

Title

Consolidated Second Quarterly Report

1. Second Quarterly Report of the Program Director (and Quarterly Reports from Country Projects)
2. Report of the TAG on Statistical Design and Data Management
3. Minutes of Institutional Council conference call
4. Schedule of Meetings, April 1982

Project

Functional Implications of Malnutrition (CRSP)
Project Number 931-1039

Type of Report

Technical Report

Submitted by: G. H. Benton

April 1, 1982

1

Second Quarterly Report

In December, 1981, it was reported that in spite of administrative delays and uncertainties of the past, the Nutrition CRSP was active and would initiate field work in the near future. Administrative problems have continued but, as predicted work has proceeded, albeit not as smoothly as had been hoped.

Funding mechanisms are in place for the Kenya project (University of California, Los Angeles and Berkeley) and negotiations are underway with Kenyan colleagues. California sent subagreement documents to Connecticut and Purdue in January. During the last week of March responses have been received from both. Connecticut is forwarding a subagreement incorporating some modifications from the original offer. This will be reviewed by California. In the meantime, discussions have been initiated toward an agreement with Mexican colleagues. Purdue has submitted a request for a number of changes in the offered agreement, some quite major and potentially requiring modification of the other agreements. These are under consideration by California at present. Formal steps to negotiate an agreement with Egyptian colleagues have been deferred pending resolution of a Purdue-California agreement. Purdue has been in discussion with Arizona and Kansas and it is expected that agreements among these institutions can be consummated rapidly after a California-Purdue agreement is signed. It is not possible to predict how much longer will be required to implement the multiple tiers of institutional agreements.

In spite of the major administration delays, the projects have moved ahead. The Kenya project has two US staff permanently in the field working with Kenyan colleagues in developing field logistics and working toward confirmation of community selection. This work has been supported by short-term visits by one of the Principal Investigators, Dr. Charlotte Neumann and other senior personnel. Plans are now underway to bring Kenyan staff to UCLA for training in preparation for the CRSP work.

A similar situation holds for the Mexico project. Co-Principal Investigator Lindsay Allen is now based in Mexico working with Dr. Chavez and colleagues in developing the logistics and infra-structure of the research program and in collecting further data relevant to community selection. In the Mexican project considerable background information exists, from previous studies by the same investigative groups, about the probable community site. Again work in Mexico has been supported by shorter term visits by other senior investigators.

Progress with the Egypt project encountered a major set-back when advice was received from AID that work related to community selection could not proceed until a formal agreement with the Egyptian institution was in place. As noted above, progress with this awaits progress in the Purdue-California agreement. US investigators have had no direct involvement in Egypt since that time. This is not to say that the Egypt Project is not moving ahead. The delay in field work has permitted further progress in US-based research planning.

Field directors have been selected for the Mexican and Egyptian projects. It is expected that they will participate in planning meetings scheduled for April to increase their awareness of the collaborative research program. Recruitment of a field director for the Kenya project is now at an advanced level. If a decision is taken in the immediate future, he/she also will participate in the New Orleans meetings.

A major disappointment has been the inability of California to recruit the internationally recognized Program Director that has been sought. Six strong individuals expressed a strong interest in the position. Five of these expressed a sincere desire to take on the position and discussion of actual arrangements was undertaken. In each case, the individuals finally withdrew their names because they were unable to break ties with their existing positions or because California was unable to provide the guarantee of tenure or the financial arrangements that would be required to permit their transfers. Dr. Beaton will leave Berkeley on April 10 but has agreed to participate in the April meetings in New Orleans and in the tentatively planned June meetings (see below). California is now considering alternative arrangements for program directors involving a different mix of skills and experience and perhaps drawing upon consultative support for a program coordinator.

The Management Entity has recruited a very well experienced Business Officer, Lin Schlyer, who has worked diligently to establish administrative systems to facilitate CRSP business and to control budget. Recently a book-keeper/secretary Felomena Brown has been recruited to support this activity in anticipation of the signing of agreements and the flow of financial documents. We regret that our secretary, Alicia Elwin, the first staff person recruited for the Nutrition CRSP, is transferring to another position in the University. A replacement is under recruitment. For the immediate future, and at least until the evolution of the ME role to data management takes place, this will constitute the full core staff. The Program Director's office continues to be indebted to the support provided by the staff of the Institute of International Studies in which the Nutrition CRSP is based. We acknowledge also the support provided by the senior administrative staff of the University of California (Berkeley) and thank them for the degree to which the unusual needs and demands of the Nutrition CRSP have been fitted into the administrative structures of the University.

A major task in recent months has been the organization of a series of planning meetings. Technical Advisory Groups (TAGS) on Statistical Design and Data Management, on Activity and on Cognition were convened in January, February and March. The report of the first of these groups is submitted with this quarterly report. Other reports are not yet available.

Five more TAG meetings (Social Competence, Reproduction, Food Intake, Disease, Nutritional Status) are scheduled in April during the meetings of the Federation of American Societies of Experimental Biology in New Orleans. A meeting of the Scientific Coordination Board will be held at that time as well. For this purpose, PIs from the country institutions as well as all US

PIs will be present. A major agenda item will be a review of the status of field studies and consideration of the timing and approach to a major Nutrition CRSP Research Design meeting (tentatively scheduled for June). Investigators from all of the country projects, as well as external advisors, have been very active participants in these meetings, again documenting the collaborative spirit of the Nutrition CRSP.

With the concurrence of BIFAD, the nominee of the SCB Pro Tem, Dr. Elise Boulding was invited to join the External Advisory Panel, but has declined. In April, SCB will consider another appointment to bring the EAP to its required quorum. BIFAD has suggested it would be advantageous (albeit expensive) to increase representation from developing countries. California sees the need to include additional statistical expertise in the membership. It has been made clear that it would be unwise to appoint persons to the EAP who have had membership on the TAGs (or visa versa), a very unfortunate constraint. Recently AID staff recognized that the present wording of the AID-California agreement precluded the involvement of the EAP in the review of the research design for Phase 2, a role that all had anticipated. AID staff have agreed to initiate an amendment to permit this important function.

Discussions have been held with Dr. M. Forman with regard to the composition and operation of the Policy Advisory Group. It is expected that a meeting of a core group will be convened in connection with the CRSP Research Design meeting (tentatively in June).

The Institutional Council "met" by telephone conference call to discuss some matters that had arisen in the area of interinstitutional policy (minutes attached). The first actual meeting is planned in connection with the CRSP Research Design meeting (tentatively in June) and the IC will meet annually thereafter.

The rotation of SCB membership among US PIs has now been decided and a plan for the next several years is attached.

