design consideration was that this generic hypothesis and its many specific subhypotheses should be examined in different biological and cultural ecosystems. Accomplishment of the most basic goal of the CRSP calls for interproject analyses of the data collected with devoted care and effort by the three projects.

It must be stated that while the goal remains unchanged, it has not been accomplished to date. In the following pages, a brief history of the evolution of thinking about analyses, often driven by the realities of work in the field and the multiplicity of pressures on projects, is offered. Some comments on statistical issues that impact upon interpretation of work accomplished to date are presented. These may be seen both as explanation and as background to a preliminary cross-project review and comparison of findings of intraproject analyses, the closest that the CRSP program can come, at this time, to its original goal of interproject analyses.

None of what is said in these pages should detract from the work that has been, and is being, done to address within-project goals. Ultimately, the scientific and policy contributions of the CRSP will arise from a blend of detailed analyses within project data sets to develop understanding of relationships within each setting, coupled with carefully planned interproject analyses designed to establish whether observed relationships are generalized or specific to one ecosystem. The interproject analyses can assist in understanding the intraproject analyses. There is a true symbiosis between the two levels of analyses.

The regret of the present is that both levels of analyses have not progressed in tandem. The challenge of the future is to ensure that it happens. Only then can the scientific community capitalize on what has been accomplished both in aggregation of unique data sets and in the preliminary exploration of those data sets. While the original CRSP goals have not been accomplished, the ground has been prepared.

B. The Evolution of Analysis Strategies: A Hierarchy of Priorities

When the original Technical Advisory Group on Statistical Design and Data Management met, it was foreseen that the real strength of CRSP analyses would lie in the interproject comparisons. That group saw a plan to do three highly comparable investigations in three different country settings as requiring organization and standardization of the study design, as well as data collection instruments and procedures. It was assumed that variables would have common definitions (and meanings) across projects and that, fully parallel analyses could be conducted to examine comparability of relationships across projects. Conversely, from such analyses, the non-comparabilities would provide important clues as to issues that required further exploration

through project specific analyses before generalized interpretations could be offered.

Some of the original aspirations seem naive in retrospect. As has been discussed in preceding sections of this report, the variables actually collected are not always comparable. Clearly, for many, the meanings are very different. This happened in spite of sincere efforts on the part of both ME and PIs to balance field constraints, field opportunities, conceptual goals, and perceived necessities to accomplish those goals. As field experience of realities accumulated, the thinking about analysis strategies evolved. The evolution of approach was augmented by the continued urging of the EEP that statistical and analytical resources be augmented at the project level.

In 1984 (SCB, May 1984), the unanimously approved working plan was that "interproject (multiple country) analyses shall be performed at Berkeley where the collaborative data sets are stored, such analyses being guided by the SCB or in default of SCB guidance, by such advisory sources as may be appropriate. It is to be expected that SCB would be involved in interpretation of such analyses. After full release of collaborative data sets to country projects, interproject analyses may be conducted at locations other than Berkeley." The process assigned the responsibility for primary country project analyses to the projects and for interproject analyses with ME. The actual procedures whereby the two levels of analyses would be coordinated and interdigitated were not defined.

By 1985, the perception of feasible approaches had evolved and a much more decentralized approach was proposed. It was proposed by PI's, statisticians and data managers that "preliminary data analysis as well as the more formal statistical modelling and analysis of each project's data should be carried out primarily by that project. At the same time, there is an explicit understanding that it is imperative to coordinate analysis across projects in order to achieve the desired generalizations from the CRSP. It was also agreed that Berkeley should serve as catalyst for inter-project comparisons and should be a resource center and clearing-house for good statistical ideas." (Report of the Data Analysis Group, April 1985.) Two major considerations impacted on this decision. The flow of data from the field to any point of analysis was much slower than had been anticipated; the problems inherent in the "cleaning" of data and development of variables were much greater than had been anticipated. US centers were under a growing pressure of conflicting priorities between the management of data and the development of analysis strategies, exacerbated by Berkeley's repeated requests for data so that it might fulfill its mission. Further, it was becoming increasingly apparent that decisions on the treatment of data (management) required consideration of the analyses to be performed with the data. There was a perceived major advantage to integration of statistical resources and management resources at the project level. This had clear budgetary implications for the allocation of resources between ME and projects. Thus, it can be seen as a critical turning point in the program's approach to analysis.

Given the constraints of both time and funding, the SCB established priorities among the many original hypotheses of the CRSP design. These analysis

priorities would establish working priorities for selection of portions of the data sets to be "worked up" first, as well as providing a basis for the collaboration between analytical staff working at the levels of the individual projects. The strategy, then, was to conduct analyses in parallel at the project level with no actual analyses to be conducted in Berkeley. The role of ME would be to monitor analyses, foster collaboration among analysts, and ultimately, receive and interpret across projects the analyses conducted within projects. This approach, including the formation of a Data Analysis Group, was reviewed and strongly supported by the EEP.

To test the system of full coordination, all projects participated in an exercise involving the production, at project level, of standardized summary statistics and simple contingency tables. The major benefit of this exercise was the demonstration to all concerned that indeed variables differed across projects and that collaborative analyses would require a great deal of effort and care in interpretation. The Data Analysis Group worked to develop a series of analytical models, focusing upon growth and morbidity, in a spirit of true collaboration and sharing of experiences. The early stages of the process carried the markers of potential success.

However, the pressures of time and funding became insurmountable. Data flow problems for two projects held back analysis development and as the data sets became available for true analyses, there was insufficient time for the collaborative process to operate. The net result was that while each project addressed in their analyses the priority questions established by the SCE, they did not do so in a uniform way. The foreseen integration of project level analyses into an interproject analytical interpretation was not feasible.

It was the conscious decision of the ME, given the circumstances that had evolved, that the most important role it could play at the end of the CPSP was to examine and compare the key variables used in project analyses, both as an aid to interpreting cross project comparisons and as a basis for others to plan future analyses. That is the thrust and goal of the present report.

C. Statistical Issues Impacting on Interpretation and Planning of Future Analyses

All projects have declared that the analyses presented in their reports are preliminary. The bulk of the reports deal with descriptive statistics, an essential background to the planning of analyses. The final chapters of each report present the analyses conducted to date. It is germane to ask "what have we learned about the pitfalls and possibilities of analyses to date?" The question is relevant to both interpretation of analyses conducted and the planning of future analyses. Of major importance is recognition of the fact that different types of analyses are addressing subtle, but often fundamentally important, differences in questions. The answers obtained must be compared in the light of the analyses performed.

One generalization that must be offered is that the basic processes of data validation and refinement are not yet complete. Both projects and ME have conducted analyses that at least suggest the presence of both biases (e.g.

interviewer effects) within project data sets and differences in the meanings of variables across data sets (e.g. the morbidity variables). In the same connotation, the approaches to dealing with missing data (often by omission of cases with incomplete data at the expense of statistical power in an already limited sample size) and with outlying values (often by screening and rejection with relatively arbitrary, and sometimes undeclared criteria) are not yet refined. Although possible approaches to these aspects of data validation were discussed in the meetings of the Data Analysis Group, testing and implementation of solutions is preliminary. It is noted that the Mexico and Egypt projects have examined regression procedures to derive variables that are either missing or displaced in time and to identify outliers (see analyses of toddler weight data), the Egypt project has implemented the SAS LSMEANS analysis of covariance procedure to develop descriptors of variables adjusted for other variables, and the Mexico project developed an adjustment procedure to provide descriptive statistics that compensated for a methodological bias in morbidity data (underreporting of short duration illnesses), and the Kenya project invested heavily in devising a procedure to correct RMR measurements found to be erroneous due to machine error. Although there has been clear progress in all projects, the procedures of "data validation" will have to be taken further as analyses move into more refined stages. In particular, the impact of any adjustment procedures on those analyses must be taken into consideration. It is emphasized that effort in this direction is warranted. There are problems within the data sets, as is to be expected in any field data collection. If these problems can be understood and suitable measures taken (rather than discarding problem data), there is much to be gained in statistical power and interpretive potential.

Since project reports present a considerable volume of descriptive statistics as a background to analysis (by prior agreement), there is temptation to interpret these presentations. Caution is warranted, overinterpretation is a constant danger. While pairwise comparisons are useful descriptions of relationships within a data set, their clear limitation is the inability to take into account multiple relationships or to examine direction (causality) of associations in a very effective manner. In tables of correlation matrices, the correlations themselves are often correlated. This, and the associated issue of establishing criteria of statistical significance in the presence of multiple tests, make accurate biological interpretation difficult. A further issue to be considered in the application and interpretation of correlation and regression analyses is the impact of "measurement error". This can attenuate the relationships seen (see discussion of this effect within the food intake variables in Section VI.D.9) with the net effect that the observed associations do not portray correctly, or perhaps even detect with statistical significance, the underlying biological relationships.

The major advantage of the rich data sets collected as part of the CRSP is the possibility of precisely defining the role of a number of factors in the nutrition, morbidity, and growth of individuals. Complex relationships cannot be adequately explored with pairwise statistics, regardless of the number of pairwise statistics calculated. Although applying multivariate techniques (preferred in most situations but not all) requires careful attention and can create subtle problems in both the application and interpretation, it has the potential to exploit the very richness of the data sets collected by the CRSP.

The point to be made at present is merely the limitation to, and danger in overinterpretation of, simple correlation and regression analyses.

The statistical model employed most frequently in the project reports is a multivariate linear model. In one way or another all projects have attempted to examine models that take some account of the longitudinal nature of the data collected and hence have the potential to infer time-related sequences, the direction of relationships. From a review of project reports, two issues arise from the application of these techniques. The first is conceptual. It is not clear in all cases that there was a biological, as well as statistical, logic driving the design of the model. Where this occurs in the reports, it may be as much or more a matter of the narrative description than the original strategy. The second is both statistical and conceptual. In the inference of a time relationship, the assumption is that the lagged variable does not persist into the current time (that is, that it is a true variable rather than a state descriptor).

The first problem may be illustrated by the application of stepwise regression procedures. These select among variables offered to develop the "best" prediction of the dependent variable. A variable may be excluded from the equation for two reasons. The test variable is unrelated to the dependent variable or the test variable is related to variables already in the equation (it explains no additional variance and has little predictive worth). The order of entry of variables affects the final results. Unless there is a careful consideration of variables that should be in the model (in terms of the hypotheses under consideration), the results may be very difficult to interpret. In many instances, projects examined multiple models, with different combinations of variables, to take these effects into account. From the discussion of "measurement error" terms above, it might be suggested that these be taken into account in deciding whether variables should be forced into regression models and in interpreting the outcomes of those models. Ultimately, the definitive statistical models will have to be based upon explicit formulations of the CRSI (or other) hypotheses translated into the variables that have been collected.

The second problem may be portrayed by two examples drawn from project analyses. In the Kenya project, relationships between dietary variables and outcomes are apparent only when the dietary measures are taken as averages over relatively long periods of time (at least three or more, usually six or 12 months). This may be because of the high variance ratios in that project's dietary data (see discussion in section VI.D.9). In the Egypt and Mexico projects, associations were, however, noted across three month or, surprisingly, even one month periods. In interpreting the observed relationships, it may be that time is so extended that a state condition rather than a deviation from that condition is what is being examined; the interpretive power would then be blunted. In some projects, an analytical strategy involving moving sequentially the averages of the needed three- or six-month time blocks has yet to be tested. In two projects, Kenya and Mexico, measures of diet quality relate fairly consistently to a number of the outcomes. In the two projects the measures are different. The observation of different measures in the two projects leads to the necessity to ask whether they may both be serving as proxies for an as yet unidentified, perhaps state

variable rather than a specific dietary variable entering in a causative relationship. In each case it will require more intensive analyses to follow up on the clear clues provided by the preliminary analyses.

In the earlier stages of the CRSP, in discussions of the Data Analysis Group and in discussions with members of the EEP, more elaborate models, often involving carefully considered transformations of data and sometimes based on experiences in econometric analyses, were discussed. Clearly such models are still a part of the thinking of project level statisticians and PIs. That they have not been presented in the reports is the result of lack of time. It seems clear, as will be discussed in a later section, that food intake does affect the outcome variables of original interest to the Nutrition CRSP program. However, the effects are subtle in the populations under study and it will require very careful, well conceived, statistical analyses to disentangle the relationships from the many associations and confounding factors. The data sets are sufficiently rich that there is real prospect for this to be accomplished if strong teams that combine statistical and biological/social science skills can be maintained, or assembled de novo, and adequately funded.

D. Review of Reported Analyses: Where to from Now?

1. Perspective

By prior agreement, each project focused attention on three areas of CRSP hypotheses: the effect of food intake on morbidity (and physical growth) in toddlers, the effect of food intake on psychological development (and social behavior) in toddlers, and the effect of food intake on pregnancy outcome (and lactation performance).

In the following pages, a very preliminary review of the results of the project analytical chapters is presented. The essential purposes of this review are to identify commonalities of findings where they exist (and thereby begin the process of generalization) and to identify differences where they appear to exist and either attempt to explain them in terms of methodologic differences between the projects or begin the process of identifying specific questions that must be addressed in future analyses. Interproject comparisons, without additional analyses, are difficult since the actual analytical strategies applied by the individual projects were not uniform and, in some instances, the variables have important differences in meaning. These matters have been discussed in previous sections of this report.

2. The effect of food intake on physical growth and morbidity

a. Physical growth and development. In the conceptualization of the CRSP, physical growth was seen variously as a direct outcome of the level of feeding, as an outcome effect of morbidity and hence a measure of the import and impact of morbidity, or as an index of all that had gone before. It was not seen as a proxy measure for intake as has been the case in many other studies. A major thrust of project analyses relating to growth can be seen as an attempt to dissociate these various possible dimensions and hence to better

understand what the variable means and how it might be used in future analyses.

At the level of correlation analyses, all projects found associations between measures of intake and achieved body size. The relationships were not always consistent within projects or across projects. Nevertheless there remains an underlying basis for assuming that there is a relationship expressed in broad terms.

The individual projects adopted different analytical strategies in their further exploration of these relationships. The Egypt and Mexico projects adopted variants of an autoregression model with lagged food intake being used as a predictor of later weight or height. The lag periods were generally short (three months or less). In this model, Egypt could detect no consistent relationships (although it was reported that when an earlier intake data base was used, some relationships were evident; there is obvious interest in rerunning the analyses when the final food composition data base becomes available). The Mexico project could not establish relationships between energy intake and growth measures using this model. In the Kenya project, for reasons discussed earlier, it was necessary to use intakes averaged over longer periods of time and therefore to ask a somewhat different question. Inconsistent relationships between energy intake level and achieved body size or growth measures were seen. Conversely, both Kenya and Mexico reported more consistent relationships when "diet quality" variables were used. In Kenya, fat intake of the toddler or mother were the most consistent variables. In Mexico, measures of the consumption of animal source foods by the toddler were the consistent variables (animal energy as Kcal/d or as % total energy; animal protein as g/d or as % total protein). In the Kenya analyses, a disturbing observation was that measures of protein intake tended to carry negative associations at the same time that fat intake carried positive associations. That is, there was no clear commonality across the two projects except that measures other than energy seemed to emerge as stronger predictors. While this might be explained in part by consideration of variance ratios associated with the different measures in Kenya, that does not seem to be a viable explanation for the Mexico project (see discussion in section VI.D.9).

Two complications of analyses were uncovered by projects. Both appear to operate across projects. First, total food intakes (estimated energy intakes) do not relate directly to existing body size when age and sex are controlled. When intake is expressed per kg body weight, a conventional nutritional approach, negative correlations may be generated with both present and future body weight (since present weight is a strong predictor of future weight). This phenomenon was explored in detail by the Mexico project. An implication of this is that the preferred approach to controlling for body size would be inclusion of weight or height as a variable in regression analyses rather than expressing intakes per unit body size. The second phenomenon is that children who are small at 18 months of age tend to show greater growth rates (same or greater absolute weight changes; greater changes relative to existing weight). Again, this can confound relationships since associates of low initial weight may appear as predictors of higher growth rates.

Projects examined the relationship of reported morbidity with growth outcomes. In both Egypt and Mexico, diarrhea was the class of morbidity that seemed to influence physical growth. The effect was not strong but it was consistent and persisted after weight, sex, and age were included in autoregression models. The Kenya project also reported low order associations between proportion of days with diarrhea and the slope of weight gain across the full year. The Mexico project undertook an interesting exploration of the relationship between stability of weight and the slope of weight gain. This falls under the hypothesis that impaired physical growth may reflect the accumulation of frequent episodes of temporary impairments of growth. It was demonstrated that the root mean square error of regression analyses of weight on time or age correlated negatively with the slope of the regression thus supporting the hypothesis. However, when this was followed with additional analyses, it was found that children with higher intakes and with less diarrhea tended to have greater variability of weight. Although the observations remain without explanation at present, the initiative of the Mexico project is worthy of pursuit. Specifically it may be of considerable interest to explore the predictors of differences in within-person variance (stability or instability) between subjects.

A clear and important finding from all projects was that weight and height at 18 months were very strong predictors of the same variables at 30 months. When the Mexico project included initial weight, sex, and age in the autoregression model, about 80% of the variance in weight in the subsequent month was explained; in Egypt, 88% of the variance was explained. In Mexico, about 50% of the variance in weight at 30 months was explained by weight at 18 months alone; in Kenya, this was even greater. These observations have at least two important implications. The expected effects of dietary variables, or morbidity, must be quite small and will be difficult to demonstrate. Second, weight or height at age 18 months is the outcome of the effect of all variables that have gone before. Since at least some of those are "state" variables that continue essentially unchanged, care must be taken to separate their effect on 18 month weight and on weight changes after 18 months. The "state" variables may themselves be correlated making it difficult to discern which are truly effectors and which are proxies for other conditions. After examining relationships among dietary patterns and body size parameters that seemed to change with subtle differences in the mode of expression or analytical approach, the Mexico project commented that "the most probable explanation for these relationships is that size of the children is primarily determined prior to 18 months of age and that the dietary pattern and quality that is associated with size at 18 months is continued through the 18 to 30 month period."