AID has committed \$1400 K in addition to the previous commitment of \$750 K. This will support planned expenditures through November. Further funding will be requested, based upon anticipated needs, in the early fall.

Reports from the three country projects are attached for information as a part of this quarterly report of the Program Director.

In approaching the termination of my service as Interim Program Director of the Nutrition CRSP, I would be remiss if I did not go on record with a personal statement and comment. I accepted this voluntary role in the firm belief that the purpose and goal of the Nutrition CRSP was important as well as scientifically worthwhile. I came with the strong conviction that we in the scientific community had a unique opportunity to collaborate in the development of truly interdisciplinary research. I recognized that in the process we had accepted a major responsibility not only to our sponsors but also to our colleagues and to the people we would study and their counterparts throughout

the world. I will leave this position with the same belief, conviction and sense of responsibilities. This Nutrition CRSP is important, relevant and unique. I regret very much that the CRSP has been plagued by administrative difficulties and, at times, by tendencies to fall into the trap of unidisciplinary thinking. I offer a plea to my colleagues, to their institutions and to AID--let us not lose sight of the meaning, the sensitivity and the immense potential benefit of true collaboration among scientific colleagues with a sincere interdisciplinary spirit. Let us not allow institutional considerations or necessary bureaucratic aspects of governmental, and intergovernmental operations destroy or even seriously hinder the true Nutrition CRSP--the collaboration of biological and social scientists sharing a common purpose and goal. I remain convinced that however difficult the path may be the goal is achievable. I thank all for having given me the opportunity to contribute to this program in my own small way and I look forward to the growth and success of the Nutrition CRSP.

George H. Beaton

NUTRITION CRSP
SCIENTIFIC COORDINATION BOARD
Membership by Calendar Year

Category of Membership	Year				
	1982	1983	1984	1985	1986
<u>U.S. PI</u>					
Egypt	G. Harrison	R. Watson	N. Jerome	G. Harrison	R. Watson
Kenya	C. Neumann	C. Neumann [†]	C. Neumann [†]	C. Neumann [†]	C. Neumann [†]
Mexico	L. Allen	G. Pelto	L. Allen	G. Pelto	L. Allen
Social Scientist	G. Pelto	--	--	--	R. Jerome
<u>Country PI</u>					
Egypt	*	*	*	*	*
Kenya	J. Kagia	J. Kagia	J. Kagia	J. Kagia	J. Kagia
Mexico	A. Chavez	A. Chavez	A. Chavez	A. Chavez	A. Chavez

The Nutrition CRSP Principal Investigator and the Program Director are ex officio members.

[†] Rotation not yet defined

* To be named

RECEIVED MAR 19 1982

TRAVEL

- 1) Site - U.C. Berkeley - January 15, 1982
Purpose - Field orientation for Dr. E. Carter and D. Cattle
Others conferred with - G. Beaton; D. Calloway, J. King, K. Beros, A. Little
M. Halderman
- 2) Site - Kenya - 2/21/82 - 3/15/82

Purpose of Trip

1. Finalization of collaborative arrangements with the University of Nairobi, Faculty of Medicine.
2. Implementation of phase I - preliminary survey.

Activities

1. "Settling in of 2 UC field staff Drs. Dorothy Cattle and Eric Carter"
 - i. Introduction of team to the Kenyan group at the University of Nairobi School of Medicine, Dept. of Community Health and Pediatrics; USAID Kenya Mission, University of Nairobi Computer Center; Central Bureau of Statistics; Medical Research Center; Public Health Laboratories; Ministry of Health; Embu district provincial and district health and nutrition officers.
 - ii. Banking temporary transportation and living arrangement.
 - iii. Orientation to Nairobi and Embu district.
 - iv. Discussion of honorary teaching appointments in Dept. Community Health and integration into the department.
2. Collaborative Arrangements with the University of Nairobi Faculty of Medicine.
 - i. Initiation of Formal Research clearance by the Council of Science and Technology, Office of the President, Kenya.
 - ii. Review of the protocol, organizational charts, lines of authority, designating of counterparts for each of the project personnel, review of time tables and overall work plan.
 - iii. Review and revision of University of Nairobi budget, particularly the pay scale changes.
 - iv. Review and exchange of ideas between Kenyan colleagues and UC group re: the phase I and phase II activities.
 - v. Setting up of CRSP office in Dept. Community Health.
 - vi. Review of draft sub contract document for the University of Nairobi subcontract which is to be reviewed by the legal and fiscal offices of the University of Nairobi prior to submission to UCLA subcontract office.
 - vii. Initiation of permit for tax exemption and duty-free import privileges for research supplies and equipment.

- viii. Discussion of content for an operating procedure manual for the UC and University Nairobi team.
- ix. Discussion of fiscal management and accounting requirements by USAID and UCLA and Univ. Nairobi.
- x. Initiation of application for a UC External Dollar in Kenya for local field expense.
- xi. Examination of University of Nairobi Computer facility and discussions and initiation of arrangements for design of research form, data management and analysis for preliminary survey.
- xii. Discussion of counterpart + training in USA.

3. Survey

- i. Identifications of background health, nutritional, economic and information from the Ministries of Health, Central Bureau of Statistics, Ministry of Economic and Financial Planning and Institute of Studies, Embu district Provincial Offices of Health, Statistics and Nutrition.
- ii. Construction of preliminary survey protocol jointly with Kenya - UC group and the Computer Center staff.
- iii. Preliminary site visit to Embu Provincial three sub-locations for potential project sites and meeting with Provincial and district health officials who guided the "tour". The visit was primarily of the semi-arid lowland areas of Embu district, a subsistence farming area.
- iv. Interview of Kenyan and in country nutritionists interested in working on Nutrition CRSP.

Individuals Visited:

1. University of Nairobi School Medicine

- Dr. James Kagia - Chairman Dept. Community Health and Co-Principal Investigator of the Nutrition CRSP. Interests - morbidity patterns and epidemiology
- Dr. Kasiji - Epidemiologist and infectious diseases, Dept. Community Health
- Dr. Nimrod Dwibo, Chairman, Dept. Pediatrics and Co-Principal Investigator Nutrition CRSP. Interests - Lactation performance physical growth of children.
- Dr. Julius Meme - Dept. Pediatrics Co-Investigator - Child Development.
- Dr. A. Kinoté - Dept. Pediatrics, Co-Investigator - Child nutrition, growth and development and food intake.
- Dr. A.A. Jansen - Medical Research Center and Dept. Community Health - Co-Investigator. Interests - Nutritional Assessment and Nutritional Aspects of Pregnancy, Pregnancy outcome and lactation, food intake.
- Dr. Bowry - Dept. Immunology - Co-Investigator - Immunologic assessment

Dr. Thairu - Chairman - Dept. Immunology, Co-Investigator - Work Capacity and Neurosensory development and nutritional status.

Computer Center - University of Nairobi

Dr. Scott - director

Mr. Corhill - data analysis, programmer

Mr. Kathito - programmer

Ministry of Health

Mrs. Ngui - director of Community Nutrition

Dr. Sehmi - Food analysis and food composition

Central Bureau of Statistics

Mr. Kekovale - director of Statistical services and nutritional and economic surveys - country-wide.

USAID - Kenya Mission

Dr. Rose Britanak - Health officer

Dr. Ned Greeley - Program Officer

Dr. Patrick Fleuret - Consultant, Regional Office East Africa

Mr. Varley - head of vehicles and procurement.

Dr. Armstrong - Program Manager - Agricultural CRSP programs - Kenya

Interviewees for possible positions:

Dr. Anne Fleuret - Nutritional Anthropologist

Mrs. O'nenygi

Eastern Province

Dr. Oyoo - Provincial Medical Officer

Mrs. R. Gathu - Provincial Community Nutrition Worker

Embu District

Health Officer - Dr. Muriithi

Clinical Officer - Mr. Nabwea

Public Health Technician (Sanitarian) - Mr. Mbungu

Community Nutrition Worker - A. Namu

Chief - Ishera Sublocation

Projected activities for Drs. Cattle and Carter in conjunction with Kenyan group.

- i. Finalization of protocol
- ii. Further exploration of proposed research sites and on site work and living arrangements.
- iii. Further planning for preliminary survey.
- iv. Continuing collection of background material
- v. Site visits to several ongoing nutrition research field sites.
- vi. Follow-up on initial arrangements.

RECEIVED MAR 7 9 1982

Other UC Team Members to assist with the above in Kenya:

1. Dr. Alfred Neumann
March 21 - April 6, 1982
Finalization of fiscal, managerial, administrative aspects of CRSP negotiations with the University of Nairobi.
Assistance in research site selection and relevant living and work site arrangements.
2. Dr. Fred Zerfas
April 8 - June 15, 1982
Carry out and direct preliminary survey.
Sampling, selection, training of field staff and carrying out of survey.
Data analysis and writeup of findings for June SCB Meeting at UC Berkeley for project evaluation and final design.

Recruitment

Recruitment is now going on for the Field Coordinator position.

Candidates to date

Richard Brown, M.D., M.P.H. - Extensive experience in Cameroons and Zaire and extensive field nutrition experience.

Drs. Patrick Ann and Patrick Fleuret - Both Nutritional Anthropologists have had extensive field experience in Kenya and Tanzania. Dr. Patrick Fleuret also is now doing a short term project for USAID in Kenya. They have indicated an interest in sharing the position although it would be preferable for Dr. P. Fleuret because of his administrative experience to be the full-time coordinator.

QUARTERLY REPORT 12/1/82-2/28/82
UNIVERSITY OF CONNECTICUT/INN, MEXICO

RECEIVED APR - 3 1982

PROJECT STATUS

Progress on the project was restricted by the fact that the University of Connecticut (UConn) is still in the process of signing a contract with Berkeley (UCB) (see below). Nonetheless, there have been significant advances in initiating the project.

Travel

1/29 - 2/3 L. Allen, G. Pelto visited INN, Mexico City

Purpose: To discuss subgrant between UConn and INN
To visit potential field-site
To plan preliminary survey of field-site

Results: Mutual agreement on major issues with subcontract
Decision that INN must have advanced funds on hand throughout the project; this must be negotiated between UConn and UCB.
Decision that the potential field-site appeared suitable, and should be further investigated.

2/4 - 2/18 L. Allen, remained at INN, Mexico City, (sabbatical leave)

Purpose: To write INN-UConn subgrant
To visit potential field-site for one week
To plan preliminary survey

Results: Draft of INN UConn subcontract written and tentatively approved by director of INN (Dr. Carlos Gual).
Selection of research site; we have tentatively chosen the Valley of Solis, State of Mexico, approximately 2.5 hours drive Northeast of Mexico City. 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) The 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. When funds are accessible, we will conduct the preliminary survey on this community.

NEGOTIATIONS WITH MEXICO

These include:

- Approval of the project by Dr. Carlos Gual, the director of INN
- Agreement on a detailed schedule of research plans up to March, 1983
- Tentative approval of INN-UConn subgrant
- Approval by INN Human Use Committee

RECRUITMENT OF STAFF

AT INN, we identified a suitable field director (Alfonso Mata, M.D., M.P.H.) and a pediatrician who may work in the field clinic. Other staff are available at INN, but await funding.

At UConn, lack of funds has precluded recruitments or hiring.

CRSP INVOLVEMENT

1/4 - 1/6 P.J. Peltó attended Technical Advisory Group (TAG) on Statistics in Boston

2/4 - 2/5 L.H. Allen planned and chaired TAG on Activity at Purdue University

SUBGRANT NEGOTIATIONS

UConn-Berkeley; UConn received a copy of the subgrant from UCB on 1/8/82. UConn suggested a number of subgrant revisions, of both a scientific and an administrative nature. Scientific revisions were tentatively approved by G. Beaton. Signing by UConn has been delayed because of our insistence on advancement of funds to Mexico. (As of 3/17/82, this problem appears to have been resolved).

UConn - Mexico; see "negotiations with Mexico"

UConn - US Institutions; negotiations are in progress

Second Quarterly Report - Egyptian Project
December 1981 - March 1982

Submitted by Dr. J.C. Wolgemuth

RECEIVED MAY - 3 1982

Although the contract between Purdue and the ME (Berkeley) has not yet been finalized, each scientist from the three universities (Purdue, Arizona and Kansas) and from Egypt has been active in the project. Their input demonstrates their eagerness and dedication to furthering the objectives of this research. It is anticipated that all contracts will be finalized shortly and that actual field work will begin in Egypt in the near future.

A. Travel

In January, Dr. A. Ismail visited Egypt. Although his purpose was for professional activities other than those of this project, he was able to meet with Dr. Hekmat Aly and Dr. Abdel Kader of the Nutrition Institute in Cairo, with Dr. Gabr (MOH) and with Dr. Ikrahim Kandil, Dean of the College of Physical Education in Helwan University in order to discuss arrangements for the initiation of the project. Dr. Ismail reported that the Egyptian scientists are eager to participate and to begin the project field research.

Dr. T. Wachs traveled to Austin, Texas, for the TAG meeting on cognitive assessment. Dr. D.R. Smith traveled to the Universities of Arizona and Kansas to discuss contract details.