Returning to the "diet quality" variables, and considering their relationship to growth, the Mexico project demonstrated that dietary patterns (proportion of energy or protein from animal sources) carried strong associations with measures of personal appearance, literacy, sanitation, SES, and other variables that might be seen as describing the state of the microenvironment of the child. In regression analyses, the project demonstrated that if these variables were entered along with the intake variables, there was still an explanatory power attributable to intake, but it was greatly diminished. The Kenya project described relationships of maternal diet quality variables to

both growth and morbidity of toddlers. These are difficult to understand as true effectors of the outcomes. Again, these dietary variables associated with SES, but the Kenya project did not attempt to analyze both together.

b. Morbidity as an outcome. Mention has been made above of morbidity as an effector variable in growth outcomes. Turning to the CRSP hypothesis that intake affects morbidity in the toddler period, analyses are presented by only the Egypt and Kenya projects. The obverse question, "does morbidity affect intake" was addressed in Chapter VI.D with an affirmative answer for certain types of morbidity. The Egypt project was able to explain about 25% of the variance in days ill (total morbidity, total - skin diseases which are very long duration, diarrhea and respiratory illness were examined separately) in models including subject, age, and season. Dietary variables offered no additional explanatory power. In Kenya, 40% of the variance in low grade illness in 18-24 month old toddlers was explained by the household sanitation/hygiene measures (although in another analysis with additional cases, the correlation was much lower than this). In the subsequent six months, this measure did not enter the prediction equation. Illness of the toddler in the previous six months, maternal low grade illness and sex explained the greatest part of the variance but toddler fat intake/kg, maternal fat intake per day, and toddler animal protein intake in previous six months entered. The last had a negative sign; all others were positive. When "significant illness" was examined, energy intake/kg was a significant predictor (negative sign; see comments above) in first six months but not in next period; there the history of illness in the previous six months and body length at 18 months were the selected explanatory variables. Unfortunately, the analytical strategy adopted by the Kenya project is difficult to interpret in terms of assessing the specific contribution of a predetermined set of variables (see previous discussion of approaches to regression analyses). Nevertheless, as for the growth outcome, there are suggestions that food intake may play a role during the toddler period. Since that project reported that illness depressed food intake, the final analyses may have to direct close attention to the signs of any relationships.

c. Interim conclusions. A tentative conclusion that may be drawn from the project analyses relating intake and morbidity to "growth", and relating intake to morbidity, is that there are dietary associates, but the analytical strategies applied to date are not sufficient to clarify the exact relationships. It is clear that all that has occurred prior to the 18th month of life has a continuing effect on growth (and morbidity) after 18 months; it is essential then that analytical strategies separate more distant effects of diet and other factors from more recent effects. The demonstrable effects of morbidity are surprisingly small and seem to be restricted primarily to effects of diarrhea. With further refinement of the morbidity data bases, it may be that additional or stronger effects will be demonstrable. Since it appears that food intake is depressed during periods of fever and digestive illness, at least in the Kenya project, (no effect of morbidity was seen in a limited series of cases examined for Mexico. See Chapter VI.G) it might be expected that there would be some interaction between intake and morbidity variables. In future analyses it will necessary to direct further attention to seasonal effects. The Egypt project demonstrated clear seasonal differences in morbidity (peaks were not in phase across classes of morbidity)

and attained body size but not food intake. Reported Egypt project analyses controlled for season. The Kenya analyses also included a variable for "season", but it would appear that full analytical advantage has not been taken of the major effect of the drought period. Admittedly, this will be an analytically complex task but it is one that should be explored. A preliminary examination of the effect of the period of food shortage on toddler weights suggested a surprisingly small effect (see Chapter VI.F). Both the Mexico and Kenya projects have focused attention on the apparent role of the "diet quality" variables. It is the judgement of ME that while these are highly interesting and worthy of much more investigation, it would be premature to conclude that these diet quality measures are causally related to the growth outcomes. It will be necessary to continue and expand the examination of the relationships between these variables and other dietary and household characteristics initiated in the Mexico project; in the end it is likely that a judgmental interpretation will have to be made as to whether the diet quality variables are proxies for one or more specific nutrients that are limiting in the diet and/or are sensitive indicators of the microenvironment surrounding the early development of the child. They seem to show a relatively greater variance across individuals than does energy intake (see Chapter VI.D). Finally, the exploration of predictors of 18 month body size initiated by two of the projects should be expanded and intensified since all projects demonstrate convincingly that achieved body size at this age is the strongest predictor of later body size identified. The important policy questions are (i) what are the factors that associate with early growth failure (perhaps identified through analyses of both toddler and infant data), and (ii) is the early failure irreversible or can factors influencing growth (recovery) during the toddler period be identified?

3. The effect of food intake on psychological development

All projects collected measures of both cognitive development (Bayley and Uzgiris-Hunt tests) and of social (play) behavior. In the analyses reported, the Bayley mental and motor tests at 30 months were a common denominator across projects and constitute the focus of the present interproject comparison. A discussion of the variables has been presented in section VI.H.

Projects fully recognized that there were cultural differences associated with the implementation and application of even tests as widely used as these (projects made culture-specific modifications of the instruments and indeed of the components of the tests administered, and adopted different approaches to scoring the results; none standardized the results against North American norms). Nevertheless, there were some striking similarities in the project results.

The Egypt and Kenya projects both noted that energy intake (kcal/d) in the 18-24 month (Egypt) or 24-30 month (Kenya) period associated with the 30 month Bayley mental score (and motor score in Kenya). When the Egypt analyses were controlled for body size (kcal/kg), the relationship disappeared. Mexico reported only trials of the relationship of energy intake per kg body weight during the 27-29 month period with Bayley scores. They found no relationship. The contributions of body size and energy intake were not assessed

concurrently in the Kenya project. For reasons discussed in the preceding section, the mode of expressing energy intake (method of controlling for body size) may make important differences in the results. Across projects, it was not possible to demonstrate that total food intake (energy intake) was associated with Bayley scores independently from achieved body size. Conversely, a consistent finding across all three projects was that measures of achieved body size (length or weight) in the early toddler period or proximate to the 30 month Bayley measures were reasonably strong associates of the scores.

Failure to demonstrate persisting association between intake and cognitive measures in Kenya might be attributable to the very high intraindividual variance relative to between-subject variance of energy intake in that project's data set (see discussion in Chapter VI.D). This could not explain the lack of associations in the Egypt and Mexico data sets where this component of variance was much lower than for Kenya.

Although energy intake was not demonstrably linked to the Bayley measures when analyses were controlled for body size, both Kenya and Mexico found that measures of "dietary quality" did associate. In Kenya, fat intake in the 18-24 or 24-30 month periods was a stronger predictor of 30 month Bayley scores and subscales than was energy (note however, that the intraindividual variance of fat intake was less than that for energy intake; this may have contributed to a stronger relationship for fat than for energy). In Mexico, % of energy or % of protein from animal sources served to explain part of the variance in Bayley scores even though energy intake per kg body weight had no explanatory power. In the Mexico project, this relationship persisted when achieved length was included in a linear regression model along with the diet quality measure. Such diet quality associations were not reported out in the Egypt project analyses and may not have been tested.

It might be speculated that these diet quality measures were serving as proxies for one or more micronutrients in the diet. Iron was examined in Kenya and Mexico without demonstrating associations, but this leaves many other potential variables that might change with makeup of the diet. It is not clear what dietary component would associate with fat intake in Kenya and with animal food intake in Mexico if indeed a common nutrient variable were involved. It is not logical that fat intake per se was a biologically active variable; if energy density of the diet were the effector variable, it might be expected that ingested energy (from a higher fat, more energy dense, diet) would emerge as a direct variable.

An alternate explanatory hypothesis, apparently favored by both projects, is that the diet quality measures may be serving as a proxy for the "household environment" in which the child's development has taken place. The Kenya project noted a reasonably strong correlation between fat intake and SES but did not actually test both measures in a regression model. However, from multivariate analyses, that project concluded that child rearing variables had greater explanatory power than did fat intake and tended to displace fat intake as a predictor of Bayley scores. In the case of the Mexico project, regression analyses were undertaken with various combinations of achieved length (expressed as the Z score), animal protein, or energy as percent of

total protein or energy and their SES measure. When used without SES, both length and diet quality were statistically significant predictors of 30 month Bayley scores as was also SES. Used individually with SES, these measures continued to explain a part of the variance. However, when all three were included, the proportion of total variance explained was additive but only length remained significant as a specific term. It might be inferred that there were important, untested, interactions among these variables. This may suggest that the diet quality measure and SES jointly reflected some attributes of the household microenvironment in which the child was developing, as was suggested from the Kenya analyses.

All three projects have initiated analyses to begin exploring the precise characteristics of the household environment that may relate most strongly to the Bayley scores. The Kenya project has completed, and submitted for publication, a detailed study of home rearing conditions and Bayley scores. The Mexico project used a "child appearance" measure as a proxy for child rearing variables and found that it contributed independently of length, greatly reduced the contribution of SES, and displaced, almost totally, the diet quality measure. In all projects, there are tantalizing hints of relationships, and the beginnings of development of analytical models (see for example the Egypt project report) but not yet definitive answers.

Demonstrable relationships between morbidity (classified in various ways) and Bayley scores were quite weak in all three projects. This is somewhat surprising, since it has been generally assumed that illness is an important determinant of at least some aspects of psychological development (but perhaps only at younger ages - the Kenya project could detect an association but only for morbidity in the 18-24 month period; morbidity after that time did not seem to influence the 30 month Bayley tests). It is also surprising in the Kenya project, since it had been reported that some types of common morbidity associated with an appreciable fall in food intake; energy intake had been associated with the psychological measures. It would be of interest to undertake an analysis in which 18-24 month morbidity and food intake are examined concurrently as predictors of 30 month Bayley scores.

Although not reviewed here, all projects collected measures of toddler behavior and interactions. Analyses of these variables has begun. They may shed light on interpretation of the cognitive tests and their relationships as well as exploring a very interesting, and perhaps more important, dimension of development of the child.

There remains the consistent finding of a relationship between achieved body size and cognitive measures at 30 months (and indeed at 24 months). Traditionally this would be taken as an association between "undernutrition" and cognitive development, since the past literature has assumed that anthropometric indices are measures of nutritional status. The CRSP data sets provide some opportunity to explore this further. As discussed in Chapter VI.F., it appears that the "break" in physical development takes place after about 4 months and prior to the 18 month age at which toddler studies began. The actual age at which growth faltering begins may differ between projects. The environmental factors (including nutritional factors) that may associate

with this break are likely to differ between projects. If then all projects seem to demonstrate similar associations between the measure of growth failure (achieved body size) and psychological measures, the CRSP may provide an opportunity to begin sorting out the environmental factors that are particularly important through careful interproject analyses of both the infant and toddler periods. Until such analyses are performed, we can tentatively agree with the conclusion of the Mexico project, "It appears that....the features of life experience that lead to small size also lead to delayed cognitive development."

The results to date do not exclude food intake as having an independent effect on psychological development and behavior but they suggest that to demonstrate such an association, it may be necessary to refine the understanding of the other variables that impact on cognitive development - a major refinement of models based on such understanding may be essential and is underway at the project level.

In the words of the Egypt project, "in terms of the overall question phrased by the CRSP, the project results indicate a low but consistent relation between food intake and cognitive-behavioral development during the toddler period. However, the nature of the results is quite complex, suggesting that the relation between food intake and toddler development is not a simple one, and [analyses to date] perhaps raise more questions than they answer." Among the areas still to be explored is another of the CRSP hypotheses - that maternal intake affects the mother-child interaction and through this route the psychological development of the child. To unravel this path will require examination of, and control for, a number of other factors such as "household crowding" (measured by proximity of others during the child observation studies) and indeed the identity of primary care-givers (and reasons for the assignment of this responsibility in the household).

There remain many questions to be asked and many analyses to be performed - not just with the Bayley measures but also with the Uzgiris-Hunt measures of cognitive development and the several measures of play behavior and social interactions that were collected in these projects. These analyses must be conducted with a clear understanding of the relationships to be expected based on studies in other population groups, the nature (meaning) of the actual variables collected in the field, and the opportunities and strategies available to the analyst. There would be distinct advantage to an interdisciplinary team approach to such analyses since they will have to cross disciplinary boundaries and a breadth of understandings will be needed. The data have been collected; the questions have been phrased; the opportunities cannot be allowed to escape. The Nutrition CRSP may be one of the very few comprehensive nutrition-psychology data sets in the world that remain unanalyzed or are even being collected.

4. The effect of food intake on reproductive outcome

Conventional thinking is that women lay down additional fat reserves, if they are able, during pregnancy and draw on these reserves to meet the substantially increased energy demand of lactation. A more reasoned view would suggest that energy balance of the mother can be maintained in the face

of reproductive demand by any temporal combination of food energy intake and deposition or withdrawal of energy reserves (within some reasonable limit) coupled, where possible, with compensatory shifts in physical activity. In the CRSP populations, maternal fatness/weight increase with parity and age, clear evidence that energy intake meets or exceeds demand over time. Longitudinal data are required to determine how this balance is achieved, and if the mechanisms for its achievement have different consequences for the welfare of the mother, the infant, and the other family members.

The sample populations of pregnant and lactating women in the three countries differed in at least two respects that are thought to affect reproductive outcome. Firstly, the Egyptian women were heavier in relation to height than were Mexican women, who were heavier than Kenyan women. All three samples include some women who are obese and others who are thin, but the Egyptian and Mexican women were, on average, respectively, about 10 and 5 kg heavier for their heights than the Kenyan women (who would be considered lean but not excessively thin at 25% body fat). While the heavier body weights would increase the energy cost of activities that involve movement of body mass (and hence energy requirement for equivalent activity patterns), the tissue constitutes an energy reserve that could attenuate the effect of low energy intakes. The CRSP data cover, with varying degrees of completeness, the components energy intake and body weight change. Skinfold data are available but have not been fully utilized in conjunction with these relationships in all projects. The skinfolds appear to carry a substantial measurement error term which will blunt their utility; nevertheless, they should assist in documenting the inference that weight loss during lactation is attributable to changes in adipose tissue (and by comparing early pregnancy and early lactation skinfolds, some inference should be possible about net deposition of fat during pregnancy).

A most significant gap is the paucity of data projects were able to gather on maternal activity patterns. The relationship between maternal intake and maternal child-care and hygiene-related activities held considerable interest in the original CRSP design. It had been postulated that shortfalls in maternal energy intake might be made up through decreases in key activities. Clearly, those maternal activities that related to child care and to the maintenance of a sanitary environment could provide an important linkage between maternal intake and child health and cognitive development. The key additional variables required for these analyses are the maternal activity variables. It appears that the Kenya project may have collected the requisite information in a form that might be related to individual women; indeed in this project, it appears that information could be developed with regard to the allocation/sharing of these activities among other household members, if this is a part of the coping mechanism in the face of the energy demands of pregnancy without major increase in intake. Other projects have collected time allocation information that can address categorical questions, but such approaches to analysis are unlikely to address the original CRSP hypothesis. All projects have collected information that can be developed into indices of child care and household sanitation but do not provide for specific inference of the mother's contribution.

The second significant difference among projects is the level of food energy intake. The data (given in section VI.I. Table 87) show that pregnancy intake was greatest in the Mexico project, particularly in relation to body weight, and lowest in Kenya (41-43 kcal/kg versus 25-32 kcal/kg); intake in Egypt was higher per day than in Kenya, but approximately the same per kg body weight. Energy intake during lactation was higher than during pregnancy, the same per kg in Egypt and Kenya (about 35 kcal/kg) and much higher in Mexico (47-48 kcal/kg). Intake of pregnant women was not higher than intake of the population of nonreproducing women (it may have been less, but some refinement of analyses will be required to be certain that comparison samples do not include lactating women). Kenya and Mexico report that RMR increased during pregnancy, so energy available for activity and tissue deposition was less than in the nonpregnant state. The increase in intake during lactation (300-400 kcal/d) was about half the amount thought to be required to meet infant energy needs (9). The occurrence of a period of food stringency complicates the Kenya data set, but the relative changes are the same in all three locations.

Despite the initial differences in maternal size and body composition and the differences in energy intake between projects, pregnancy weight gain and infant weights at birth and six months of age were nearly the same (see section VI.F. and I.). Women gained little or no weight in the first trimester of pregnancy (some lost weight) in Mexico and Kenya (data for Egypt are not presented). During the second and third trimesters, Egyptian women gained 7.4 kg (equally deposited in trimester 2 and 3); Mexican women also gained 7.4 kg (4.5 kg in trimester 2 and 2.9 kg in trimester 3); and Kenyan women gained 6.4 kg (equally in the two trimesters). The 1.0 kg difference in gain between Kenya and the other two projects, if it were all adipose tissue, would represent an additional storage of about 40 kcal/d. Infant male birthweights were 3.3±0.4 kg in Egypt and Kenya, and 3.2±0.4 kg in Mexico. Average male weights at six months of age were 7.1 kg in Egypt and Kenya and 7.0 kg in Mexico. Female weights were somewhat less, noticeably more so in Mexico. Variance was about the same in the three data sets; CV of pregnancy gain was about 50% and of infant weights about 12-15%. The existence of disparate initial states and concurrent inputs but similar outcomes in several key measures presents an interesting challenge to cross-project analyses, which is yet to be faced.

Within-project analyses to date have suggested that there are demonstrable associations between aspects of diet, maternal weight gain during pregnancy, and birth weight. They have suggested also associations between weight gain during pregnancy, intake during lactation, weight loss during lactation and infant weight gain with inference of a relationship between available energy supplies (exogenous and endogenous), infant energy intake and infant weight gain. As noted in the previous discussion of project-reported analyses, there are several areas in which refinements of analytical procedures would be desirable. Those comments are not repeated.

Each project addressed the question of a relationship between food intake and maternal weight gain during pregnancy. Kenya reported a positive association between weight gain and prepregnancy energy intake, and intakes of fat, protein, and energy in the third trimester. Significant variables identified

in a multiple regression model were fat intake in the third trimester, change in skinfold thickness between the second and third trimesters, occurrence of significant illness in the second trimester (negative), and thyroid size. Various Mexico project models identified the variables maternal energy intake in the second trimester, kcal from animal foods, pregnancy complications/vomiting (negative), maternal % body fat, and SES. Neither dietary nor morbidity variables emerged in the Egypt models. In Egypt, negative correlations were reported in bivariate analyses between postpartum maternal weight (not weight gain) and intakes of energy , plant protein (a proxy for energy intake in that data set), and total protein during pregnancy; that is, heavier women ate less during pregnancy. (Heavier women were also older, see above.)

All projects confirm the expected association between maternal body weight and infant birth weight. Birth weight was not, however, significantly correlated with maternal weight gain, except in Egypt. Maternal serious illness was a negative factor in Kenya, but maternal morbidity (% of time ill) was not associated with birth weight in Egypt. Birth weight did not significantly correlate with SES in Egypt or Kenya. A multiple regression model in Mexico included SES as a third factor after maternal weight and maternal energy intake, but the model accounted for only about 20% of the variance. The strongest variable identified in a Kenya model was energy intake in the second and third trimester; all factors (energy intake in all trimesters, skinfold thickness in trimester 1, and significant illness (negative)) accounted for 28% of the variance in birth weight. Egypt models identified body weight or body mass index and maternal weight gain as being significant, but no dietary variables emerged as significant.