B. Negotiations of Institutions

At Purdue, the participants reviewed ME/Purdue contract and offered suggestions for improvement. Items to be included in the Purdue/Egypt contract were discussed by Purdue participants.

Dr. D.R. Smith visited Arizona and Kansas to clarify those contract specifications which will be negotiated with each institution.

Dr. Watson continued work on the selection and confirmation of the Egyptian PI. At this point Dr. Hekmat Aly will serve as the acting PI and will participate in the TAG and SCB meetings in New Orleans in April. Dr. Watson assisted in making arrangements for her travel to Egypt. Dr. Watson participated in the TAG on work capacity in setting it up and assisting in its proceeding forward.

C. Recruitment of Staff

Dr. Jerome corresponded with and interviewed 5 candidates for the position of Research Assistant at the University of Kansas. Ms. Judith A. Ricci from the University of Pennsylvania, Anthropology, was selected and began work in January of 1982.

Dr. Kirksey received an application for her post-doctorate position to assist her in methodology refinement for Phase II of the project.

Dr. June Wolgemuth was hired in February by Purdue University to later assume the position of Project Director/Research Coordinator in Egypt. Dr. Ron Watson, in particular, as well as Kirksey, Wachs, and Ismail assisted in individual orientation sessions with Dr. Wolgemuth to familiarize her with their research activities and project involvement.

D. Methodology Development

Drs. Harrison and Ritenbaugh and their associates at Arizona have developed and drafted documents for food intake methods. They have also developed a field manual for anthropometric standardization. A tape from CDC containing the anthropometric reference data to be used during Phase I was acquired. The above documents and information on anthropometric equipment have been or will be shared with Purdue, Kansas and Egypt.

Dr. Jerome and Ms. Ricci have identified a variety of instruments which may be used to measure social assessment. They continue to search and evaluate measures of social competence.

Dr. Wachs has developed and submitted to TAG participants preliminary cognitive development plans and assessment procedures. Replies from TAG members were coordinated in preparation for the March TAG meeting.

Dr. Kirksey collected laboratory data, recent papers, schedules, and questionnaires for use in Phase II in reproductive competence measurements. Topics include milk volume, energy content of milk, and milk intake of infants.

Dr. Watson continued planning on disease and immunological measures. Dr. Ismail is coordinating project activities on worker capacity in Egypt.

The Egypt project scientists should be commended on the degree of cooperation which has taken place between institutions. They continue to share knowledge, skills and publications in a manner which will contribute greatly to the successful completion of this project.

E. TAG Activities

The worker productivity TAG was held at Purdue University on March 5 and 6, 1982, hosted by Dr. A. Ismail, and chaired by Dr. Lindsey Allen. Drs. Watson, Kirksey, and Wolgemuth attended (silent observers) with the purpose of gaining a greater understanding of the nature and direction of this work.

Dr. Wachs chaired the TAG session on Cognitive development in Austin, Texas. He is preparing a report of their meeting to be submitted shortly.

Dr. Ruth Hassanein participated in the January meeting of the statistical design TAG. The report of this meeting has been circulated.

Dr. Harrison has accepted membership in the new TAG on nutritional status. The first meeting of this TAG will be held at a later date.

SECOND QUARTERLY REPORT
Nutrition - CRSP
12/1/81 - 3/1/82

University of Arizona

A. Country Project Status

1. Travel - none
2. Negotiations with counterpart institutions - none
(Purdue is handling this)
3. Recruitment of staff - none
4. Placement of staff in field - none
5. Methodology Development

The Arizona group has developed draft documents for food intake methods which have been shared with Purdue, Kansas and Egypt. We have similarly disseminated information on equipment for anthropometry. We have developed a draft of a field manual for anthropometry standardization, which will be sent to the other institutions in early March.

B. CRSP Involvement

1. Participation in TAG's

Harrison has accepted membership in the new TAG on nutritional status. Neither this nor the other TAG in which we are involved (Reproduction) has met.

2. Interproject activities

The materials developed on food intake and equipment (A5 above) have been shared with all country projects and the ME. In addition, we have generated and made available for field use tables of anthropometric values for Phase I based on criteria agreed upon by the SCB in October 1981. All of the above documents are part of "Phase I and Preliminary Survey Summary" sent to all projects in February by the ME, for reference in conducting Phase I fieldwork.

The field manual for anthropometry standardization (see A5) will be made available to all projects and sent to TAG on nutritional status for review and revision.

We have acquired from CDC the data tapes containing the anthropometric reference data agreed upon by the SCB for Phase I reference. We have been sent tapes by UCLA and ME for transfer of data and have obtained permission from CDC to copy the tapes for all CRSP projects.

3. Subgrant negotiations

Planned initiation of Phase I has been delayed because of the delay in these negotiations. Dr. Richard Smith of Purdue University visited our campus on January 29, at which time we discussed in detail the subcontract proposed between ME and Purdue. Since that time we have heard nothing about the process. Planned travel to Egypt has been delayed.

C. Others

We remain optimistic about the project, but are concerned about the effects of the delays in contracting. No Phase I work can begin until the contract between Purdue and the Ministry of Health in Egypt is signed, as we understand it. Unless this is completed quickly, it will be impossible to have any useable data by the June SCB meeting. Currently we plan that Harrison will spend approximately May 15 - June 15 in Egypt, coordinating with Jerome (late April to late May) and Wolgemuth.

SECOND QUARTERLY REPORT
Nutrition - CRSP
12/1/81 - 3/1/82

University of Kansas

A. Country Project Status

1. Travel - none
2. Negotiations with counterpart institutions:

Hold meetings with Dr. Richard Smith, Assoc. Director, International Education and Research, Purdue University at his request to discuss the contract issued by the University of California's (Berkeley) contract office.

3. Recruitment of staff:

Corresponded with a number of people interested in the position, Research Assistant. Interviewed 5 candidates and succeeded in recruiting Ms. Judith Ann Ricci, at the time, a graduate student at the University of Pennsylvania's Dept. of Anthropology. Ms. Ricci assumed the position as R.A. in the project on January 18, 1982.

4. Placement of staff in field - none
5. Methodology Development:

Social Competence TAG - We have embarked on a comprehensive search for measures of social competence and for specialists in that field. Obtained a variety of instruments used to measure "social assessment" for detailed discussions at the TAG meeting in April. Also shared concepts and procedures with Ted Wachs (Cognitive Development TAG) to ensure that both TAGs proceed along similar lines and remain faithful to the study design.