All projects have examined the hypothesis that maternal intake affects attained weight and length of infants and the rate of change in these parameters. All projects found infant weight at six months of age to be associated with maternal energy intakes during lactation, either in bivariate or multiple regression analyses. In Egypt, the relationship held for energy intake and for intakes of carbohydrate and plant protein, but not animal protein; in Kenya, both maternal energy and protein intake were related to infant weight at some periods; and in Mexico, energy intake during lactation and animal protein intake during pregnancy and lactation were significant correlates of infant weight. Maternal weight change (loss) during lactation was positively associated with infant weight at six months in Kenya and maternal weight at term with six month weight in both Kenya and Mexico. Weight at six months was correlated with birth weight in Egypt and Kenya but not in Mexico.

There are some interesting reversals among projects in other factors examined. SES was significantly associated with attained infant six-month weight in Kenya and Egypt, but not in Mexico. Infant morbidity was negatively correlated with six-month weight in Egypt (% of time with diarrhea) and attained length in Kenya (episodes of febrile illness). In Mexico, infant weight was positively associated with maternal morbidity and more infant diarrhea was reported in households with good sanitation scores. The Mexico project believes this to be a perception-of-illness phenomenon, rather than a true difference in morbidity. As in the case of toddlers, rates of change in

weight and length are confounded with size, larger infants having a somewhat slower rate of growth. Dietary supplementation of infants may also be a confounding factor, although much remains to be examined in this regard. Egypt found that the heaviest infants at six months received fewest supplements, but this did not reach statistical significance; in Kenya, infant supplementation was negatively associated with infant length. Provision of sugar and similar nutritionally poor foods, displacing breast milk calories, may adversely affect infant growth and development, while sparing maternal energy demand. No other factor has been identified that offers as much promise of explaining why growth falters so early in life in these populations and why so many infants die.

In the original hypotheses, maternal food intake was seen as a variable that would impact directly, biologically on infant endowment at birth, not only as to size but also immune status and cognitive potential. Maternal intake was also seen as influencing morbidity and the psychological development of the infant by the indirect behavioral route discussed above. Turning to the psychology variables, there were both measures of endowment at birth (the Brazelton measures) and developmental markers at six months (Bayley Motor Scales and aspects of the Infant Behavior Record). Dubowitz assessments (of gestational age) have been archived for most infants by the Egypt and Kenya projects; analyses in Kenya suggest that this may not be a reliable index of gestational age in the study population, although further analysis of this is warranted. Brazelton scores are available for most infants in Egypt and Kenya and for about half of the infants in Mexico. Bayley Motor and Infant Behavior Record data have been archived for most six month-old infants in Kenya and Mexico and about one-quarter of the infants in Egypt. All projects collected information on infant-caregiver interactions, an important explanatory variable for psychological development as well as a potential marker of development (the observations can and are being scored from both perspectives). In Kenya and Egypt, observational data were collected at two, four, and six months but for only about one third of expected measures in Egypt. In the Mexico project, near-complete data appear to be available for three and six month-olds. Most of these psychology variables remain essentially unexamined at present. Because of missing data or non-collection, the opportunity for cross-project analyses of the original CRSP hypothesis is limited. However, there are ample opportunities for analyses within the projects and for more comprehensive analysis in certain situations. All of this remains to be done.

On their own initiative, all projects collected breast milk samples to be used for immunologic and other assays. It appears that only the Kenya project was able to arrange for immunologic analyses (on about one-third of the collected samples). All projects report that the milk samples are being held in frozen storage. At present, it is not possible to explore the subhypothesis that maternal nutrition influenced the passage of immune factors to the infant. Projects have not yet examined infant morbidity in relation to maternal intake and caregiving activities or maternal activity patterns in general. The data available for these analyses is limited.

The analyses to date have not yet probed in appropriate depth the possible association of both food intake and outcomes with the SES variables. It is of

obvious programmatic interest to know whether the gradients across SES within and between these communities are comparable and to explore the extent to which income-related resources can influence the manner in which women cope with the energy intake-utilization pressures of pregnancy and lactation.

E. A Futuristic View: What Remains to be Explored

It was emphasized early in this chapter that, in the face of time and budgetary constraints, the SCB took a decision to assign priority to the exploration of three areas of CRSP hypotheses. In so doing, priorities were set also for the development of data bases that would be required in these analyses. In the preceding sections of this chapter, some of the preliminary analyses have been reviewed. It is absolutely clear that much remains to be done before the analyses of the selected priority hypotheses can be said to be complete. What may not have been as clear is the scope of data, and research hypotheses, that have not yet been examined. The work to date represents but a small fragment of the work that can be done. There has been a major investment of time and dedicated effort by a very large number of people. There has been a monumental contribution by the subjects in the three communities that were studied. What is now needed, to capitalize upon this investment, are equally dedicated teams of skilled investigators prepared to invest major effort in the analyses of these data and funding bodies that are prepared to support the analysis of data already collected rather than the collection of new data.

It is the sincere hope of all in Management Entity that the opportunity, indeed the obligation, to complete the task initiated by the Nutrition CRSP will not be allowed to pass.

IX. REFERENCES CITED

1. Steering Comm., NRC Study on World Food and Nutr. World Food and Nutrition Study: The Potential Contributions of Research. Chapter 2. High Priority Research. Natl. Acad. Sci., Washington, D.C., 1977.

2. Commission on Internat'l Relations, NRC. Supporting Papers: World Food and Nutrition Study. Vol. IV. Study Team 9, Nutrition. Natl. Acada. Sci., Washington, D.C., 1977.

3. Commission on Internat. Nutr. Programs, Food and Nutr. Board. Summary Report: Workshop on the Need for Research on Nutrition and Function, July 18-21, 1977. Natl. Acad. Sci., Washington, D.C., March, 1978.

4. Calloway, D.H., C. Wood et al. Final Report: Collaborative Research Support Program on Intake and Function, Contract No. AID/DSAN-C-0002, Univ. of Calif., Berkeley, 31 May 1980.

5. Beaton, G.H. Phase I, "Preliminary Survey", memorandum to CRSP, US Investigators from MF, University of California, Berkeley, February, 1982. (Documentation available in microfiche at USAID.)

6. Merrill AL, Watt BK. Energy value of foods - basis and derivation. USDA Handbook No. 74. 1955.

7. Paul AA, Southgate DAT. McCance and Widdowson's Composition of Foods, 4th rev edition. London: HMSO. 1978.

8. United States Department of Agriculture. Provisional table on percent retention of nutrients in food preparation. Hyattsville, MD: USDA. 1982.

9. FAO/WHO/UNU Energy and Protein Requirements. Report of a Joint Expert Consultation. WHO Tech Rept Ser No. 724. Geneva. 1985.

10. National Academy of Sciences (1986) "Nutrient Adequacy: Assessment Using Food Consumption Surveys" Report of a Subcommittee on Criteria for Dietary Evaluation. National Academy Press, Washington, 1986.

11. FASEB Life Sciences Research Office (1986) "Guidelines for Use of Dietary Data" Report Prepared for the Center for Food Safety and Applied Nutrition, Food and Drug Administration. Federation of American Societies for Experimental Biology, Bethesda, MD.

12. Beaton, G.H. Milner, J. Carey, P. McGuire, V. Cousins, M. Stewart, E. de Ramos, M. Hewitt, D. Grambasch, P.V. Kassim, N. Little, J.A. Sources of variance in 24-hour dietary recall data: Implications for nutrition study design and interpretation. Am J. Clin. Nutr. 32: 2546-2559. 1979

13. Duggan, M.B., Alwar, J. and Milner, R.D.G. Arch. Diseases Childhood 61: 61-66. 1986.

14. Sukhatme, P.V. and Margen, S. Am J Clin Nut 35: 355-365. 1982.

15. Energy Requirements in Pregnancy. An Integrated Study in Five Countries. Nestle Foundation Annual Report, Lausanne, Switzerland, 1986.

16. Dean, J.A.,ed. Lange's Handbook of Chemistry, 13th ed., McGraw-Hill, 1985. pp. 10-26/10-28.

17. Hamill, P.V.V., Drizd, T.A., Johnson, C.L. Reed, R.B. and Roche, A.F. (1977) NCHS Growth Curves for Children Birth-18 Years. Vital and Health Statistics Series 11, No. 165. US Department of Health, Education and Welfare.

18. Hamill, P.V.V., Johnston, F.E. and Grams, W. (1970) Height and Weight of Children. Vital and Health Statistics Series 11, No. 104. US Department of Health, Education and Welfare.

19. Whitehead, R.G. and Paul, A.A. (1984) Growth charts and the assessment of infant feeding practices in the western world and in developing countries. Early Human Development 9: 187-207.

20. Fomon, S.J. (1986) "Protein Requirements of Term Infants" in Fomon, S.J. and Herod, W.C. Energy and Protein Needs During Infancy, Academic Press, pp 55-68.

21. Martorell et al. Am J Clin Nutr 33:345. 1980.

22. Egypt Project. Six Month Report, Nutrition CRSP - October 1, 1984-March 31, 1985.

23. Durnin, J.V.G.A., Womersley, J. Br. J. Nutr. 32: 77-96, 1974.

APPENDIX A

MEMBERS OF CRSP ORGANIZATIONAL BODIES

[SCB, IC, TAGs]

1. SCIENTIFIC COORDINATION BOARD (SCB)

Egypt Project:
- Osman Galal, M.D., Ph.D. (Nutrition Institute, Egypt)
- Gail Harrison, Ph.D. (University of Arizona)
- Norge Jerome, Ph.D. (University of Kansas)
- Avanelle Kirksey, Ph.D. (Purdue University)

Kenya Project:
- Nimrod Bwibo, M.D. (University of Nairobi)
- Charlotte Neumann, M.D. (UCLA)

Mexico Project:
- Lindsay Allen, Ph.D. (University of Connecticut)
- Adolfo Chavez, M.D. (Instituto Nacional de la Nutricion, Mexico)
- Gretel Pelto, Ph.D. (University of Connecticut)

2. INSTITUTIONAL COUNCIL

Egypt Project:
- Billy Hudson (University of Kansas)
- C. John Marc, DVM, Ph.D. (University of Arizona)
- D. Woods Thomas, Ph.D. (Purdue University)

Kenya Project:
- Philip Costic (UCLA)

Mexico Project:
- Alexandra van Gelder (University of Connecticut)

Management Entity:
- Robert Edwards (Chair) (University of California, Berkeley)

3. TECHNICAL ADVISORY GROUPS

TAG on Cognition

Egypt Project: Theodore Wachs (Chair)
Kenya Project: Emmy Werner, University of Calitornia, Davis
Mexico Project: Tiffany Field, Gordon Findlay
Consultants: Ernesto Pollitt, University of California, Davis; Craig Ramey, University of North Carolina, Chapel Hill

TAG on Disease

Egypt Project: Ronald Watson
Kenya Project: Nimrod Bwibo, Charlotte Neumann
Mexico Project: Ranjit Chandra, Alfonso Mata

TAG on Nutritional Status

Egypt Project: Gail Harrison
Kenya Project: Charlotte Neumann
Mexico Project: Adolfo Chavez (Chair)

TAG on Reproduction

Egypt Project: Avanelle Kirksey, Cheryl Ritenbaugh
Kenya Project: Janet King
Mexico Project: Lindsay Allen
Consultants: Nancy Butte, USDA, Baylor, Texas; M.S. Read, University of Maryland, College Park

TAG on Food Intake

Egypt Project: Norge Jerome, Hekmat Aly
Mexico Project: Gretel Pelto (with input from Adolfo Chavez and A. Mata)
Management Entity: Doris Calloway, Chair; George Beaton

TAG on Social Competence

Egypt Project: Norge Jerome (Chair) (with input from H. Aly, Judi Ricci, and Cheryl Ritenbaugh)
Kenya Project: Claudia Carr (with input from Nimrod Bwibo)
Mexico Project: Perti Pelto
Management Entity: George Beaton

TAG on Statistical Design and Data Management

Egypt Project:	R. Hassanein
Kenya Project:	Ann Coulson
Mexico Project:	P. Pelto
Management Entity:	George Beaton
Consultants:	J. Ware (Chair), Harvard University, Cambridge Massachusetts; W. Rand, Massachusetts Institute of Technology, Cambridge, Massachusetts; W. Reinke, Johns Hopkins School of Hygiene, Baltimore, Maryland

TAG on Activity

Egypt Project:	A. Ismail
Mexico Project:	Lindsay Allen (Chair)
Consultants:	A. Ferro-Luzzi, Institute of Nutrition, Rome; E. Messer, Yale University, New Haven, Connecticut; G. Spurr, Wood Veterans Administration Medical Center, Milwaukee, Wisconsin

APPENDIX B

MEMBERS OF THE EXTERNAL EVALUATION PANEL
NUTRITION CRSP

Professor Richard Adams, Department of Anthropology, University of Texas, Austin, TX 78712

Mr. Sol Chafkin, Executive Vice President, Local Initiative Support Corporation, 666 Third Ave. (32nd Floor), New York, NY 10017

Professor Gerald Lousch, Professor of Medicine, Chief, Division of Geographic Medicine, Tufts University School of Medicine, 136 Harrison Ave., Boston, MA 02111

Mr. Philip Payne, Department of Human Nutrition, London School of Hygiene and Tropical Medicine, Nutrition Policy Unit, Keppel Street (Gower St.), London, WC1E 7HT, ENGLAND

Professor T.N. Srinivasan, Department of Economics, Economic Growth Center, Yale University, New Haven, CT 06520

Professor Donovan Thompson, Department of Biostatistics, University of Washington, Seattle, WA 98105

Dr. Marian Radke Yarrow, National Institutes of Health, 9000 Rockville Pike, Bethesda, MD 20014

APPENDIX C
MANAGEMENT ENTITY STAFF AND CONSULTANTS

BADGER, Diana - Secretary 3/86 - 11/86; Editorial Assistant, 12/86 - 10/87

BALDERSTON, Judith - Deputy Program Administrator, 6/82 - 10/84; ME Coordinator for Data Management and Analysis, 10/84-5/86; Associate Research Economist, 6/82 - 5/86

BEATON, George - Acting Program Director, 11/81 - 4/82; Principal consultant, 5/82 - termination of CRSP

BELL, Michael - Student Assistant (Senior Typist Clerk), 9/85 - 12/85

BERGS, Karin - Management Services Officer, inception (9/81) - termination

BICKLE, Amanda - Summer Intern, 6/85 - 8/85

BRODSKY, Joel - Associate Research Statistician, 4/83 - 6/83

BROWN, Felomena - Administrative Assistant, 3/82 - 7/85

CALLOWAY, Doris Hones - CRSP Program Administrator, inception (9/81) - termination

CASTELLA, Tony - Undergraduate Nutritional Sciences Assistant, 1/87 - 5/87

CATTLE, Dorothy - Assistant Research Anthropologist, 9/81 - 12/85

CHEN, Shan-Shan - Data Management Programmer, 8/84 - 2/85

CHERNOW, Harmeen - Graduate Student Assistant, 8/85 - 6/86

CHURCH, Kathleen - Student Assistant, (Senior Typist Clerk) 9/85 - 12/85

CONDON, Katherine - Business Manager, 11/82 - 5/86

CRAIG, Deberah - Principal Typist Clerk, 6/82 - 6/83

EDWARDS, Robert - Sponsored Projects Officer, inception (9/81) - termination

ELWIN, Alicia - Secretary, 11/81 - 3/82

FOSLER, Patsy - ADP Systems Analyst, 7/82 - 11/84

GEIST, Judith - Assistant Research Political Scientist, 8/82 - 11/82

HANSEN, Hyrum - Administrative Analyst, 10/83 - present

HESSELMAN, Henrik - Student Assistant, 7/86 - 5/87

HOLBROOK, Janet - Assistant Specialist, 11/87 - termination

HORAN, Hugh - CRSP Program Coordinator, 9/84 - 11/87

KEFFELEW, Atnaf - Data Management Analyst/Programmer, 11/86 - 11/87

KOYAK, Robert - Statistical Research Assistant, 6/83 - 8/84

KREBS, William - Research Assistant, 6/83 - 12/84

LAST, Ross - Undergraduate Nutritional Sciences Assistant, 1/87 - 5/87; Senior Clerk/Typist, 6/87 - termination

LEE, Lily - Research Assistant, 5/83 - 6/83

LEIN, David - Data Management Analyst/Programmer, 5/85 - termination

LEROY, Susan - Statistical Research Assistant, 7/83 - 8/84

LIEEL, Birgit - Assistant Nutrition Specialist, 3/87 - 11/87

MERRILL, Katherine - Administrative Assistant, 10/87 - termination

MURPHY, Suzanne - ME Data Manager and Assistant Research Nutritionist, 9/1/85 - termination

NELSON, Peggy - Secretary, 10/81 - 11/81, 4/82

PETERSON, Christine - Sr. Typist Clerk, 4/82 - 12/83, 5/86 - 10/86, 1/87

RITTENHOUSE, Duane - Student Assistant, 10/86 - termination

ROSBERG, Carl - Co-Program Administrator, inception (9/81) - termination

ROSE, Don - Research Assistant, 3/87 - 11/87

SAKOVICH, Maria - Secretary, 1/81 - 12/82

SCHLYER, Srinual Lin - Business Manager, 1/82 - 11/82

SELVIN, Steve - Statistical Advisor, 7/85 - 12/87

SHARP, Leslie - Research Assistant, 7/85 - 8/85

SMITH, Joanne - Secretary, 11/82 - 4/83

WATERS, Claudia - Data Management Analyst/Programmer, 7/83 - 3/87

WEAVER, Susan - Research Assistant - 6/83

WILHITE, Judith - Senior Coder, 1/86 - 2/86

WILSON, Alan - Research Educator, 6/83

YOUNG, Patricia - Secretary, 6/82 - 9/82

CONSULTANTS TO THE MANAGEMENT ENTITY

Dr. Edmundo Alvarez, M.D., Mexico Consultant, Oficina Sanitaria Panamericana, Los Cedros 269, San Isidro, Lima, PERU

Professor David E. Barrett, Clemson University, Department of Elementary and Secondary Education, 400 Pillman Hall, Clemson, SC 29634-0709

Dr. George Beaton, Principal Consultant, Department of Nutritional Sciences, Faculty of Medicine, Fitzgerald Building, University of Toronto, Toronto, Ontario, CANADA M5S 1A8

Dr. Leo Breiman, Statistics, Department of Statistics, 409 Evans Hall, University of California, Berkeley, CA 94720

Dr. Tiffany Field, Psychology/Cognition, Milkman Center for Child Development, University of Miami, P.O. Box 016820, Miami, FL 33101

Dr. Stephen Fienberg, Statistics, Department of Statistics, Carnegie-Mellon University, Schenley Park, Pittsburgh, PA 15213

Dr. Raymond Neff, Mexico Consultant, Information Systems and Technology, 209 Evans Hall, University of California, Berkeley, CA 94720

Dr. Steve Selvin, Statistics, School of Public Health, 120 Warren Hall, University of California, Berkeley, CA 94720

Dr. Marian Sigman, Psychology/Cognition, Neuro-Psychiatric Institute, Room 68-237E, University of California, Los Angeles, CA 90024

Dr. Elliot Turiel, Psychology/Cognition, School of Education, 4317 Tolman Hall, University of California, Berkeley, CA 94720

Dr. Theodore Wachs, Psychology/Cognition, Department of Psychological Services, Pierce Hall, Purdue University, West Lafayette, IN 47907

APPENDIX D

CHRONOLOGY OF NUTRITION CRSP
SEPTEMBER 1981 - JANUARY 1988

1981

September
- o Signing of Grant Contract by USAID.
- o D. Cattle opens CRSP Management Entity (ME) office at UC Berkeley.

October
- o SCB pro tem meeting (Los Angeles) to plan for preliminary field studies and to establish mechanisms for exchange of information on methodologies.

November
- o Installation of G. Beaton as Interim Project Director at Berkeley until 4/82.

December
- o UC Berkeley accepts Contract with USAID.

1982

January
- o TAG Meeting on Statistical Design and Data Management
  (J. Ware-Chair, G. Beaton, A. Coulson, R. Hassanein, P. Pelto, W. Rand, W. Reinke)
  Review and discussion of research design, data management system, procedures for refinement of design study, plans for development of recommended data management system, organization matters related to design and data management, and implementation recommendations.

- o Egypt and Mexico project subagreements issued by UC Berkeley.

March
- o TAG Meeting on Cognition
  (T. Wachs-Chair, T. Field, G. Findlay, E. Pollitt, C. Ramey, E. Werner)
  Discussion of problems and procedures, including design and statistics, coding, and interproject standardization.

- o Signing of subcontract between Thomas Cook Travel and UC Berkeley.

April
- o SCB Meeting, New Orleans, LA
  To confer on status of preliminary Phase I activities, which included: selection of communities for the study according to specific criteria; formulation of technical plans; and formulation of plans for the Phase II research design and data collection based on the field experience of the three projects.
  Review of TAG reports.

o TAG Meeting on Food Intake
(D. Calloway-Chair, H. Aly, G. Beaton, A. Chavez, N. Jerome, A. Mata, G. Pelto)
Recommendations made on methodology for collecting dietary data and for management and analysis of dietary data across projects.

o TAG Meeting on Nutritional Status
(A. Chavez-Chair, G. Harrison, C. Neumann)
Determination of nutritional status measures to be made by all projects, based on discussion of various indicators - anthropometric,clinical, and biomedical.

o TAG Meeting on Disease
(N. Bwibo-Chair, R. Chandra, A. Mata, C. Neumann, R. Watson)
Establishment of a core set of measurements and methodologies for evaluating effects of marginal malnutrition on immunologic function and of a standard for documenting clinical morbidity.

o Signing of contractual agreement between ME at UC Berkeley and the University of Connecticut for the Mexico Project.

o Signing of contractual agreement between ME at UC Berkeley and Thomas Cook Travel, Inc.

May

o TAG Meeting on Physical Activity
(L. Allen-Chair, A. Ferro-Luzzi, A. Ismail, E. Messer, G. Spurr)
Discussion of research approach to measuring work performance and activity patterns of daily living, covering areas of physical activity and physical fitness.

o TAG Meeting on Social Competence
(N. Jerome-Chair, G. Beaton, C. Carr, P. Pelto)
Discussion of the concept of social competence as it relates to the CRSP, of relevant research questions for Phase I preliminary survey, and effective mechanisms for obtaining accurate, reliable, quantitative data.

o Acceptance of UC Berkeley subcontract by University of Connecticut for the Mexico project.

June

o Installation at Berkeley of J. Balderston as CRSP Deputy Program Administrator.

o Meeting of US PIs and Statistical Design and Data Management Committee.
To discuss construction of integrated Interim Design from the separate TAG proposals. Decisions made on Phase I activities, instruments tested in Phase I for use in Phase II. Discussion of design resulted in draft proposal which included schedules, protocols, and observational techniques.

o Signing of contractual agreement between ME and Purdue University for the Egypt Project.

o Issuing of subagreement with National Nutrition Institute of Mexico (INNSZ) by University of Connecticut for the Mexico project.

July

o TAG Meeting on Reproduction
(L. Allen, N. Butte, A. Kirksey, J. King, C. Ritenbaugh, M.S. Read)
Discussion of methodology for assessing reproductive function as measured by maternal and infant parameters.

o Signing of contractual agreement between ME and UC Los Angeles for the Kenya Project.

August

o Subagreement with the Egyptian Ministry issued by Purdue University for the Egypt Project.

September

o D.H. Calloway and G. Beaton site visit to Mexico City and Solis (9/1-9/30).

November

o SCB Meeting (Airlie House, Virginia): Formulation of Phase II research plan, discussion of appropriate unit of study, choice of target subjects, sampling strategy, and specific functions to be included in the design.

o Institutional Council (IC) and Finance Committee (FC) Meeting (Airlie House, Virginia).

o External Advisory Panel (EAP) renamed to External Evaluation Panel (EEP) by revision of the CRSP Guidelines.

December

o First design report completed by SCB and ME, forwarded to EEP.

o Data Management Group meeting, Berkeley, with representatives of ME and all projects.
Discussion of data management process, methods of quality control and data transfer.

1983

January

o EEP Meeting with project representatives and ME, Berkeley, CA. EEP recommended simplification of design to concentrate on the functions of cognition and morbidity, and non-core studies to be limited to 15% of each project's budget.

	o Deloitte, Haskins & Sells (D,H&S) subcontract signed with ME for audits of all overseas projects.
March	o D,H&S First Audit of CRSP non-U.S. collaborating institutions covering period 9/1/81 - 2/28/83.
April	o EEP Report of January meeting submitted to USAID and ME.
May	o US PIs meet in Berkeley to reformulate design according to EEP recommendations and to make basic sampling decisions.
June	o Revised Research Design prepared. Refocussed around five main hypotheses reflecting EEP-recommended priorities. Sample size established.
	o Meeting of the Berkeley Committee for the Protection of Human Subjects - subcommittee formed to resolve questions relative to the CRSP.
July	o Submission of Deloitte, Haskins & Sells audit report on Egypt and Mexico host country institutions, for the period 9/1/81-3/31/83.
August	o Meeting of EEP chairperson Sol Chafkin, USAID program officer, and ME, in Berkeley; agreement that the refocused research design as suggested by the EEP was acceptable and that the CRSP should proceed to Phase II.
	o EEP approval of Revised Research Design.
	o M. Forman site visit to Embu and Nairobi.
September	o Completion of "Manual on Data Collection & Data Flow" by Berkeley Data Management team (First version of "Red Book").
	o USAID approval of Revised Research Design.
October	o Seminar presentation by D.H. Calloway on CRSP to USAID staff, Washington, D.C.
	o CRSP Directors' Meeting at USAID, Washington, D.C.
November	o Umbrella Human Subjects Protocol developed by the Berkeley Committee for the Protection of Human Subjects.
December	o D.H. Calloway and J. Balderston site visit to Mexico to evaluate operations: meet with field staff and data management personnel, observe interview techniques, and visit the Institute of Nutrition (12/5 - 12/9).

1984

January	o J. Balderston site visits to Kenya and Egypt to assess Phase II operations (1/10-2/2).
February	o D.H. Calloway site visits to Egypt and Kenya to evaluate progress on implementation of Phase II research (1/31-2/12).
March	o D.H. Calloway site visit to Mexico to evaluate progress on implementation of Phase II research (3/20-22).
	o D.H. Calloway trip to Rome, Italy to consult with FAO staff (3/3-7) on CRSP-related matters.
May	o SCB, FC, IC meet in Berkeley. Discussion of design changes; plans initiated for reorganization of responsibilities and procedures of decision-making for the remainder of the CRSP; Professor Bwibo designated as Kenya PI, replacing Dr. Kagia.
	o ME meeting in Berkeley with M. Forman to discuss reorganization.
June	o Submission to AID of Proposal for the Administrative Reorganization of the CRSP.
	o Definition of Basic Data Set proposed to simplify research design.
July	o EEP members P. Payne and G. Keusch site visit and evaluation of Egypt Project (7/15 -18).
August	o EEP members M. Yarrow and D. Thompson site visit and evaluation of Mexico Project (8/18-21).
September	o EEP members P. Payne and G. Keusch site visit and evaluation of Kenya Project (7/15-18).
	o SCB meeting (Berkeley, CA) Review variables lists, data management and data flow, timetables for second year of Phase II, publications agreement; consider time allocation and activity protocols; share EEP site visit experiences.
	o EEP meeting (Berkeley, CA) Review and evaluate projects' verbal reports of progress, discuss analytical plans: consider reorganization of ME and terms of reference of Program Coordinator.
	o Installation of CRSP Program Coordinator H. Horan.
	o Submission of Deloitte, Haskins & Sells audit report on Mexico and Egypt project host country institutions, covering the period 4/1/83-3/31/84.

October — o Revision of Basic Variables List completed.

November — o H. Horan site visit to Mexico City, Solis (11/3-13).

o Submission of Deloitte, Haskins & Sells audit report on Kenya project host country institution, covering period 7/1/83-3/31/84.

December — o ME Consultant G. Beaton and EEP member S. Chafkin appear before the Joint Committee on Agricultural Research (JCARD), Washington, D.C.

o D.H. Calloway meets with AID representatives M. Forman and S. Kahn in Washington, D.C. to report on CRSP activities and progress.

1985

January — o Meeting of PIs, H. Horan, and G. Beaton in Cairo, Egypt to discuss data flow, management, and analysis of food intake data.

o H. Horan site visit to Embu and Nairobi, Kenya (1/14-30).

February — o SCB, IC/FG meetings in Berkeley; formative meetings of Data Analysis/Statistics group. Discussion of analytical strategies and organization of activities by projects and monitoring by ME; rebudgeting among projects and ME agreed.

o G. Beaton site visit to Embu, Kenya (2/7-8).

April — o CRSP Managers meeting at USAID, Roslyn, VA.

o EEP member M. Yarrow meeting with project psychologists M. Sigman and T. Wachs, and G. Beaton, in Toronto to review and evaluate projects' progress and analytical plans relating to cognitive and behavioral variables and report to EEP.

o SCB and EEP meeting with ME and consultants G. Beaton and S. Fienberg in Querretero, Mexico. Review and approval of project data analysis plans; presentation of preliminary analysis and progress report.

May — o M. Forman, S. Kahn, and D. Calloway meet at AID offices in Washington, D.C.

o F. Peters site visit to Kenya (5/1-31).

June — o Inter-CRSP meeting, sponsored by the Peanut CRSP, in Georgia, attended by H. Horan.

July	o ME data management and analysis planning meeting, UC Berkeley.
August	o Informal meeting of CRSP PIs attending IUNS Congress, Brighton, England, together with Professor Mamadou Gabr and M. Forman.
September	o Site visits to Storrs, CT and Mexico City/Solis site by the special Mexico consultancy team to advise on data flow and management activities and field science question.
October	o Data Analysis Group meeting at University of Connecticut, Storrs, including project statisticians, ME consultants S. Selvin and G. Beaton, S. Murphy, data managers, and some PIs.
November	o Visit of ME Data Manager S. Murphy and programmer C. Waters to UCLA to confer with Kenya project on data management.
	o Visit of S. Murphy to University of Kansas to attend Egypt project data management meetings.
December	o Visit of G. Beaton and G. McCabe to Berkeley to consult with ME on data management procedures.

1986

February	o Meeting of the Scientific Coordination Board (SCB) at UCLA, concurrent with meetings of the statisticians, data managers, ME (D. Calloway, H. Horan, G. Beaton, J. Balderston, S. Murphy, and S. Selvin), and the Institutional Council.
March	o Program effectiveness audit at ME in Berkeley by USAID.
	o Visit of S. Murphy to UCLA to consult with Kenya project data management team.
April	o Meeting of Egypt project personnel in St. Louis, MO, attended by G. Beaton and S. Murphy.
May	o Visit by G. Beaton to UCLA to confer with data management and C. Neumann on Kenya project data, and to Berkeley to confer with ME (D. Calloway, H. Horan, S. Murphy, J. Balderston, S. Selvin, and programmers C. Waters and D. Lein) on data management and analysis.
June	o Visit to Berkeley by Kenya project PI C. Neumann and researcher S. Weinberg to confer with D. Calloway, S. Murphy, and research assistant K. Mulligan.
	o Statisticians' meeting at Purdue University, Lafayette, Indiana, including G. Beaton and S. Selvin.

o Submission of Deloitte, Haskins & Sells audit report on Mexico and Egypt project host country institutions, covering the period 4/1/84-9/30/85 and on Kenya project host institution covering the period 4/1/84-12/31/85.

July

o Visit of G. Beaton to Berkeley to consult with Data Management team (S. Murphy, S. Selvin, and programmers C. Waters and D. Lein).

o Congressional presentation on all CRSPs sponsored by the National Association of State Universities and Land Grant Colleges (NASULGC), attended by Mexico PI G. Pelto.

August

o Visit of G. Beaton to the University of Connecticut to confer with PIs, data managers, and statistician on the Mexico project.

o Visit of S. Murphy and ME Data Management Research Assistant K. Mulligan to UCLA to consult with Kenya project staff.

September

o Chicago meeting of statisticians and data managers, Egypt project PI G. Harrison, and ME consultants S. Fienberg and G. Beaton.

November

o Scientific Coordination Board meeting at Berkeley, including EEP representatives G. Keusch and D. Thompson, AID, PIs, psychologists, statisticians, ME, and data managers to review project analytical plans and progress.

o Formal request submitted to AID for no-cost extension of CRSP to 1/30/88.

December

o Informal request submitted to AID for further funds for the country projects.

1987

March

o AID Grant Modification #12 received at UC Berkeley, authorizing continuance of CRSP beyond January 31, 1987, and authorizing the last tranche of the $11.8 million award.

o Meeting of D. Calloway and S. Murphy with those US PIs attending the FASEB meeting in Washington, DC.

April

o Berkeley analytical meetings between Drs. McCabe, Beaton, Calloway, Murphy, and Horan. Preliminary discussions of final format of archive data sets and documentation, RMR issues, Final Report narrative, and analytical agenda through the end of the CRSP.

	o EEP member M. Yarrow meeting with project psychologists M. Sigman and T. Wachs to review and evaluate progress.
June	o Review and discussion of CRSP psychology analyses, Chicago, including G. Beaton, T. Field, T. Wachs, G. Pelto, G. McCabe, and M. Sigman.
July	o CRSP Directors and Program Managers Workshop in Washington, D.C., with Hugh Horan representing NCRSP.
	o PCG meeting in Berkeley, chaired by EEP-member S. Chafkin, including PIs G. Harrison, L. Allen, C. Neumann, O. Galal, A. Chavez, and N. Bwibo; overseas officials M. Gabr and M. Okoth and program Co-Administrator C. Rosberg.
	o Berkeley consultation with Dr. George Beaton; review of the first draft of project reports.
August	o Continued Berkeley consultation with by Dr. George Beaton, further review of final reports and start of preparation of ME Final Report.
	o IDECG meeting, Guatamala. Attended by Dr. George Beaton.
October	o Submission of the final draft of the Kenya Final Report to ME. Reviewed and edited by ME staff.
	o Submission of the final draft of the Mexico Final Report to ME. Reviewed and edited by ME staff.
	o Submission of the final draft of the Egypt Final Report to ME. Reviewed and edited by ME staff.
	o Consultation with Dr. George Beaton in Berkeley, for preparation of the ME Final Report.
November	o Review of printer's proof of the Kenya Final Report by ME and Dr. Charlotte Neumann.
	o First sections of ME Final Report sent to PI's for review.
December	o Review of printer's proof of the Mexico Final report by ME and Drs. Lindsey Allen and Gretel Pelto.
	o All sections of ME Final Report sent to PIs for review.
January	o Final copy of the Kenya Final report produced. Copies sent to projects and USAID.
	o Final copy of the Mexico Final Report produced. Copies sent to projects and USAID.

- o Final copy of the Egypt Final report produced. Copies sent to projects and USAID.
- o Final copy of ME Final Report produced. Copies sent to projects and USAID.
- o Final archive of all data sets and documentation. Copies sent to projects and USAID.
- o CRSP officially closes.

APPENDIX E

CRSP BASIC FILES AND MASTER INDEX OF BASIC VARIABLES

Following is a list of the Basic Files that were defined for the revised Basic Data Set, September, 1984. The page numbers refer to the section of this appendix that details the types of variables within each Basic File. Also included in each section is an indication of which projects collected a variable type, and the number of variables collected. Full details on the Basic Variables and Files are given in the document, "Nutrition CRSP, Basic Data Set", September, 1984. Project variable information is based on the final data archive in Berkeley, 11/30/87.