B. CRSP Involvement

1. Participation in TAGs:

Ruth Hassanein participated in the January meeting of the Nutrition CRSP TAG on Statistical Design. The TAG report has been circulated.

2. Interproject activities:

Shared equipment information (for measures of food intake and anthropometry) with other PIs for possible use in all country projects.

3. Subgrant negotiations:

Negotiations have been delayed due to some hold up in the grant negotiations between Purdue and California, as I understand it. This caused a serious setback in field research activities. Dr. Jerome's planned trip to Egypt in January - February for site selection and preliminary fieldwork had to be postponed indefinitely due to the protracted negotiations.

C. Others

We have very serious concerns about the protracted contract negotiations. Phase I work has been held up, thus putting us off schedule. If the contract problems are not solved soon we may have to consider some alternate arrangements.

REPORT OF THE
TECHNICAL ADVISORY GROUP ON STATISTICAL DESIGN
AND DATA MANAGEMENT

Jan. 4, 5, 6, 1982
Harvard School of Public Health

Members:

J. Ware (Chairman)
W. Rand (Consultant)
W. Reinke (Consultant)
A. Coulson (Kenya Project)
R. Hassanein (Egypt Project)
B. Peltó (Egypt Project)
G. Beaton (Interim Program Director)

I. Preamble	1
II. Study Design Considerations	2
A. Selection of Community	2
A.1 Community size	2
A.2 Homogeneity	3
A.3 Stability	3
B. Selection of Households	
B.1 Definition of Households	3
B.2 Census	6
C. Type, Frequency and Duration of Data Collection	6A
D. Other TAG Input to Study Design	7
E. Special Studies	7
F. Review of Design	8
III. Control of Supplementary Data Collection	8
IV. Mechanisms for Modification of Approved Study Design	9
V. Data Management Considerations	11
A. Some Principles to be Observed	11
B. Data Transfer Requirements and Schedule	12
C. Idealized Data Management Systems	13
C.1 The Normal Situation - Field Coded Data	13
C.2 Special Data Analyses in U.S. and Third Country Institutions	13
C.3 Special Data Analyses in Country Institutions	13
D. Nutrient Intake Data	15
E. Precoded Reporting Forms	15
F. Data Management Committee	15
G. Implications for Nutrition CRSP Core Staffing	19
VI. The Timetable of Preparatory Work	20
A. Development and Testing of Instruments	21
B. Preliminary Field Study	21
C. Data Management Capability	22
D. Dietary Data Management Capability	23
VII. Other Organizational Considerations	23
A. Inter- and Intra-Project Communication and Collaboration	23
B. Publications and Shared Data Use	24
C. Household Socio Economic Characteristics	26

ANNEXES

A.	An Example of the Operation of Data Management at the Field Level	28
B.	Some Types of Considerations in Data Collection Forms	32
C.	A Check List of Questions for TAG	33

I. Preamble

The CRSP TAG on Statistical Design and Data Management met in Boston, Massachusetts on January 4 - 6, 1982 to review and advise upon the following general areas.

- i) The Nutrition CRSP study research design
- ii) The required data management system
- iii) Procedures for refinement of study design
- iv) Plans for development of the recommended data management system
- v) Organizational matters related to design and data management
- vi) Implementation recommendations

The background documents of the consultation were: the research design section of the Nutrition CRSP proposal to AID (the grant application) and an administrative summary of this by the Interim Program Director. Some members had reviewed also the Final Report of the Planning Contract from which the research proposal had been drawn. The presence of members of the country project teams permitted the introduction of additional background knowledge relevant to individual projects and encouraged a true collaborative approach in this important element of the Nutrition CRSP implementation.

The administrative structure of the management and operation of the CRSP was presented as background.

It was made clear that the role of this and other TAG's was advisory to the Scientific Coordination Board (SCB) and the Management Entity (ME) who jointly must take the final decisions on design and operation of the Nutrition CRSP.

For working purposes, two deadlines were accepted as follows: mid June, 1982, for "finalization" of the design by SCB and ME; September, 1982, for initiation of data collection. It was explained and accepted that country-projects would be initiating preliminary data collection, which would aid in design decisions, in January or February, 1982. The role of the present TAG was not to take decisions but to facilitate the implementation of the Nutrition CRSP by identifying decisions that would have to be taken and, very importantly, by identifying the background information that would be needed in taking those decisions.

The TAG on Statistical Design and Data Management enthusiastically supports the concept of intensive collaboration among the three CRSP country projects to develop highly comparable protocols instruments and other study procedures to maximize commonalities in the three investigations. The currently proposed study design is seen as representing an excellent beginning. Conversely, the adequacy of that study design for meeting different research goals must be carefully reviewed by other TAGs in the next few months. There is an obvious need for rapid development and field testing of standardized instruments, coding procedures and other data collection and management plans.

The present report discusses some of these issues and offers a number of suggestions and recommendations concerning planning activities and study procedures. Among these is a recommendation that the study design be reconsidered and its effectiveness reevaluated after TAG recommendations on instruments and substudy design requirements are available. The present TAG emphasizes that this is not in any sense an expression of doubt about the interim design or

invitation for fundamental change. Rather, it is a recognition of the need for operation of the normal process of refinement of design as additional information becomes available.

Some of the major issues and recommendations addressed in this report are:

1. Project feasibility, especially regarding data collection and management (need for, and characteristics of, a data management system).
2. The major responsibility of the subsequent TAG; to review the study design and to develop specific and concrete hypotheses that can be satisfactorily addressed through the final design.
3. The need for continuing review, and revision as necessary, of the study timetable recognizing the considerable volume of work to be accomplished before initiation of Phase 2.
4. The need for, and potential problems of, careful coordination and harmonization of this multi-institution, multi-country project.

II. Study Design

The goal of the CRSP Project is the completion of three highly comparable investigations in three different country settings. This will require standardization of study design, data collection instruments, and coding procedures across projects to the extent that is compatible with local circumstances. The standardization will be achieved only through close and frequent collaboration among the investigators both before and after field testing of study procedures.

A. Selection of Communities

The project investigators will be selecting study communities in the next few months. Basic criteria for selection are thoughtfully spelled out in the proposal and the Statistical Design TAG concurs in their identified importance. The significance of stability, appropriate range of nutrient intake, and adequate infrastructure is especially underscored. The community must be stable if the study is to avoid problems of attrition of households, thereby reducing sample size below tolerable levels and, more importantly the validity of the study. The range of nutrient intake must be sufficient to test the hypothesized association between intake and the study outcome variables. Adequate infrastructure is necessary to ensure feasibility of the study.

Accepting the basic thrust of the proposed criteria for community selection, the following paragraphs elaborate upon certain features and implications of the selection process that the TAG considers to be especially critical.