BASIC FILE	PAGE
Household entry/exit/change	269
Mortality notification and diagnosis	272
Household food intake summary	273
Individual food intake summary	274
Supplemental feeding - infants	276
Anthropometry - target subjects	277
Anthropometry - non-target subjects	278
Metabolic adaptation (RMR)	279
Historical assessment: Disabilities and chronic illness	282
Reproductive history	285
Physical assessment: nutrition related	287
Physical assessment: during pregnancy	289
Pregnancy outcome	290
Monthly conception survey	292
Monthly pregnancy survey	293
Lactation and infant feeding	295
Laboratory assessment, including immunology	297

BASIC FILE	PAGE
Morbidity; illness episodes	300
Morbidity summary	303
Psychological function:	
Infant	304
Toddler	306
Schooler	308
Adult	310
Household sanitation	311
Individual demographic data	313
Household socioeconomic data	315
Community climatic data	317
Childcare and sanitation activities of lead female	318

HOUSEHOLD ENTRY, CHANGE, OR EXIT MASTER INDEX

PROJECT	FILES	NUMBER OF RECORDS	NUMBER OF HOUSEHOLDS	ESTIMATED COLLECTION FREQ
Kenya	Entry/Change/Exit	1062	292	as needed
Egypt	Entry/Change/Exit	230	230	once
Mexico	Target Individuals(1)	1068	1068	once

BASIC VARIABLES	KENYA	EGYPT	MEXICO
o Code # of HH for which entry, change, or exit is recorded	1 var	1 var	
o Type # of HH before this entry, change, or exit	1 var	1 var	
o Type # of HH after this entry, change, or exit	1 var	1 var	
o Section code: 01 HH entry 02 HH change/exit	1 var	1 var	
o ID of lead female		1 var	
o ID of lead male		1 var	
o Is this a record of HH entry to study?	1 var	1 var	
o Has a target infant been born or left target status?	1 var	1 var	
o Has a toddler entered or left target status?	1 var	1 var	
o Has a school-aged child entered or left target status?	1 var	1 var	
o Has the lead female entered or left the study or changed reproductive status?	1 var	1 var	
o Has the lead male entered or left the study, or returned to or left the HH?	1 var	1 var	
o Have non-target persons moved into or out of the HH?	1 var	1 var	
o Is the HH leaving the study?	1 var	1 var	
Household Entry:			
o Date of HH entry	1 var	1 var	
o # of target hh members at entry	1 var	1 var	
o # of non-target HH members at entry	1 var	1 var	
o Total # of HH members at entry	1 var	1 var	
Infant:			
o ID of target infant	1 var	1 var	
o Date change in status of target infant occurred	1 var	1 var	
o Does current change include birth of			

target infant?	1 var	1 var
o Has target infant completed the study?	1 var	1 var
o Has target infant left the study before completion?	1 var	1 var
Toddler:		
o ID of target toddler	1 var	1 var
o Date change in status of target toddler occurred	1 var	1 var
o Does current change include entry of toddler as a target subject?	1 var	1 var
o Has target toddler completed the study?	1 var	1 var
o Has target toddler left the study before completion?	1 var	1 var
School-Age Child:		
o ID of target school-age child	1 var	1 var
o Date the change in status of target school-age child occurred	1 var	1 var
o Does current change include entry of school-age child as a target subject?	1 var	1 var
o Has target school-age child completed the study?	1 var	1 var
o Has the target school-age child left the study before completion?	1 var	1 var
Lead Female:		
o Date change in status of lead female occurred	1 var	1 var
o Does this change include entry of lead female in the study?	1 var	1 var
o Does current change include anticipation of lead female's target pregnancy?	1 var	1 var
o Does current change include confirmation of lead female's target or non-target pregnancy?	1 var	1 var
o Does current change include termination of lead female's target or non-target pregnancy?	1 var	1 var
o Does current change include failure of lead female to conceive a target infant?	1 var	1 var
o Has lead female completed the study?	1 var	1 var
o Has the lead female left the study before completion?	1 var	1 var
Lead Male:		
o Date the change in status of lead male occurred	1 var	1 var
o Does this change include entry of the lead male in the study?	1 var	1 var
o Does current change include lead male leaving the HH temporarily?	1 var	1 var
o Does current change include return of lead male to the HH?	1 var	1 var
o Has the lead male completed the study?	1 var	1 var
o Has the lead male left the study		

before completion?	1 var	1 var
Non-Target:		
o Date non-target subject(s) moved into or out of the HH	1 var	1 var
o # of non-target subjects who moved into the HH with this change	1 var	1 var
o # of non-target subjects who moved out of the HH with this change	1 var	1 var
Household Exit:		
o Date of HH exit	1 var	1 var
o Reason HH is leaving the study	1 var	1 var

(1) Mexico does not have a file in the format of the file "Household Entry, Change, or Exit". The file Target Individuals identifies the Mexico Project target individuals. The Mexico Project has defined two types of target individuals: A. those in the cross-sectional data set, and B. those in the longitudinal data set.

NON-BASIC VARIABLES	KENYA	EGYPT	MEXICO
o # of non-targets in HH that are potential targets	1 var		
o Individual's ID and sex			4 var
o Cross-section: Begin of window(SAS date)			1 var
o Cross-section: End of window (SAS date)			1 var
o Longitudinal: Begin of window (SAS date)			1 var
o Longitudinal: End of window (SAS date)			1 var
o Birth date of subject			1 var
o Individual type: cross-section			1 var
o Individual type: longitudinal			1 var
o Position in Family			1 var(a)
o Father's ID, family position, birth date			3 var(b)
o Mother's ID, family position, birth date			3 var(b)

(a) for lead males and lead females only
(b) for infants, toddlers, schoolers only

MORTALITY NOTIFICATION AND DIAGNOSIS MASTER INDEX

PROJECT	FILES	NUMBER OF RECORDS	NUMBER OF INDIVIDUALS	ESTIMATED COLLECTION FREQ
Kenya	Mortality Notification(1)	4	4	once
Egypt	(2)			
Mexico	Mortality(3)	603	??	once

BASIC VARIABLES	KENYA	EGYPT	MEXICO
o Date of death	1 var		1 var
o What was cause of death? (accident, illness, other)			14 var(a)
o Was a morbidity questionnaire filled out?			
o Was medical care sought?			

(1) A Kenya Mortality file with 4 records exists. A codebook for that file is not available in Berkeley, though complete range checks have been archived. Deaths were also recorded as a morbidity outcome on the individual morbidity file (forms 313 and 314) (see Kenya Final Report Chapter 8 and Kenya Morbidity Manual)
(2) Information on mortality for the Egypt Project is in the Data Documentation Archive under Mortality.
(3) The Mexico Mortality file contains records for 201 households (3 records/HH). 24 households report that at least one death had occurred in the past 5 years. Individual ID's are given for 27 subjects, including 13 with ID=9999 (non-targets). The date of death is reported for 25 subjects. According to the dates 16 individuals died within the study period (01/84-03/86).

(a) In the Mexico Mortality file the cause of death is written in characters (in Spanish) and a diagnosis code is reported. Recordings for two deaths per household are possible.

NON-BASIC VARIABLES	KENYA	EGYPT	MEXICO
o Death in family in past 5 years (y,n)			1 var
o Individual ID of person #1,#2			2 var
o Sex of person #1,#2			2 var
o Age of person #1,#2			2 var
o Date of death of person #1,#2			2 var

HOUSEHOLD FOOD INTAKE SUMMARY MASTER INDEX

PROJECT	FILES	NUMBER OF RECORDS	NUMBER OF HOUSEHOLDS	ESTIMATED COLLECTION FREQ
Kenya	Food intake: HH Energy Ratio	8421	247	2 days/mo
Egypt	(1)			2 days/mo
Mexico	(2)			2 days/mo

BASIC VARIABLES	KENYA	EGYPT	MEXICO
o Date intake data collected	1 var		
o Day of week			
o Code for method of measurement			
o HH adjusted intake: g/24 hrs for all food consumed by HH members from HH supply			
o Total kcals consumed by HH	3 var(a)	1 var(b)	
o Total g protein consumed by HH			
o Total g fat consumed by HH			
o Total g carbohydrates consumed by HH			
o Total kcal consumed by lead male from HH supply			
o Total g protein consumed by lead male from HH supply			
o Total g fat consumed by lead male from HH supply			
o Total g carbohydrates consumed by lead male from HH supply			
o Household energy intake adjusted by household BMR	3 var(a)		

(1) Food items archived; energy values were sent to Berkeley but later determined to be based on incorrect food composition data and therefore withdrawn from the archive.
(2) Food items archived but not nutrients.

(a) Kenya nutrients are divided into dish nutrients, snack nutrients, and non-household nutrients.
(b) Not revised to reflect corrected food composition table values.

INDIVIDUAL FOOD INTAKE SUMMARY MASTER INDEX

PROJECT	FILES	NUMBER OF RECORDS	NUMBER OF INDIVIDUALS	ESTIMATED COLLECTION FREQ
Kenya	Food Intake: Indiv. Nutr.	39192	1358	2 days/mo
Egypt	Food Intake: Indiv. Nutr.	15961	995	2 days/mo
Mexico	Food Intake: Indiv. Nutr.	20736	1116	2 days/mo

BASIC VARIABLES	KENYA	EGYPT	MEXICO
o Date intake data collected	1 var	1 var	1 var
o Day of week	1 var	1 var	1 var
o Collection method (regular, special)		1 var	(c)
o Food composition code (to indicate nutrient conversion method used)			1 var
Intake:			
o Kcals per day	3 var(d)	1 var(a)	1 var
o Proteins in grams per day	3 var	2 var(a,b)	1 var
o Fats in grams per day	3 var	2 var(a,b)	1 var
o Carbohydrates in grams per day	3 var	1 var(a)	1 var

(a) Not revised to reflect corrected food composition table; both the Egypt Project final report and the ME final report use energy values calculated from protein, fat, and carbohydrate, using the 4-9-4 factors.
(b) Specified: from animal source, from plant source
(c) The variable is supposed to be in the file "Mexico Kcals". No value has been recorded here.
(d) Kenya nutrients are divided into dish nutrients, snack nutrients, and non-household nutrients.

NON-BASIC VARIABLES	KENYA	EGYPT	MEXICO
o Daily kcals from animal source			1 var
o Daily kcals from plant source			1 var
o Protein (from animal source): g/day	1 var	1 var	1 var
o Protein (from plant source): g/day		1 var	1 var
o Fat (from animal source): g/day		1 var	
o Fat (from plant source): g/day		1 var	
o Total grams of water per day		1 var	
o Calcium: mg/day		1 var	1 var
o Iron (from animal source): mg/day		1 var	
o Iron (from plant source): mg/day		1 var	
o Total daily iron: mg/day	1 var		1 var
o Sodium: mg/day		1 var	
o Thiamin: mg/day		1 var	1 var

o Riboflavin: mg/day	1 var	1 var
o Niacin: mg/day	1 var	1 var
o Retinol equivalents: micrograms/day		1 var
o Retinol: micrograms/day	1 var	
o Retinol equivalents of retinol/day	1 var	
o Beta-Carotene: micrograms/day	1 var	
o Retinol equiv. of beta-carotene/day	1 var	
o Ascorbic acid: mg/day	1 var	1 var
o Crude fiber: g/day	1 var	
o Ash: g/day	1 var	

SUPPLEMENTARY FEEDING: INFANT MASTER INDEX

PROJECT	FILES	NUMBER OF RECORDS	NUMBER OF INFANTS	ESTIMATED COLLECTION FREQ
Kenya	Food Intake: Indiv.(1)	13613	140	2 days/mo
Egypt	Food Intake: Indiv.(1)	15901	127	2 days/mo
Mexico	Food Intake: Indiv.(1)	20809	97	2 days/mo

BASIC VARIABLES	KENYA	EGYPT	MEXICO
o ID of infant's mother			
o Date of observation	1 var	1 var	1 var
o Day of week	1 var	1 var	1 var
o Did infant consume anything other than breast milk?	Lact.*	Lact.**	
o Did infant consume formula or milk from bottle, cup or spoon?	Lact.*		
o Did infant consume any water?	Lact.*		
o Was sugar added to the water?	Lact.*		
Other nutrients:			
o Kcal intake exclusive of breast milk per 24 hour period	3 var	1 var	1 var
o Grams of protein exclusive of breast milk per 24 hour period	3 var	2 var(a)	1 var
o Grams of fat exclusive of breast milk per 24 hour period	3 var	2 var(a)	1 var
o Grams of carbohydrates exclusive of breast milk per 24 hour period	3 var	1 var	1 var

(1) No project archived a separate file for supplementary feeding of infants. However, all collected information on infant nutrient intakes, exclusive of breast feeding, as part of the regular food intake files.

* Information available in file: Lactation Questionnaire
** Information available in file: Lactation/Infant Feeding

(a) Specified: from animal source, from plant source

ANTHROPOMETRY - TARGET SUBJECTS MASTER INDEX

PROJECT	FILES	NUMBER OF RECORDS	NUMBER OF INDIVIDUALS	ESTIMATED COLLECTION FREQ
Kenya	Anthropometry (1)	13613(2)	1022	monthly
Egypt	Anthropometry	9320	1252	monthly
Mexico	Individual Anthropometry	6751	987	every 3 mo(3)

BASIC VARIABLES	KENYA	EGYPT	MEXICO
o Height in centimeters	1 var	1 var	1 var
o Weight in kilograms	1 var	1 var	1 var
o Supine length for infants and toddlers in centimeter	1 var	1 var	1 var
o Arm circumference in centimeters	1 var	1 var	1 var
o Head circumference in centimeters	1 var	1 var	1 var

(1) The Kenya Anthropometry file contains the data for target and non-target individuals. Above counts are for targets only.
(2) Records for target individuals only.
(3) For infants, toddlers, pregnant and lactating women: monthly

NON-BASIC VARIABLES	KENYA	EGYPT	MEXICO
o Skinfold: triceps	1 var	1 var	2 var
o Skinfold: biceps	1 var	1 var	2 var
o Skinfold: subscapular	1 var	1 var	2 var
o Skinfold: suprailiac	1 var(a)		2 var
o Skinfold: abdomen	1 var(a)		
o Skinfold: thigh	1 var(a)		
o Age in years		1 var	
o Age in months		1 var	
o Age of infant in days		1 var	
o Age of infant in months		1 var	
o Length of lactation in days		1 var	
o Length of pregnancy or lactation in mo		1 var	
o Circumference of lower arm in cm			1 var
o Circumference of abdomen in cm			1 var
o State of subject's health (well, ill)			1 var

(a) Females only

ANTHROPOMETRY - NON-TARGET SUBJECTS MASTER INDEX

PROJECT	FILES	NUMBER OF RECORDS	NUMBER OF INDIVIDUALS	ESTIMATED COLLECTION FREQ
Kenya	Anthropometry(1)	943	915	once
Egypt	Anthropometry:Non-targ.	1112	431	twice(?)
Mexico	Anthropometry:non-targ.	699	698	once

BASIC VARIABLES	KENYA	EGYPT	MEXICO
o Height in centimeters	1 var	1 var	1 var
o Weight in kilograms	1 var	1 var	1 var
o Supine length for infants and toddlers in centimeter	1 var	1 var	1 var
o Arm circumference in centimeters	1 var	1 var	1 var
o Head circumference in centimeters	1 var	1 var	1 var

(1) The Kenya Anthropometry file contains records for target and non-target individuals.

NON-BASIC VARIABLES	KENYA	EGYPT	MEXICO
o Skinfold: triceps	1 var	1 var	2 var
o Skinfold: biceps	1 var	1 var	2 var
o Skinfold: subscapular	1 var	1 var	2 var
o Skinfold: suprailiac	1 var(a)		2 var
o Skinfold: abdomen	1 var(a)		
o Skinfold: thigh	1 var(a)		
o Age in years		1 var	
o Age in months		1 var	
o Age of infant in days		1 var	
o Age of infant in months		1 var	
o Length of lactation in days		1 var	
o Length of pregnancy or lactation in mo		1 var	
o Circumference of lower arm in cm			1 var
o Circumference of abdomen in cm			1 var
o State of subject's health (well, ill)			1 var

(a) Females only

METABOLIC ADAPTATION MASTER INDEX

PROJECT	FILES	NUMBER OF RECORDS	NUMBER OF INDIVIDUALS	ESTIMATED COLLECTION FREQ
Kenya	Resting Metabolic Rate	4288	633	every 3 mo
Egypt	Resting Metabolic Rate	798	482	every 3 mo (?)
Mexico	Resting Metabolic Rate	1930	510	every 3 mo

BASIC VARIABLES	KENYA(a)	EGYPT(b)	MEXICO(c)
o Date measurements taken	1 var	1 var	1 var
o Weight of subject in kg	1 var*	1 var	1 var
o Height of subject in cm		1 var	1 var
o Body temperature of subject	1 var	1 var	
o Kcal/kg of body weight: minute 18	1 var**	1 var(d)	
o Kcal/day: minute 18	1 var	1 var(e)	
o Oxygen consumption/minute: minute 18	1 var***	1 var	
o Respiratory quotient: minute 18	1 var	1 var	
o Heart rate: minute 18	2 var****	1 var(f)	
o Kcal/kg of body weight: minute 19		1 var(d)	
o Kcal/day: minute 19		1 var(e)	
o Oxygen consumption/minute: minute 19	1 var	1 var	
o Respiratory quotient: minute 19	1 var	1 var	
o Heart rate: minute 19		1 var(f)	
o Kcal/kg of body weight: minute 20		1 var(d)	
o Kcal/day: minute 20		1 var(e)	
o Oxygen consumption/minute: minute 20	1 var	1 var	
o Respiratory quotient: minute 20	1 var	1 var	
o Heart rate: minute 20		1 var(f)	
o Kcal/kg of body weight: minute 28	1 var	1 var(g)	
o Kcal/day: minute 28	1 var	1 var(h)	
o Oxygen consumption/minute: minute 28	1 var	1 var	
o Respiratory quotient: minute 28	1 var	1 var	
o Heart rate: minute 28		1 var(i)	
o Kcal/kg of body weight: minute 29		1 var(g)	
o Kcal/day: minute 29		1 var(h)	
o Oxygen consumption/minute: minute 29	1 var	1 var	
o Respiratory quotient: minute 29	1 var	1 var	
o Heart rate: minute 29		1 var(i)	
o Kcal/kg of body weight: minute 30		1 var(g)	
o Kcal/day: minute 30		1 var(h)	
o Oxygen consumption/minute: minute 30	1 var	1 var	
o Respiratory quotient: minute 30	1 var	1 var	
o Heart rate: minute 30		1 var(i)	

(a) The Beckman Metabolic Measurement Cart was used to measure resting energy expenditure in Kenya. Analysis of Kenya RMR for both the Kenya Final Report and the ME Final Report was produced with a series of corrections to the data. See these reports for specific documentation. A copy of the computer program to apply corrections has been provided on tape (see Kenya tape documentation).

(b) The Beckman Metabolic Measurement Cart was used to measure resting energy expenditure. The protocol followed in Egypt deviates from the original protocol for RMR measurements (see documentation).

(c) In Mexico the Beckman Metabolic Cart was not used and therefore data from minutes 18-20 and 28-30 of respiration test are not available. See documentation for the protocol followed.

(d) Kcal/kg body weight (second 10 minutes) -- one variable for the second 10 minute period

(e) Kcal/day (second 10 minutes) -- one variable for the second 10 minute period

(f) Heart rate for second 10 min. period -- measured at the beginning of the period

(g) Kcal/kg body weight (third 10 minutes) -- one variable for the third 10 minute period

(h) Kcal/day (third 10 minutes) -- one variable for the third 10 minute period

(i) Heart rate for third 10 minutes -- measured at the beginning of the period

* May be derived from kcal/kg body weight variable

** Kcal/kg body weight and kcal/day are each provided once for each of the second and third 10 minute periods.

*** Tidal volume, oxygen uptake, carbon dioxide output, frequency of respiration, minute ventilation, oxygen consumption/minute and respiratory quotient are each provided three times for each of the second and third 10 minute periods. Note that these measurements may not be for the 18,19,20 or 28,29,30 minute intervals. See Kenya Final Report, Chapter 12 for specific documentation.

**** Heart rate is provided twice: once as a pre-test measurement and once as a post-test measurement.

NON-BASIC VARIABLES	KENYA	EGYPT	MEXICO
o Age of subject in years		1 var	
o # of months pregnant		1 var	
o # of months lactating		1 var	
o Body surface area	1 var	1 var	
o Kcal/day predicted	2 var	1 var	
o Minute ventilation (liters/min) for minutes 18-20	3 var	3 var	
o Frequency of respiration (breaths/min) for minutes 18-20	3 var	3 var	
o Tidal volume (ml/breath)			

for minutes 18-20	3 var	3 var	
o Oxygen uptake (ml/kg/min) for minutes 18-20	3 var	3 var	
o Carbon dioxide output (ml/min) for minutes 18-20	3 var	3 var	
o Basis of calculation for variables of the second 10 min. period (steady state, final min. of 2nd period)	1 var	1 var	
o % kcal/day predicted (second 10 min.)	1 var	1 var	
o Kcal/m² body surface area (2nd 10 min.)	1 var	1 var	
o Minute ventilation (liters/min) for minutes 28-30	3 var	3 var	
o Frequency of respiration (breaths/min) for minutes 28-30	3 var	3 var	
o Tidal volume (ml/breath) for minutes 28-30	3 var	3 var	
o Oxygen uptake (ml/kg/min) for minutes 28-30	3 var	3 var	
o Carbon dioxide output (ml/min) for minutes 28-30	3 var	3 var	
o Basis of calculation for variables of the third 10 min. period (steady state, final min. of 3rd period)	1 var	1 var	
o % kcal/day predicted (third 10 min.)	1 var	1 var	
o Subject's health (well,ill)			1 var
o Subject's physiological state			1 var
o Cardiopulmonary health			1 var
o Room temperature (°C)	1 var		1 var
o Barometric pressure (mm Hg)	1 var		1 var(j)
o Water vapor pressure	1 var		1 var
o Initial volume - first sample			1 var
o Final volume - first sample			1 var
o Volume expired air, first 3 minutes			1 var
o Initial volume - second sample			1 var
o Final volume - second sample			1 var
o Volume expired air, second 3 minutes			1 var
o Percent oxygen - first sample			1 var
o Percent oxygen - second sample			1 var

(j) Barometric pressure not measured after May 1985 as it never changed from 586-590 mm Hg in a year of measurements.

HISTORICAL ASSESSMENT: DISABILITIES AND CHRONIC ILLNESS
MASTER INDEX

PROJECT	FILES	NUMBER OF RECORDS	NUMBER OF INDIVIDUALS	ESTIMATED COLLECTION FREQ
Kenya	Disabilit./Chronic Disease Quarterly Update	727	265 HH	every 3 mo
Egypt	Medical hist./Phys.Exam:			
	Medical History Data	1327	749	targ.:entry,
	Physical Exam	49		6 mo,exit
	Med.His./Phys.Exam	5902		non-t.:entry
Mexico	Clinical Histories(1)	4293(2)	958(2)	every 3 mo (pg,inf monthly)

BASIC VARIABLES	KENYA(a)	EGYPT(c)	MEXICO(d)
o Does the subject have a chronic organic disease such as diabetes, cardiovascular disease, epilepsy, cancer or any other chronic organic disease?			(e)
o ICD code of 1st chronic disease	1 var	(e)	
o ICD code of 2nd chronic disease	1 var		
o ICD code of 3rd chronic disease	1 var		
o Does the subject have an acute illness such as tuberculosis, respiratory disease, sexually transmitted disease, gastrointestinal disease or genito-urinary disease other than STD?		(e)	
o ICD code of 1st acute disease	1 var(b)	(e)	
o ICD code of 2nd acute disease	1 var(b)		
o ICD code of 3rd acute disease	1 var(b)		
o Does the subject have a permanent physical handicap?			
o Is the subject blind?		1 var	
o Is the subject deaf?			
o Does the subject have an arm or hand missing or useless?		4 var	
o Does the subject have a leg or foot missing or useless?		4 var	
o Does the subject have any other permanent physical handicap?			@100 var
o Is the subject's mental development below normal for age?			

(1) Section 2 of the Clinical Histories files: Personal Medical History
(2) Summation of all records/individuals from section 2 of all Clinical Histories files.

(a) In Kenya the following question was asked:
Does anyone in the household have or has had any of the following: Trouble hearing/seeing/walking, paralysis, lost an arm/hand/food/leg, tuberculosis, asthma/wheezing, severe backache, cancer (site), discharging ear, diabetes, epilepsy, heart disense, high blood pressure, stroke, peptic ulcer, vomiting blood, mental disturbance, other. If yes, ICD code was recorded.
(b) It is not specifically asked for occurrence of sexually transmitted diseases and genito-urinary diseases.
(c) Information about physical handicaps is on Egypt's Physical Examination form.
(d) Section 2 of Mexico's Clinical Histories files does not give information about physical handicaps, the subject's mental development, cancer, and acute illnesses. No ICD codes are listed.
(e) Egypt's Physical Examination form has space for up to 8 final diagnosis codes, not separated by chronic and acute.

NON-BASIC VARIABLES	KENYA	EGYPT	MEXICO
o Condition: new (within last 3 mo)			
precexisting	6 var		
o History of infectious childhood diseases/scarlet fever/rheumatic fever?		8 var	
o Immunizations		5 var	
o History of Malaria/Schistosomiasis		2 var	
o Schistosomiasis treated? With what?		2 var	
o History of parasites		6 var	
o Hist. of diabetes/tuberculosis/asthma/ pneumonia/hepatitis/high blood pressure		6 var	
o Health problems in last 6 months		64 var	
o Any operations? Type and year of surgery		5 var	
o Taking any medications?		1 var	
o Medications currently taking		1 var	
o Taking any vitamins?		1 var	
o Vitamins currently taking		1 var	
o History/type/source of diagnosis of cardiopathology			3 var
o History/type/source of diagnosis of hypertension			3 var
o History/type/source of diagnosis of diabetes			3 var
o Presence/type/source of diagnosis of allergy			3 var
o History/type/source of diagnosis of convulsions			3 var

- History/type/source of diagnosis of surgery 3 var
- History/type/source of diagnosis of trauma 3 var

REPRODUCTIVE HISTORY MASTER INDEX

PROJECT	FILES	NUMBER OF RECORDS	NUMBER OF INDIVIDUALS	ESTIMATED COLLECTION FREQ
Kenya	Reproductive Hist.	462	283	taken once
Egypt	Reproductive Hist.	334	258	taken once
Mexico	Clinical Histories(1)	250	250	taken once

BASIC VARIABLES	KENYA	EGYPT	MEXICO
o Duration of present marriage	1 var	1 var	
o Age at first marriage (Kenya,Egypt)	1 var	1 var	
o Age when first pregnant (Kenya,Egypt)	1 var	1 var	1 var(a)
o Number of marriages (Kenya,Egypt)	1 var	1 var	
o # of living children		1 var	1 var(a)
o # of pregnancies	1 var	1 var	1 var(a)
o # of live births	1 var	1 var	1 var(a)
o # of abortions and miscarriages -- until 28th week of pregnancy	1 var	1 var	1 var(a,e)
o # of stillbirths -- after 28th week of pregnancy	1 var	1 var	1 var(a,f)
o # of breech deliveries	1 var		
o # of deliveries with forceps	1 var	1 var	1 var(a)
o # of deliveries by Caesarean Section	1 var	1 var	1 var(a)
o Present birth control measures	1 var	1 var	1 var(b)
o Age at menarche	1 var	1 var	1 var(c)
o # of days between menstrual periods	1 var	1 var	
o Duration of menstruation (days)	1 var	1 var	1 var(c)
o Regularity of menses	1 var	1 var	1 var(c)
o Pain during menses		1 var	1 var(c)
o Is the subject pregnant?	1 var	1 var	1 var(d)
o Is the subject lactating?	1 var	1 var	1 var(d)
o Was any child not breast fed?		1 var	1 var(d)
o # of days after birth usually started breast feeding		1 var	1 var(d)
o # of months usually breast fed children		1 var	1 var(d)
o # of months after birth usually started suppl. feeding in addition to breast milk or formula		1 var	

(1) The variables have to be taken from 4 different sections of the Mexico Files "Clinical Histories: Women" and "Clinical Histories: Pregn. and Lactat. Women".

(a) Variable is in section 4
(b) Variable is in section 10
(c) Variable is in section 3
(d) Variable is in section 5
(e) Miscarriages not counted; Not specified: "until 28th week of pregnancy";
(f) Not specified: "after 28th week of pregnancy";

NON-BASIC VARIABLES	KENYA	EGYPT	MEXICO
o Date of last menses (day, mo, yr)	1 var		1 var(g)
o Date of last delivery (day, mo, yr)	1 var		1 var(h)
o Past pregnancies: Edema?	1 var		
o Past pregnancies: High blood pressure?	1 var		
o Past pregnancies: Convulsions?	1 var		
o Past pregnancies: Anemia?	1 var		
o Past pregnancies: Vaginal bleeding?	1 var		
o Past pregnancies: Sugar in urine?	1 var		
o Age at first sexual activity			1 var(h)
o # of births			1 var(h)
o # of periosteotomies			1 var(h)
o # of premature births			1 var(h)
o # of births with low birthweight			1 var(h)
o Date of last C-section			1 var(g)
o Date of last abortion			1 var(g)
o Date of last Legrado?"			1 var(g)
o Previous gestational complications (3 previous pregnancies)			1 var(i)
o Previous birth complications (3 previous deliveries)			1 var(i)
o Previous parturition complications (3 previous deliveries)			1 var(i)
o Previous complications during lactation (3 previous lactations)			1 var(i)
o Contraceptives used for how long?			1 var(j)

(g) Variable is in section 5
(h) Variable is in section 4
(i) Variable is in section 6
(j) Variable is in section 10

PHYSICAL ASSESSMENT: NUTRITION RELATED VARIABLES
MASTER INDEX

PROJECT	FILES	NUMBER OF RECORDS	NUMBER OF INDIVIDUALS	ESTIMATED COLLECTION FREQ
Kenya	Clin. Exam - Summary	2301	891	every 6 mo
Egypt	Medical hist./Phys. Exam:			
	Physical Examination	49		targ:entry,
	Med. Hist./Phys. Exam	5902	749	6 mo; non-t.:entry
Mexico	Physical Examination:			
	Adults & children	17841	697	??
	Women w/ newborn	6040	240	??

BASIC VARIABLES	KENYA	EGYPT(a)	MEXICO(b)
o General nutritional status	5 var		
o Subject has vitamin A deficiency?			
o Subject has deficiency of B-vitamins?			
o Subject has vitamin C deficiency?			
o Subject has rickets?	1 var		
o Subject has iodine deficiency?	1 var(c)	1 var(d)	1 var(d)
o Subject has anemia?	1 var		
o Subject has fluorosis?	1 var		
o Hearing (normal, decreased)	1 var		2 var(e)
o Right eye vision: 20-xxx	1 var		1 var
o Left eye vision: 20-xxx	1 var		1 var
o Both eyes vision: 20-xxx	1 var		

(a) Form #2 of Egypt's Physical Examination Data contains the variables "Final Diagnosis: 1st to 7th". However, the diagnosis codes are not specific for these nutritional deficits.
(b) Section 5 of Mexico's Clinical Examination files gives information about "problems identified at the clinic" (problem code and treatment code). As of June 1987 it is not known whether the problem code is identical to the "Symptoms and illness codes" provided with the Morbidity files. If this is the case no information about recognized nutritional deficits is provided.
(c) Specific nutritional diagnosis - Goiter
(d) Enlargements of the thyroids
(e) Abnormality of eardrum, response to auditory stimulation

NON-BASIC VARIABLES	KENYA	EGYPT	MEXICO
o General nutr. diag. - Stunting	1 var		

o General nutr. diag. - Other	1 var		
o Specific nutr. diag. - Bitot's spots	1 var	1 var	
o Specific nutr. diag. - Xerophthalmia	1 var	1 var	
o Specific nutr. diag. - Keratomalacia	1 var	1 var	
o Specific nutr. diag. - Cheilosis	1 var	1 var	
o Spec. nutr. diag. - Angular stomatitis	1 var		
o Spec. nutr. diag. - Atrophic tongue	1 var		
o Spec. nutr. diag. - Magenta tongue	1 var		
o Spec. nutr. diag. - Follicular keratosis	1 var		
o Systolic and diastolic blood pressure	2 var	2 var	2 var
o Height and weight	2 var		
o Body temperature	1 var	1 var	1 var
o Size of tonsils/thyroid/liver/spleen	4 var		
o Immunization	7 var		
o Non-nutritional diagnosis	6 var		
o Appearance (normal, acutely/chron. ill)		3 var	
o Condition of the hair		5 var	
o Eyes (present, sight: yes,no)		3 var	
o Condition of the lids		6 var	
o Condition of the conjunctiva/cornea		9 var	
o Lips: angular scarring		1 var	
o Condition of the gums		3 var	
o Condition of the tongue		9 var	
o Condition of the skin		9 var	
o Condition of the ears		21 var	
o Skin texture, skin lesions			2 var
o Site of skin lesions			1 var
o Changes in the scalp/hair			2 var
o Color of the sclerotic/conjunctiva			2 var
o Alterations of the eyelid			1 var
o Alterations of the background of the eye			1 var
o Condition of the lips/gums/palate			3 var
o Condition of the oral mucosa/tongue			2 var
o Alteration of carotid arteries			1 var
o Tachycardia			1 var
o Alterations in articulations			1 var
o Site of the alteration of articulation			1 var
o Alterations of the extremities			1 var
o Site of the alterations of the extremities			1 var
o State of alertness			1 var
o Muscle tone			1 var
o Other medical conditions		111 var	38 var

PHYSICAL ASSESSMENT SPECIFICALLY DURING PREGNANCY MASTER INDEX

PROJECT	FILES	NUMBER OF RECORDS	NUMBER OF INDIVIDUALS	ESTIMATED COLLECTION FREQ
Kenya	Clin. Exam - Summary	2301(1)	891	2x/year
Egypt	Medical Hist./Phys. Exam:			
	Physical Examination	49(1)		
	Med. Hist./Phys. Exam	5902	749	
Mexico	Physical Examination:			
	Adults(2)	685(3)	78(3)	monthly

BASIC VARIABLES	KENYA	EGYPT	MEXICO
o Systolic blood pressure			1 var
o Diastolic blood pressure			1 var
o Were fetal heart sounds observed?			1 var

(1) It is not possible to identify pregnant women directly in the Egypt and Kenya files. Longitudinal pregnant females may be identified by use of the longitudinal pregnancy/infant master files developed by ME (see section VI B).
(2) By merging the variable "Individual Type" from section 5 into sections 1 to 4 it is possible to identify the pregnant women. However, the month or trimester of pregnancy can not be identified.
(3) Pregnant women only

PREGNANCY OUTCOME MASTER INDEX

PROJECT	FILES	NUMBER OF RECORDS	NUMBER OF INDIVIDUALS	ESTIMATED COLLECTION FREQ
Kenya	Pregnancy Outcome	138	138	taken once
Egypt	Pregnancy Outcome	124	110	taken once
	Dubowitz	110	110	taken once
Mexico	Clinical Histories(1)	109	109	taken once

BASIC VARIABLES	KENYA	EGYPT	MEXICO
o Date of pregnancy termination	1 var	1 var	1 var(c)
o Outcome of pregnancy	1 var	1 var	
o Difficulty of delivery	}	1 var	}
o Did the delivery require forceps?	}1 var	1 var	}1 var
o Was the delivery by Caesarean section?	}	1 var	}
o Did the delivery involve hemorrhage?		1 var	
o Plurality	1 var	1 var	1 var(d)
o Persons attending the birth	1 var	1 var	1 var
o Birth weight of infant in kg	1 var	1 var	
o Infant's age (in hours) when birth weight measured	1 var	1 var	
o Birth length of infant in cm.			
o Infant's age (in hours) when birth length measured			
o Infant's weight in kg at approximately 8 days			
o Infant's supine length in cm at approximately 8 days			
o Head circumference of infant in cm at approximately 8 days			
o Left mid-arm circumference of infant in cm at approximately 8 days			
o Infant's age (in days) when anthropometry measurements taken			
o Gestational age in weeks: by menstr. hist.	1 var	1 var	Cl.Ex.(e)
o Gestational age in weeks: by Dubowitz score	1 var	1 var	
o Age in hours when scored	1 var	1 var	
o Did the infant cry immediately after birth?			
o Did the infant exhibit respirat. distress?	}		
o Did the infant have an infection?	}		
o Did the infant have a birth injury?	}3 var(a)		
o Did the infant have a congenital abnormality?	}		
o If yes, identify (projects to provide codes)	}		
o Does the mother require medical follow-up?	3 var(b)		

(1) Some of the variables can be found in the file "Clinical Histories. Children", section 7: History of last birth. 109 of the 390 records contain infant data. However, no da'a is available for cases of stillbirth, miscarriage and abortion.

(a) 3 variables: Abnormal findings - infant (ICD name & code)
(b) 3 variables: Abnormal findings - mother (ICD name & code)
(c) Date of birth
(d) Variable: Numero de Producto??
(e) The gestational age of 42 infants can be found in the file "Clinical Examination: Children", section 1. The method by which the gestational age was derived is not mentioned.