A.1 Community Size

In order to ensure necessary flexibility in establishing appropriate criteria for exclusion of households from the sample, the community must contain substantially more than the 200 households to be studied. Moreover, a large community will permit adequate sample representation of relatively uncommon, but important characteristics. For example, a characteristic present in 20% of community households but to be included in 10% of sample households for reasons of efficiency and sensitivity indicates the need for a community of at least 400 households in total. Considering the likelihood of refusals to participate and modest attrition a community of more than 400 households and perhaps as many as 600 households is not an unreasonable requirement. An important determinant of size will be the observed distribution of variables.

These considerations suggest that sampling from a single village may not be feasible. Rather, the "community" of interest may be a region of comparable settlements. Particular villages selected from the region should together exhibit the same characteristics of stability and variability noted earlier. In addition, they should not be distinctly different in relevant nominal characteristics, such as tribal affiliation or ethnic composition, environmental conditions (e.g., source of water), occupational distribution, or religion.

A.2 Homogeneity

The TAG concluded that homogeneity is not a requirement per se, but that important differences in community composition should be well understood, quantifiable, and weakly correlated with food intake if they are to be controlled in data analysis. The TAG recommends that semi-structured community profiles be constructed, incorporating uniformly accepted dimensions to permit comparability across projects. Community factors represent an important element in the interpretation of project findings, along with individual and household variables. It is recognized, of course, that community assessment will be necessarily more descriptive than analytical.

A.3 Stability

The following items would contra-indicate selection of a community or region.

i) Large and Relatively Enduring Out-migration of persons to employment in areas beyond reach of the project, continuing permanent out-migration concentrated in certain categories, e.g., adult males, is a special concern. Temporary out-migration on a seasonal or other basis of periodicity is likewise highly undesirable. A modest amount of temporary out-migration randomly distributed in time among population members could be tolerated. It will be important to collect information for the preliminary study about the expected rates and durations of individual out-migrations.

ii) Rapid Changes in the Economic Base of the community through industrial development, road construction, energy development, etc.

iii) Political Instability in the region.

iv) Prospects of crop failures or other causes of serious food shortages.

B. Selection of Households

The proposal highlights the important principle that household sampling should result in over-representation of the extremes of nutrition status (short of severe clinical malnutrition). Reproductively active females and pre-school children should also be adequately represented. Criteria for achieving these goals have not been specified in detail, and they may differ somewhat among projects. The TAG recommends that projects develop operational criteria for household selection in the course of conducting field work leading to community selection. The criteria should be presented for SCB approval, along with information depicting expected distributions of criterion variables. This documentation will permit judgements to be made concerning comparability between projects and will enhance confidence that unintended distortions will not be present in the findings.

The TAG notes that the planned design calls for a two stage stratification in household sampling. The first level of stratification relates to nutritional status as a proxy for the independent variable, food intake. This is discussed above and this approach is endorsed as an important feature of the proposed design. The second level of stratification was to be based on

measures of socio-economic status. While the members of the TAG appreciated the intent of this plan they questioned the feasibility in practice - feasibility of developing a scaling system that would have interproject meaning before the scheduled time of implementation, and feasibility of collecting data required for such a scaling system before sampling. It was recommended that this aspect of household selection be reconsidered.

The TAG recognized that at the time of data analysis there might be need for description of the relative position of households on some type of socio-economic scale. Such data might be gathered once during the two-year period as a "fixed" description of the household. This possibility is discussed more fully in section VII C.

B.1 Definition of Households

In the design the key unit of observation is individuals but within the context of households. The key unit of sampling is households. The TAG considered this carefully and notes the absolute necessity for each country project to develop an operational definition of households to be used in sampling and in subsequent data analyses. This must be done during the preliminary phase. It must be confirmed by the SCB. The TAG draws attention to the fact that there are both analytical implications and logistical (and cost) implications to these definitions. It does not prejudge decisions of SCB but wished to ensure that the implications are recognized, relevant information is collected during the preliminary census, and informed decisions are then taken.

"Household" is generally understood to be:

1. Group of people sharing a common residence "under the same roof."
2. Sharing the same food, which is organized and prepared in common. "They eat from the same pot."
3. Usually the people are genetically and socially related, with a mated pair (or more than one mated pair) as "household heads", plus their offspring.
4. A range of additional persons, usually related to one or another household head (the "spouse") are also present, especially including parents of either spouse, and siblings of either spouse.
5. Occasionally persons totally unrelated (e.g. servants, hired workers or "renters" may be included as household members.
6. The household is fundamentally an economically cooperating unit, pooling their resources of housing, food, and other materials and labor.

Certainly there will be some difficulties in clearly identifying households in some locations. The following guideline is suggested

where possible, households should be defined in terms of individual mated pairs.

Thus, in cases where multiple mated pairs are in close association, it should still be possible to identify the individual mated pairs plus dependents.

(Presence of other mated pairs in close association with the index "household" must be noted as an important "social relations" or "Kinships network" variable.)

If "households" are defined on this basis, then consideration must be given to the question of data requirements concerning other individuals who share facilities (kitchen, dwelling, etc. but have a different relationship). It may be that different research questions will demand different operational definitions of "households" - but may not require the same intensity of data collection for all individuals. The TAG's as well as the SCB will need to consider this; there will have to be a conceptual consistency across country projects. The definitions can be built only with knowledge of both the population structure and the research questions. The investigators (and TAGs) must consider the significance of the occasionally absent members of households, and consider what kinds of data are needed concerning the qualities and activities of those persons during their absence (and whether such information can be effectively collected within logistical constraints).

The TAG became aware of a major logistical issue that may not have been adequately considered in earlier consideration of design. It had been assumed, in documents placed before the TAG, that core data would be collected on all individuals in the household, with specifics of the nature and frequency varying with age, sex and indicator events. It had been assumed that household size would be 5-6 and that 200 households would generate a population of 1000 - 12000 individuals.

If instead, the operational definition of a "household" comprises 7-8 individuals (e.g. grandparents or spousal kin) the population would rise to 1400-1600 with major cost and logistical implications for data collection and analysis. Upon detailed consideration this may be beyond the resources of the projects and some compromise of design may be necessary.

The present TAG suggests that consideration be given to a final design in which:

- 1) core data on dependent variables need be gathered only in certain key individuals in the household.
- 2) some data concerning food intakes, etc. may be needed in all household members in order to define their roles as contributing to the household environment affecting the index individuals.

The implications of such a design approach (and the identification of the variables to be measured with differential frequency or with qualitative differences in information) can be addressed only in terms of consideration and precise definition of the research questions - a matter to be addressed by the TAGs and SCB.