NON-BASIC VARIABLES	KENYA	EGYPT	MEXICO
o Where was the delivery?	1 var		1 var
o Apgar (1 minute)	1 var		
o Apgar (5 minutes)	1 var		
o Infant physical condition	1 var		
o Matern. PF condition	1 var		
o Time of delivery	1 var		
o Age of infant in days for Dubowitz score		1 var	
o Total weight of infant and clothing in kg		1 var	
o Weight of infant's clothing in kg		1 var	
o Gestat. age in weeks: by date of quickening		1 var	
o Gestat. age in weeks: by fundal height		1 var	
o Type of abnormal birth			1 var

MONTHLY CONCEPTION SURVEY MASTER INDEX

PROJECT	FILES	NUMBER OF RECORDS	NUMBER OF INDIVIDUALS	ESTIMATED COLLECTION FREQ
Kenya	Pregnancy Survey	3028	287	monthly
Egypt	Pregnancy Monthly Survey	220	113	once
Mexico	Clinical Histories:(1)			
	Women	537	177	every 3 mo
	Pregn. or lact. Women	1221	180	monthly

BASIC VARIABLES	KENYA	EGYPT	MEXICO
o Visit number	1 var		
o Date of last menses	1 var	1 var	1 var
o Was date of last menses established by: Exact calender, calender of events			
o Urine sample obtained for preg. test?	1 var		
o Is the subject pregnant?	1 var		1 var

(1) Data available in section 5; Obstetric Data

PREGNANCY SURVEY MASTER INDEX

PROJECT	FILES	NUMBER OF RECORDS	NUMBER OF INDIVIDUALS	ESTIMATED COLLECTION FREQ
Kenya	Pregnancy Survey	3028	287	monthly
Egypt	Pregnancy Monthly Survey	220	113	??
Mexico	Clinical Histories(1)	7411	138	monthly

BASIC VARIABLES	KENYA	EGYPT	MEXICO
o Date of last menses	1 var	1 var	1 var(g)
o Apparent month of gestation	calc.(a)	1 var(e)	
o Has the subject experienced morning sickness?	1 var	1 var	
o Has the subject experienced cravings in the past week?	1 var(b)	1 var	
o Has the subject experienced quickening/fetal movement in the past wk?	1 var(b)	1 var	1 var(b,h)
o Does the subject have lower limb edema?	1 var(c)	1 var	1 var(i,h)
o Is the subject breast feeding?	1 var	1 var	
o Does the subject take any medications or supplements?	1 var	1 var(f)	1 var(j,h)
o Does the subject take iron supplements?	1 var(d)	1 var	
o Length of use of iron (weeks)		1 var	
o Does the subject take B-complex supplements?		1 var	
o Length of use of B-complex (weeks)		1 var	
o Does the subject take other mineral/vitamin supplements?		1 var	
o Length of use of other mineral/vitamin supplements (weeks)		1 var	
o Did the subject obtain supplements from health center?		1 var	
o Is the subject taking prescription medication?			
o Length of use of prescr. medication (wks)			
o Is the subject taking herbal medication?			
o Length of use of herbal medication (wks)			
o Is the subject taking other medication?			
o Length of use of other medication (wks)			
o Does the lead male smokes cigarettes?		1 var	

(1) Clinical Histories: Pregnant or Lactating Women, Section II: History of Present Pregnancy

(a) Month of gestation can be calculated from expected date of delivery and date of interview.
(b) Question does not specify "in the past week"
(c) Question does not specify "lower-limb"
(d) Answer is included in the previous variable (Does the subject take any medication or supplements?)
(e) Gestational age given in weeks
(f) Supplements only
(g) Answer in section 5: Obstetric data
(h) Answer is included in the variable "Code of the Activity"
(i) Edema of the legs or face
(j) Medications only

NON-BASIC VARIABLES	KENYA	EGYPT	MEXICO
o Are you pregnant now?	1 var		
o Expected delivery date	1 var		
o Do you have heaviness of body"	1 var		
o Do you have heaviness of breasts"	1 var		
o Do you have enlarged nipples?	1 var		
o If morning sickness - # of days	1 var		
o If cravings - # of days	1 var		
o If quickening - # of days	1 var		
o If edema - # of days	1 var		
o If heaviness of body - # of days	1 var		
o If heaviness of breasts - # of days	1 var		
o If enlarged nipples - # of days	1 var		
o Are you/will you obtain prenatal care"	1 var		
o Where will you deliver?	1 var		
o Urine test - pos., neg., not done	1 var		
o - Xanthomatous illness			1 var
- Transvaginal hemorrhage			
- Leucorrhea			
- Varicose veins			
- Vomiting			
- Headaches			
- Abdominal growth			
- Rupture of membranes			
- Masses in the breast			
- Pigmentation of nipples			
- Reddening of nipples			
- Painful uterine contractions			
- Increased food intake			
- Decreased food intake			

LACTATION AND INFANT FEEDING MASTER INDEX

PROJECT	FILES	NUMBER OF RECORDS	NUMBER OF INDIVIDUALS	ESTIMATED COLLECTION FREQ
Kenya	Lact. Questionnaire	867	140	??
Egypt	Lact./Infant Feeding	685	111	??
Mexico	Clinical Histories(1)	109	109	taken once

BASIC VARIABLES	KENYA	EGYPT	MEXICO
o Were milk or colostrum samples taken?	1 var		
o How is the baby fed?	1 var	1 var	2 var(a,c)
o Is the mother breast feeding another child?	1 var	1 var	
o Is the infant breast fed by someone other than his/her mother?	1 var	1 var	
o Have food supplements been introduced?	1 var	1 var	
o Did the mother receive advice about feeding the baby?	1 var	1 var	
o # of days after birth baby was first put on breast	1 var	1 var	1 var(a)
o Is the baby fed on demand?	1 var	1 var	
o Does the baby get both breasts?	1 var	1 var	
o Is the breast milk adequate?	1 var	1 var	
o Type of formula used	5 var(c)	1 var	1 var(b)
o # of times during days baby is formula fed	1 var	1 var	
o # of times during night baby is formula fed	1 var	1 var	
o Do you dilute fresh milk?	)	1 var	)
o ml of fresh milk in ratio of milk/water	)1 var	) var	)1 var(b)
o ml of water in ratio of fresh milk/water	) (d)	1 var	) (d)
o Measures of dry milk powder in ratio of dry milk powder to water		1 var	
o ml of water in ratio dry milk powder/water		1 var	
o Is the formula water boiled?	1 var	1 var	
o Why was supplementation started?	1 var	1 var	2 var(f)
o Are supplements adequate?	1 var		(a,b)
o Have there been side effects from suppl.?	1 var	1 var	
o Breast milk: secretory IgA level: mg/dl			
o Breast milk: secretory IgG level: mg/dl			

(1) Data are available in sections 8 and 9 of the file "Clinical Histories: Children". Data from 390 children were collected, 109 of which were infants.

(a) Variable is included in section 8.
(b) Variable is included in section 9.
(c) What kind(s) of other food(s) do you give your baby? (food code)

-- 5 different answers possible
(d) Ratios are given, not ml of milk and water
(e) Did child receive breast milk? Did child receive artificial milk?
(f) Reason for introduction of foods; Reason for weaning child

NON-BASIC VARIABLES	KENYA	EGYPT	MEXICO
o Plain water given to baby?	1 var		
o # of times baby is breast fed during day	1 var		
o # of times baby is breast fed during night		1 var	
o Aver. # of minutes baby feeds on each breast	1 var		
o Always start with same breast?	1 var		
o If other foods given to baby: # of days food #1 to #5 given to baby	5 var		
o Is baby fed with a bottle?	1 var		
o If baby is given formula, are preparation instructions followed?	1 var		
o If preparation instructions not followed: - more water added - more powder added than advised	1 var		
o First food fed to baby			1 var(a)
o Age at which lactation was terminated			1 var(a)
o Initiation of use of material milk substitutes (months of life)			1 var(b)
o Age at which other foods were introduced to child's diet (in months)			1 var(b)
o Food which was first introduced to child's diet			1 var(b)
o Did child have diarrhea during introduction of foods?			1 var(b)
o How long did diarrhea last? (introd. of foods)			1 var(b)
o Did child have diarrhea during weaning?			1 var(b)
o How long did diarrhea last? (no longer breast fed)			1 var(b)
o Age at which child began eating family diet			1 var(b)

(a) Variable is included in section 8.
(b) Variable is included in section 9.

LABORATORY ASSESSMENTS, INCLUDING IMMUNOLOGY MASTER INDEX

PROJECT	FILES	NUMBER OF RECORDS	NUMBER OF INDIVIDUALS	ESTIMATED COLLECTION FREQ
Kenya	Laboratory			
	Physiologic Samples	2067	866	twice
	Immunology/Water	1025	676	2x/yr
Egypt	Lab assessment/Immunology			
	Hematology/Urine	3961	1048	
	Immunology	655	167	
	Parasitology	668		
	Water Microbiology	229		
Mexico	Blood Work-Hematology	537	504	
	Blood Work-Immunology	536	504	
	Blood Work-Biochemistry	1755	504	
	Urine Analysis	9085	718	
	Fecal Analysis	2018	969	

BASIC VARIABLES	KENYA	EGYPT	MEXICO
o ID code of person drawing blood for blood test	1 var		1 var
o Date blood sample drawn		1 var	1 var
o Time blood drawn: hour/minute		(b)	
o Hemoglobin (g/dl)	1 var	1 var	1 var
o IgG level: mg/dl	(a)		
o Serum albumin level: mg/dl	(a)		

(1) Range checks archived but documentation incomplete
(2) Neither adequate documentation nor complete range checks archived
(3) File has not yet arrived in Berkeley

(a) The variables "IgG Level" and "Serum Albumin Level" are supposed to be in the file Immunology/Water Contamination. No values have been recorded here.
(b) The variable "Time Blood Drawn" is supposed to be in the file Hematology/Urine. No value has been recorded here.

NON-BASIC VARIABLES	KENYA	EGYPT	MEXICO
Hematology:			
o Birthdate and age of subject		1 var	
o Subject's health status (well, ill)			1 var
o Subject's physiological state		1 var	1 var
o Hematocrit	1 var	1 var	1 var
o Malaria	1 var		
o R.B.C. Stain	1 var		
o R.B.C. Morphology	1 var		
o R.B.C. Count			1 var
o MCV			1 var
o White cell count	1 var		1 var
o PMN	1 var		
o Lymphocytes	1 var		1 var
o Eosinophils	1 var		1 var
o Basophils	1 var		1 var
o Monocytes	1 var		1 var
o Neutrophils			1 var
o Nuc. RBC	1 var		
o Hyperseg. PMN (%) / 100 PMN	1 var		
o Ferritin	1 var	1 var	
o Other measurements			x var
o E. Rossette (%)	1 var		
o Small Rossette (%)	1 var		
Urine analysis:			
o Birthdate and age of subject		1 var	
o Subject's health status			1 var
o Subject's physiological status		1 var	1 var
o Individual is on medication?			1 var
o ID # of medication			5 var
o Glucose	1 var	1 var	1 var
o Ketones	1 var	1 var	1 var
o Protein	1 var	1 var(a)	1 var
o Hematuria	1 var		1 var
o pH		1 var	1 var
o Specific gravity		1 var	
o Density			1 var
o Total bacteria			1 var
o Gram negative			1 var
o Smell/color/appearance of urine			3 var
o Bilirubin/urobilinogen			2 var
o Nitrites			1 var
o Urate, phosphate (amorphous)			2 var
o Ammonium urate			1 var
o Epithelial cells, leukocytes			2 var
o Mucoid filaments			1 var
o E. vermicularis		1 var	1 var
o Microscopic analysis			1 var
Fecal analysis:			
o Subject's health status			1 var
o Subject's physiological state			1 var

o Medication given during previous 48 hr[a]			1 var
o ID # of medication			5 var
o Consistency of feces			1 var
o Presence of mucus/blood/parasites			3 var
o Type of parasites			1 var
o Ascaris	1 var		1 var
o Hookworm	1 var		
o Giardia	1 var		2 var
o Amoeba	1 var		
o Strongyloides	1 var		
o Trichuris	1 var		
o Taenia	1 var		
o Hymenolepis	1 var		
o Enterobius	1 var		1 var
o Schistosomae	1 var		
o Liver flukes	1 var		
o E. histolytica		1 var	2 var
o Other protoz. trophozoites and cysts		7 var (")	2 var
o Hymenolepis N.			1 var
o Other helminths types 1 and 2			2 var
o Amt. of bacteria microscop. visible			1 var
Breast milk analysis:			
o Creamatocrit (%)	1 var		3 var
Skin Tests:			
o Date administered	1 var		
o PPD - Tuberculin	3 var		
o Candida	3 var		
o Tetanus	2 var		
Household Water Contamination:			
o Date of collection/analysis	2 var	1 var	
o Source of water	1 var	1 var	
o Bacterial count - Total	1 var	1 var	
o Bacterial count - Total plate negative	1 var	1 var	

Note: A variable appearing in more than one project might have different value labels or discrepancy.

(a) Amount of albumin in urine

MORBIDITY/ILLNESS EPISODES MASTER INDEX

PROJECT	FILES	NUMBER OF RECORDS	NUMBER OF INDIVIDUALS	ESTIMATED COLLECTION FREQ
Kenya	Morb.-Indiv. Four Weekly Report(1)	11694	983	weekly recall
Egypt	Morb. Illness Episodes	3566	706	weekly recall
Mexico	Morbidity (Mexico Method)	36522	2608	1/week(2)

BASIC VARIABLES	KENYA	EGYPT	MEXICO
o Date of morbidity visit		1 var(c)	
o Code number of the main organ system affected by the illness or the category of the illness	1 var(a)	9 var(c)	11 var(c)
o ICD code of illness	1 var	1 var(d)	
o Day number of this episode of illness			
o Date this episode of illness began	1 var	1 var(d)	1 var
o Date this episode of illness ended	1 var	1 var(d)	1 var
o How severe is the illness	1 var	1 var(d)	
o Who diagnosed this illness?	1 var	1 var(d)	1 var(f)
o What was the source of the diagnosis?	1 var	1 var(d)	
Treatment (since last visit):			
o Subject received treatment			
o Subject received traditional treatm.	1 var		
o Subj. received modern symptom. treatm.	1 var		11 var(g)
o Subj. received mod. anti-biot. treatm.	1 var		
o Subject received antihelminthics	1 var		
o Change in activity during illness	1 var	1 var(c)	1 var(h)
Lead Male and Female:			
o Did subject work outside of home (employm.) since diagn. or last visit?	1 var(b)		
o Did subject do farm work?	1 var(b)		
o Did anyone also do/help with farm work because he/she was ill?	1 var(b)		
o If yes, who helped with farm work?	1 var(b)		
o Did subject do household work?	1 var(b)		
o Did anyone also do/help with HH work because he/she was ill?	1 var(b)		
o Did subject cook meals?	1 var(b)		
o Did anyone also do/help with meal preparation because he/she was ill?	1 var(b)		
o Did subject care for children?	1 var(b)		
o Did anyone also do/help with child care because he/she was ill?	1 var(b)		
Toddler:			

o Date anthropometry measurements taken
  at the beginning of the illness
o Weight in kilograms
o Date anthropometry measurements taken
  at the end of the illness
o Intake (24-hr recall): kcals
o Intake (24-hr recall): g protein
o Intake (24-hr recall): g fat
o Intake (24-hr recall): g carbohydrates
o Did the subject have diarrhea
  on the day of report?
o # of bowel movements in past 24 hrs
o Were the stools liquid?
o Were the stools bloody?
o Were the stools foul-smelling?

---

(1) Kenya data have been collected on two different survey forms (313 and 314) Although these two data sets have been archived separately, the above counts represent the pooled data set.
(2) Morbidity data for Mexico reflect illnesses on the day of interview only.

(a) Variable included in file: Morbidity - Household Four Weekly Report
(b) Information available after October 1984 (form # 314)
(c) Variable included in file: Weekly Morbidity Recall
(d) Variable included in file: Morbidity Illness Episodes
(e) Variable "Symptoms" Note: The Mexico symptoms code is not based on the ICD code of illness.
(f) Always a physician in Mexico
(g) Variable "Medication code" Note: Only treatment on day of visit is asked for.
(h) Variable "Degree of incapacity" gives similar information

Note: The values of a variable appearing in more than one project might vary from project to project.

NON-BASIC VARIABLES	KENYA	EGYPT	MEXICO
o # of days of illness duration	1 var		
o Subject received antimalarial treatm.	1 var		
o Illness outcome	1 var		
o Change in food intake	1 var	1 var(k)	
o Task reassignment/fetching water,fuel and extend of help	2 var(i)		
o Weight change	1 var		
o Other anthropom. changes	4 var(j)		
o Present this week		1 var(k)	
o Subject absent short term		1 var(k)	
o Subject absent long term		1 var(k)	
o Menses		1 var(k)	

o Pregnant	1 var(k)
o Examined	1 var(k)
o Sickness subroutine	1 var(k)
o Age in years; Age in months	2 var(k,m)
o Change in appetite	1 var
o Modification of diet	1 var
o Application of Medication	1 var
o Date of 1st visit to HH in a particular month	[illegible]
o Visits expected (for this month)	[illegible] var
o Visits completed (in this month)	1 var

(i) Information available after October 1984 (form # 314)
(j) Information available until October 1984 (form # 313)
(k) Variable included in file: Weekly Morbidity Recall
(m) Variable included in file: Morbidity Illness Episodes

MORBIDITY SUMMARY MASTER INDEX

PROJECT	FILES	NUMBER OF RECORDS	ESTIMATED COLLECTION FREQ
Kenya	Household 4-weekly report	7211 (1)	Weekly
Egypt	Morbidity weekly recall	82672 (1)	Weekly
Mexico	Morbidity summary by calender month	33165 (2)	Weekly

BASIC VARIABLES	KENYA	EGYPT	MEXICO
o Date of quarterly morbidity report	1 var	1 var	1 var
o Number of days free from all illnesses during the current quarter			
o Number of days ill during the current quarter			1 var
o Number of days for which no morbidity information was available during the current quarter			1 var
o Total number of days morbidity information was available for the current quarter			1 var

(1) Summarizes illness events for all family members, but no data on date of illness or duration. For targets, details on illnesses are in the illness episode file.

(2) Data are summarized by 4 illness categories; missing days are estimated based on missing interviews. See Mexico Final Report for details.