B.2 Census

In order to identify the appropriate households for recruitment and follow-up in the study, it would be desirable, if not essential, to define the community as an entity having geographic boundaries and to complete a census (and map) of the households in the defined community. Data from the census can serve not only to define the sampling frame but also to relate the selected households to improve generalizability.

The census should derive a roster of all members of each defined household, including age, sex, relationship to others in the household, reported health in the last six months, length of residence in the community, "susceptibility" to temporary out-migration, general occupation, individual anthropometry, and any other easily determined (or desirable surrogate) data useful in characterizing the members of the household in relationship to final study design. In addition to geographic location of the household, information should be obtained about locational relationship to other households (and if possible to kinship households) in the community, basic source of food and economic support, types of foods consumed in the household, quantity of foods used in the household per unit time, and, if possible, ethnicity and religious characteristics. As deemed appropriate in terms of final study design, if there are observable markers of economic or social status in the culture of the area, query could be made in the census questionnaire or structured observations could be made by the study staff members conducting the census.

Data instruments completed during the census should be abstracted to characterize households in a manner of relevance to the study and planned analyses. The abstracts can be divided into strata and the strata sampled heavily or lightly according to the study needs. Arbitrary identifying numbers assigned to households can be used in randomized sampling within the strata.

Census data should be edited, coded and reviewed. By entering census data, and outcomes of recruitment attempts, in the data bank, it will be possible to generate descriptions of the total community, of sampling ratios, of nonparticipation bias and other study population descriptors. The household data may find use also as a base for selection of replacement households (refusal of participation or loss during follow-up).

Census data forms and data should be distributed to country and US country project data centers. However, data leaving the country should be coded in a manner that prevents identification of individuals but permits linking of data.

C. Type, Frequency, and Duration of Data Collection

The categories of variables and general plan for frequency of measurement are defined in the current protocol and are accepted as a reasonable starting point for further detailing of the study design and development of instruments. As variables are more clearly defined and refined in the course of instrument development the frequency of measurement needs to be re-appraised from a number of perspectives.

Repeated measurement of a particular variable relative to an individual is for one of four purposes. First repetition may be to account for variation in

individuals. This purpose is exemplified by the repeated measurement of food intake to determine an individual's average consumption during a given period (which might be a season or a phase of an economic cycle, etc.). The question here is how many measurements are necessary to provide satisfactory levels of precision. The second purpose of repetition is to establish patterns of interest. For example, patterns of physical growth can only be derived from periodically administered anthropometric measures. Where this purpose is operative, consideration should be given to techniques for aggregation in the form of indices. The creation of indices will simplify subsequent multivariate analyses and will overcome the inevitable problem of missing observations. The third purpose of repetition is to permit time-phased analysis of interactions. Morbidity episodes may be associated with concurrent or subsequent increments of growth, and growth patterns may be associated with later morbidity. The fourth purpose of repeated measures is to ascertain seasonal or age-dependent effects. In this case, frequency of measurement depends upon the periodicity of the event of interest. A related consideration is the need to obtain measurements at a specified age. If a particular variable is applicable to children reaching their second birthday, for example, provision must be made for identifying the individual and making the measurement appropriately, regardless of when this time falls chronologically.

Although the time frame and repetitive nature of data collection make the study appear to be longitudinal in nature, it should be emphasized that detection of change over time is not a major purpose of the study. Rather the timing of observations is mainly for the purposes listed above. Admitting that the proposed two-year time frame is generally inadequate for recording change, it may also be too short to observe certain effects that are intended for analysis in the studies. That is, for example, it may require more than two years for an effect during early development to be expressed during later development. Other TAGs should evaluate this possibility and make appropriate recommendations for extending data collection or for limiting study objectives where extension is not feasible.

D. Other TAG Input to Study Design

Much of the needed refinement in study design can only emerge from other TAG efforts in instrument construction. A checklist of study design issues to be addressed in these TAG deliberations is presented below.

1. What are the relevant research questions and hypotheses to be tested?
2. What data are needed with what frequency for what period of time relative to what target population in order to answer these questions? How large a population of subjects will be needed to answer the question?
3. What specific measurements are to be made and how?
4. What instruments are to be used for capturing the needed data?
5. Who is to be responsible for collecting the data?
6. What are the procedures for coding and processing the data?
7. What are the staffing implications and respondent burdens imposed by the data collection?
8. What tabulations and analytical methods are to be employed in relating the data to the research questions?

In addressing 1-3 above, it is recommended that each TAG test it's answers by asking "How will these data be used to answer that question?". A practical test of considerable merit will be to draw up mock tables. As noted earlier, the TAG's must consider approaches to generating indexes that would represent consolidated data in analyses. This is the implication of 8.

Of necessity this approach will lead TAG groups to consider data and design requirements for relatively narrowly defined questions (the relationship of the food intake variable to single outcome variables). The final design, to be decided by SCB, will have to consider the requirements for these individual questions (that is, for example, incorporate data collection to meet the most demanding question - or decide to modify or abandon the question). Importantly, it will fall on SCB to consider the data requirements for interaction questions (TAG's have been constructed to generate expert advise on single outcome functions). The TAG's should consider the interaction requirements for design but should place priority on giving description of design requirements for at least the direct research questions.

E. Special Studies

The preceding checklist and discussion suggests that the same body of data may not be required uniformly with the same frequency over the same period of time from all individuals in the sample. For example, details of food intake for lactating mothers, weaning age children, and adult males may differ somewhat. Some peripheral members of the household (e.g., adult

males working in the city) may be of no concern for certain purposes except as their behavior impacts on the specific target population of principal concern. Such nuances need to be clearly elucidated.

Certain special studies may require a sample different from that identified for core purposes. For example, a project undertaking a special study on the aged may obtain some respondents from the core sample and supplement these from households identified in the census but excluded from the core sample because of the absence of reproductively active females.

F. Review of Design

The TAG considers that it may be of value to reconvene this or another group of advisors to reconsider the statistical issues and design suitability after the TAG recommendations are available (on instruments of data collection and requirements for narrowly defined research questions). This meeting might serve to focus the issues to be addressed by SCB and thereby facilitate the final design decisions.

Given the importance of design considerations (e.g. required frequency of observation, required sample sizes, etc.) by the TAG's, it is recommended that each TAG include at least one person with statistical expertise. Such an individual might be drawn from individual projects or might be drawn from the membership of the present TAG. It is recognized that such statistical advice should be provided by persons who are aware of the general approach of the Nutrition CRSP.