INFANT PSYCHOLOGY MASTER INDEX

PROJECT	FILES	NUMBER OF RECORDS	ESTIMATED COLLECTION FREQ
Kenya	Psychology, Infant:		
	Cognitive infant summary	276	taken twice
	Infant interaction	5947	every 2 mo
	Brazelton (BNAS)	132	taken once
Egypt	Psychology, Infant:		
	Infant cognitive	32	taken once
	Infant behavior	246	taken twice
	Brazelton	101	taken once
Mexico	Psychology, Infant:		
	Infant interaction, 3 mo.	95	taken once
	Infant interaction, 6 mo.	99	taken once
	Brazelton exam	61	taken once
	Mother/child ratings	97	taken once
	Bayley motor	97	taken once
	Bayley mental	97	taken once
	IBR	97	taken once

BASIC VARIABLES	KENYA	EGYPT	MEXICO
o BNAS scores:			
State	3 var	3 var	
Elicited responses	20 var	20 var	
Descriptive paragraph	3 var	3 var	
Quieting activity	1 var	1 var	
Behavior score scale	27 var	32 var	
Habituation score (Lester)			1 var
Orientation score (Lester)			1 var
Motor score (Lester)			1 var
Range of states (Lester)			1 var
Regulation (Lester)			1 var
Autonomic stability (Lester)			1 var
Reflex score (Lester)			1 var
Process of interact. (a priori)			1 var
Motor processes (a priori)			1 var
Control of states (a priori)			1 var
Response to stress (a priori)			1 var
o Caretaker/infant observations:			
Scores, 2 mo	9 var		
Scores, 3 mo		14 var	19 var
Scores, 4 mo	9 var		
Scores, 6 mo	9 var	11 var	19 var

o Bayley motor scores, 6 mo	33 var	1 var	30 var
o Bayley BIBR scores, 6 mo	30 var	11 var	10 var
o Bayley mental, 6 mo			55 var

* Egypt Project infant psychology variables have been cleaned by the project psychologist. The revised variables have not been archived as separate psychology files, but some of the infant psychology variables have been archived with the analytic data sets for infants.

NON-BASIC VARIABLES:

o Fagan		10 var
o Visual attention test	6 var	
o Other interaction obs	Interaction with other family memb.	Number of people in home; Child's location; Response to distress

TODDLER PSYCHOLOGY MASTER INDEX

PROJECT	FILES	NUMBER OF RECORDS	ESTIMATED COLLECTION FREQ
Kenya	Psychology, Toddler:		
	Cognitive toddler summary	315	taken twice
	Cognitive todd. summ.-30 mo.	230	taken once
	Toddler interaction	9933	every 2 mo
Egypt	Psychology, Toddler:		
	Toddler cognitive-18,24 mo.	234	taken twice
	Toddler cognitive-30 mo.	59	taken once
	Toddler behavior obs	2586	taken 2x/mo
Mexico	Psychology, Toddler:		
	Bayley scores, 18 mo	101	taken once
	Bayley scores, 24 mo	118	taken once
	Bayley scores, 30 mo	96	taken once
	Bayley motor, 18 mo	94	taken once
	Bayley motor, 24 mo	117	taken once
	Bayley motor, 30 mo	96	taken once
	Bayley mental, 18 mo	94	taken once
	Bayley mental, 24 mo	117	taken once
	Bayley mental, 30 mo	96	taken once
	IBR, 18 mo	94	taken once
	IBR, 24 mo	117	taken once
	IBR, 30 mo	96	taken once
	Bayley language, 24 mo	117	taken once
	Bayley language, 30 mo	96	taken once

BASIC VARIABLES	KENYA	EGYPT*	MEXICO
o Bayley mental			
18 months	25 var (verbal only)	1 var (S)	56 var
24 months	51 var	1 var (S)	66 var
30 months	48 var	1 var (S)	66 var (S)
o Bayley motor			
18 months			14 var
24 months			19 var
30 months	16 var	1 var (S)	24 var
o BIBR (Bayley behavior)			
18 months	12 var	12 var	10 var
24 months	12 var	12 var	10 var
30 months	31 var	11 var	10 var
o Bayley language			
24 months			22 var

30 months			22 var
o Observation of care-taker/child	9 var	11 var	30 var

* Egypt Project toddler psychology variables have been cleaned by the project psychologist. The revised variables have been archived with the toddler analytic files.

NON-BASIC VARIABLES:

o Toy play		
18 months		7 var
24 months		7 var
30 months	8 var	7 var
o Additional home obs.	Interact with other members, 5 var	No. people in home; Child's location; Response to distress
o Uzgiris Hunt		
18 mo	19 var	28 var 3 var (S)
24 mo	19 var	28 var 3 var (S)
30 mo	6 var	
o Vision/hearing test	2 var	

S = scaled variable(s)

SCHOOLER PSYCHOLOGY MASTER INDEX

PROJECT	FILES	NUMBER OF RECORDS	ESTIMATED COLLECTION FREQ
Kenya	Psychology, Schoolers:		
	Cognitive schooler-summary	447	taken 2 or 3X
	Schooler classroom	516	taken 2 or 3X
	Schooler playground	1465	taken 2 or 3X
	School attendance/perform.	211	taken by term
Egypt	Psychology, Schoolers:		
	Schooler cognitive	202	taken twice
	Child behavior rating	137	taken twice
	Schooler classroom	543	taken twice
Mexico	Psychology, Schoolers:		
	Cognitive tests, schoolers	292	taken twice
	Classroom observations	213	taken twice
	Playground observations	270	taken twice
	Drawings	113	taken once
	Teacher evaluation	188	
	Int/extraversion	188	

BASIC VARIABLES	KENYA	EGYPT*	MEXICO
o Wechsler:			
Digit span	3 var	3 var	3 var
Raven's matrices	1 var	1 var	1 var
Block design	2 var	1 var	1 var
o Classroom observations	5 var	5 var	4 var
o Playground observations:			
Interaction w/ others	10 var	18 var	24 var
Level of activity	4 var	4 var	5 var
Emotion	6 var	6 var	7 var

* Egypt Project psychology varinbles for schoolers have not been examined by the project psychologist and may contain errors.

NON-BASIC VARIABLES

o Wechsler:			
Verbal meaning	1 var		
Test behavior	4 var		
Vocabulary		1 var	1 var
Similarities		1 var	
Picture completion	1 var	1 var	
Coding		1 var	
Raven's impulsivity		3 var	
Auditory vigilance		5 var	

Total lines (crosses)			1 var
Arithmetic score			1 var
Mazes			1 var
o Reads and writes			1 var
o Able to write name			1 var
o Teacher's rating data		8 var	51 var
o Time taken to do tests			1 var
o School code	1 var	1 var	
o Grade		1 var	1 var
o Years in school			1 var
o Drawings			3 var
o Int/extraversion			28 var

ADULT PSYCHOLOGY MASTER INDEX

PROJECT	FILES	NUMBER OF RECORDS	ESTIMATED COLLECTION FREQ
Kenya	Psychology, adults		
	Cognitive adult-summary	622	taken once
	Adult literacy test	408	taken once
Egypt	Psychology, adults	226	taken once
Mexico	Psychology, adults	452	taken once

BASIC VARIABLES	KENYA	EGYPT*	MEXICO
o WAIS scores:			
Block design	1 var	1 var	1 var
Digit span	3 var	3 var	3 var
Raven's matrices	1 var	1 var	1 var

* Egypt Project psychology variables for adults have not been examined by the project psychologist and may contain errors.

NON-BASIC VARIABLES			
o Similarities	1 var	1 var	
o Test behavior	4 var		
o Arithmetic	1 var		1 var
o Verbal meaning	1 var		
o Vocabulary		1 var	1 var
o WAIS total		1 var	
o Auditory vigilance		7 var	
o H marking		1 var	
o Dotting		1 var	
o Raven's impulsivity		6 var	
o Proverbs		1 var	
o Literacy	24 var		
o Able to write name			1 var
o Total lines (crosses)			1 var
o Mazes score			1 var
o Minutes elapsed			1 var
o Number of years in school			1 var

HOUSEHOLD SANITATION MASTER INDEX

PROJECT	FILES	NUMBER OF RECORDS	NUMBER OF HOUSEHOLDS	ESTIMATED COLLECTION FREQ
Kenya	Sanitation and Hygiene	1293	277	every 3 mo
	Washing Observations	421	143	3 times
Egypt	Sanitation/Hygiene-HH	933	280	every 3 mo
Mexico	Sanitation/Hygiene?			
	Sanit. Conditions and Hygiene	8920	235	monthly
	Five Sanit. Measures	235	235	once

BASIC VARIABLES	KENYA	EGYPT	MEXICO
o Household hygienic adequacy score		1 var	5 var(a)

(a) - External living area scale
- Internal living area scale
- Mother's appearance score
- School-aged child's appearance score
- Young child's appearance score

NON-BASIC VARIABLES	KENYA	EGYPT	MEXICO
o Pers. hygienic practices of least female	7 var		
o Hygienic practices regarding infant	6 var		
o Hygienic practices regarding toddler	8 var		
o Hygienic practices regarding schooler	8 var		
o Clothing care	6 var		
o Hygienic practices in kitchen	6 var		
o Rodent problem in kitchen/household?	2 var		
o Hygienic practices in household/compound	6 var		
o Crops stored in household?	1 var		
o Where are poisons+sprays mixed/stored?	2 var		
o Latrine in compound? Shared?	1 var		
o Water tap in compound?	1 var		
o Current source of water?	1 var		1 var
o Animals/birds kept in house?	1 var		
o Number of animals kept in compound yard	8 var		14 var
o Manure in compound yard?	1 var		
o Animals penned near kitchen?	1 var		

- o Animal pen well kept? 1 var
- o Flies in pen/kitchen/latrine 3 var
- o Adequ. drainage/stagn. water in cmpl.? 2 var
- o Dung heap in yard? 1 var
- o Observation of house construction 5 var
- o Type of kitchen (inside, outside, sep.) 1 var
- o Any animals in kitchen? 1 var
- o # of windows in kitchen 1 var
- o latrine - distance from house 1 var
- o latrine - distance from kitchen 1 var
- o latrine - distance from water source 1 var
- o latrine - direction to water 1 var
- o latrine - concrete slab? 1 var
- o latrine - # of flies 1 var
- o Food preparation - # of flies 1 var
- o Rating of hygiene of yard 1 var
- o Types of waste observed in the yard 4 var
- o Quantity of each type of waste in yard 4 var
- o Presence of animals in yard 1 var
- o General appearance of living quarters 1 var
- o Appearance of walls (excl. kitchen) 1 var
- o Types of waste on the floors 3 var
- o Illumination inside house 1 var
- o Quantity of animals inside of house 6 var
- o Presence of garbage inside house 1 var
- o General appearance of kitchen 1 var
- o Location of stove/foods 1 var
- o Covering on foods 1 var
- o Appearance of mother 1 var
- o Appearance of toddler(s) 1 var
- o Appearance of schooler(s) 1 var
- o Appearance of father 1 var
- o Disposal of human waste 1 var
- o Times fem. doesn't wash before food prep 1 var
- o Times fem. washes without soap 1 var
- o Times fem. washes with soap 1 var
- o Times infant is observed washed 1 var
- o Times schooler feeds other 1 var
- o Times schooler washes before eating 1 var
- o Times schooler washes before feeding 1 var
- o Times food served on dishes 1 var
- o Times dishes washed with soap 1 var
- o Times dishes washed without soap 1 var

INDIVIDUAL DEMOGRAPHIC DATA MASTER INDEX

PROJECT	FILES	NUMBER OF RECORDS	NUMBER OF INDIVIDUALS	ESTIMATED COLLECTION FREQ
Kenya	Census Update	11193	292 (1)	every 3 mo
Egypt	SES - Individual	2188	2188	taken once
Mexico	Census: Basal	7134		taken twice
	Target Individuals	1068		

BASIC VARIABLES	KENYA	EGYPT	MEXICO
o Subject's relation to head of HH	2 var	1 var	1 var
o Subject's marital status	1 var	1 var	
o Is the subject literate?		1 var	1 var
o Subject currently enrolled in school?	1 var	1 var	
o # of years of school completed		1 var	Cogn.(3)
o Type of education completed	1 var	1 var	
o Subject's occupation	SES(2)	3 var	1 var
o Subject's employment status	SES(2)	1 var	

(1) Number of households
(2) Variables "HH head: salaried employment off-shamba" and "Any business owned by HH" included in file SES (Socio-Econ)
(3) Information included in file: Cognitive Tests (Adult Psychology)

NON-BASIC VARIABLES	KENYA	EGYPT	MEXICO
o Age in years	1 var(a)	1 var	
o Birthdate	1 var	1 var	1 var
o Verification of birthdate	2 var		1 var
o Place of birth			1 var
o Year of marriage (head and wife)	1 var		
o Religion	1 var		
o Three languages best spoken	1 var		
o Tribe	1 var		
o Clan	1 var		
o HH member residential status	1 var		
o Time spent out of subloc in past year	1 var		
o How often member comes home	1 var		
o Reason for absence	1 var		
o Status of HH members	1 var		
o Month of death	1 var		
o Cause of death	1 var		

o Total # of persons living in HH	1 var
o How long HH been living in subloc (yrs)	1 var
o HH counted in 1979 census?	1 var

(a) Age given in months if below 5 years

HOUSEHOLD SOCIOECONOMIC DATA MASTER INDEX

PROJECT	FILES	NUMBER OF RECORDS	NUMBER OF HOUSEHOLDS	ESTIMATED COLLECTION FREQ
Kenya	SES (Socio-Econ)(1)	1373	285	every 3 mo
Egypt	SES - Household scores	312	312	once
Mexico	SES	545	339	twice
	SES Scores	331	314	twice

BASIC VARIABLES	KENYA	EGYPT	MEXICO
o Type of dwelling		1 var	
o Material used for roofing	1 var(a)		1 var(f)
o Material used for walls	1 var(b)		2 var(f)
o Material used for flooring	1 var(c)		1 var(f)
o Type of windows	2 var	1 var	1 var(i)
o Type of doors	1 var(d)		2 var(j)
o Quality of house construction			
o Number of buildings			
o # of rooms for excl. use of household		1 var(e)	1 var(f)
o Separate kitchen?	Sa/Hy*	1 var	
o Type of cooking facilities		1 var(f)	1 var(k)
o Source of drinking water		1 var(f)	Sanit.**
o Water quality			
o Type of toilet system	Sa/Hy*	1 var(fg)	Sanit.**
o Source of light		1 var(h)	
o Type of garbage collection system	Sa/Hy*		Cl.H.***
o Social status score of household		}1 var	}2 var(m)
o Economic status score of household		}	}

(1) Kenya Project also has a file of sub-chief social and economic rankings, and a file of household economic questionnaires.

(a) Iron roof: yes,no
(b) Milled timber: yes,no
(c) Cement floor: yes,no
(d) "Purchased" door: yes,no
(e) Not specified "for exclusive use of HH"
(f) More than one response possible
(g) Bathing and latrine facilities
(h) Does the HH have electricity?
(i) # of windows with glass
(j) # of wooden/iron doors
(k) # of gas stoves
(m) Two SES measures: initial visit & second visit

* Some information included in file: Sanitation and Hygiene

** Information included in file: Sanitary Conditions and Hygiene
*** Information included in file: Clinical Histories, Section: Profile of Ambient Hygiene and Sanitation

NON-BASIC VARIABLES	KENYA	EGYPT	MEXICO
o Total acreage owned/cultivated by HH	2 var		
o # of farm animals	3 var	12 var	10 var
o Income/employment of HH head	2 var		
o Income producing business owned by HH	1 var		
o Possessions/assets owned by HH	1 var	3 var	14 var
o Did HH hire any farm or HH labor (for cash) in last 3 months?	1 var		
o How is HH compared to 3 mo ago?	1 var		
o Head is consulted in family matters?	1 var		
o Were school expenses for HH children met?	1 var		
o HH is paying secondary school fees?	1 var		
o Highest class attained by head	1 var		
o All prim. school age childr. attend school?	1 var		
o Head or wife has attended F.T.C. agric. instruction?	1 var		
o Do HH children attend Sunday school or services regularly?	1 var		
o Head or wife has in last 3 months sent or received mail, used the telephone/a bank	3 var		
o Head or wife has an NSSF card/an account	2 var		
o Meetings/activities/memberships	10 var		
o Head and/or wife permanently employed?	1 var		
o Head/wife has at least one sibling who is currently employed	2 var		
o Head is considered a clan leader	1 var		
o Skills of head	1 var		
o Has HH been visited by agric./vet/forest personnel in the last 3 months?	1 var		
o Did HH use improved agric. methods in current season?	1 var		
o Head has planned trees on their land	1 var		
o Compound yard has been made pretty	1 var		
o HH has made improvement in compound/ hh structures during last 3 months	1 var		
o # of occupants of dwelling		1 var	1 var
o Ownership of dwelling		1 var	
o Year last painted house			1 var
o # of windows			1 var
o # of doors			1 var
o Weekly expenditures on purchased food			1 var
o # of agricultural tools/machines			16 var
o Subject's health			1 var

COMMUNITY CLIMATIC DATA MASTER INDEX

PROJECT	FILES	NUMBER OF RECORDS	ESTIMATED COLLECTION FREQ
Kenya	Weather Data	566	daily (accumultd. by month)
Egypt	(1)		
Mexico	Climate	1151	

BASIC VARIABLES	KENYA	EGYPT	MEXICO
o Date climatic data obtained	1 var		1 var
o Average daily temperature (°C)	3 var(a)		1 var(c)
o Average rainfall (cm)	1 var(b)		1 var
o Average barometric pressure			1 var

(1) No data archived

(a) The temperature at 7:30 AM, the previous night's lowest temperature, and the previous day's highest temperature have been recorded.
(b) in mm
(c) Not clear if the values represent average daily temperatures.

NON-BASIC VARIABLES	KENYA	EGYPT	MEXICO
o Humidity (dry and wet readings)	2 var		

CHILD CARE AND SANITATION ACTIVITIES OF LEAD FEMALE
MASTER INDEX

PROJECT	FILES	NUMBER OF RECORDS	NUMBER OF INDIVIDUALS	ESTIMATED COLLECTION FREQ
Kenya	Care-Giving Activities	30969	264	every 2 months
Egypt	Childcare Frequency (1)	2178		
Mexico	(2)			

BASIC VARIABLES	KENYA	EGYPT	MEXICO
Lead female's proficiency score in performing care-giving activities for the target child with respect to each of the following areas:	(a)		
o Nutritional activities			
o Hygienic practices			
o Illness care			
o Clothing care			
o Supervision and safety			
o Cleanliness of infant's or toddler's environment			

(1) Data received after close of archive; included in archive, but no ME descriptions available.
(2) Data not collected

(a) The Kenya file does not contain proficiency scores. Instead, the results of two hours of observation have been directly recorded (see non-basic variables below).

NON-BASIC VARIABLES	KENYA	EGYPT	MEXICO
o Date of observation	1 var		
o Starting time (minutes and hours)	2 var		
o Ending time (minutes and hours)	2 var		
o Activity code	1 var		
o Interaction codes	6 var		
o Location of child/mother	2 var		