III. Control of Supplementary Data Collection

The focus of attention of the present TAG is the research design and data collection and analysis procedures for the core research. It is recognized that the Nutrition CRSP makes provision for supplemental studies within country projects, either funded through the Nutrition CRSP or by independent means, provided that such additional data collection does not compromise or interfere with the core design.

It would appear that mechanisms exist to control the collection of core and supplementary data insofar as budgetary implications are concerned. All institutional agreements carry the stipulated requirement for the collection and transfer of core data.

The TAG did not receive satisfactory assurances that mechanisms now exist for the control of supplementary studies from the standpoint of their potential impact upon the participating subjects of the core research design. That is, it is recognized that the core design in itself represents a major imposition and intrusion in the cooperating families. There already is a major need for field efforts to maintain this cooperation and involvement in the longitudinal design. Further intrusions, and particularly intrusions that are undesirable to the participants (e.g. blood letting, long time involvement in interviews or increased interference in patterns of daily life) may jeopardize the long term cooperation.

Any compromise of the core research is a major concern not only for investigators in the particular country project but also for all collaborators in the multi-country nutrition CRSP. Therefore, it is a concern of the SCB and the ME acting on behalf of the total program and the individual investigators.

It is recommended that, regardless of funding source, all supplementary studies which will involve core research subjects be presented to the SCB and ME for review and approval before implementation. The submission, and the review, should direct major attention to the additional impact the proposed supplemental study would have on participating subjects.

Under the terms of the grant and subagreements, all such studies, as well as the core research, must have the approval of human ethics committees. The SCB/ME review and approval would not replace or substitute for human ethics review nor should human ethics approval substitute for independent approval by SCB and ME.

To facilitate the above, a consideration might be given to providing immediate review by the Program Director and advisors as needed with authority to approve the study if deemed to have very limited, and non-detrimental impact. SCB should be informed of such action. If initial approval is not granted, the country project might withdraw the application or request that it be forwarded to SCB for consideration and recommendation.

IV. Mechanisms for Modification of Study Design

The Nutrition CRSP is a collaborative and coordinated study. To a very major degree its analytical power depends upon the ability to conduct integrated comparative analyses. In turn this ability depends upon the implementation of standardized data collection protocols which will provide comparability of data. Both the instruments and procedures of data collection and the timing of application of these instruments must be standardized.

For certain of the variables, notably food intake and some of the social and cognitive measures, it is recognized already that the actual instruments may be project-specific. Nevertheless, these must be designed and tested to establish data acquisition with comparable interpretation.

The above decisions are to be taken during Phase I of the program. Before data collection for the study design begins, there must be jointly agreed upon and approved detailed protocols. These protocols must be followed throughout the whole of the data collection unless there is an approved modification. Approval should be recommended by SCB only after it is ascertained that the same modification can be made across all projects, or that the modification in the single project will provide an improvement in the uniformity and comparability of data across projects, and assurance that the change will not affect longitudinal continuity for data analysis.

Recommendations for change in the protocols may arise from a number of sources including (see figures in section V):

- 1) field data collection or coding point (the responses obtained to the instrument are uninterpretable or unreliable or do not fit the predetermined coding system).

- ii) first level of project data aggregation and descriptive analysis (observation of unexpected distribution of values suggesting that population characteristics are different from anticipated in planning stage; observation that intraindividual variability (or measurement variability) is different than expected implying possibility of, or need for, modification of measurement frequency; missing data frequency in longitudinal sense, etc.)
- iii) first level of interproject data comparison (observation of unexpected differences between projects in distributions of variables, etc.)
- iv) first level of preliminary data analyses, intra-and inter-project levels (observation that the relationships are different in nature or strength than anticipated implying the desirability of modified duration of data collection, discontinuation of an area of data collection, or other fundamental modification of the design, etc.).

It may be anticipated that SCB will need appropriate documentation of both the problem to be addressed and the implication of the recommended change before recommending approval for a modification of the protocol. The extent of such documentation may be expected to vary with the import of the proposal. For changes which imply alterations across all projects, some supporting data comparisons may be required. The ME should be notified as soon as a proposal is generated and should serve as the coordinator or effector for any required interproject data analysis comparisons and the Program Director should ensure that appropriate steps are taken to provide the SCB with requisite background information.

Recommendations and modification of the preliminary design decisions before implementation of data collection are likely to arise as a result of the deliberations of the specific TAG's (see discussion of Study Design, Section II) or as the result of field testing of instruments. If feasible the reports of the TAG's and the results of preliminary field trials should be reviewed by the TAG on Statistical Design so that it may frame final recommendations for the SCB. It may be necessary for SCB to hold a final meeting shortly before data collection is scheduled to begin if final field testing (after the planned meeting in June) reveals specific problems with the proposed methods and instruments, and at that meeting to take the final collaborative decision on the protocols to be implemented.

The above processes refer to the core data collection of the Nutrition CRSP. Changes in protocols for project-specific research need not come to SCB except with reference to their impact upon the study population as discussed in section III.

In practice, the above processes can be facilitated by delegating SCB review authority to the Program Director with the proviso that he/she be required to consult with SCB before giving approval to a change that is believed to infringe upon interproject data analytical power or to require modification of protocols for other projects, and with the proviso also that a veto by the Program Director should be referred to the SCB on

request by the project. If the Program Director approves a modification SCB should be notified promptly and individual members should respond immediately if they identify unrecognized issues.

V. Data Management Considerations

The TAG noted that, in contrast to research design, the data management systems to be implemented in the Nutrition CRSP were hardly discussed in the background documents. The Interim Program Director provided minimal additional background about decisions that had been taken. It was deemed appropriate and essential that the TAG address this crucial aspect of the research and develop a statement of principles that should hold for whatever system is implemented, to recommend an idealized system that would seem to fit the Nutrition CRSP program, and to offer some examples of the operation of such a system based upon past experience.

A. Some Principles to be Observed

The first principle is that scientific research can be no better than the quality of the data collected and analysed. With a program having the complexity and scope of the Nutrition CRSP, there must be major concern about, and provision for, quality control in data collection and management. The first level of quality control should be at the field.

The second principle is that all data transfers from the first transfer at the field level to the final transfer into the core data base, and all edit queries and responses, must be fully documented - to the extent that all data could be regenerated at any time from these records.

The third, and related, principle is that there must be redundancy of data bases. In the event that one data base is lost, there must be mechanisms for its regeneration. Explicit within this is the concept that there are duplicated hard copy data bases, recognized repositories of primary data derived from this hard copy material, and repositories of secondary or restructured data files that will provide easy entry for analysis.

A fourth principle may seem obvious but has been overlooked in the past, is that all data in this system must be demonstrably identical. That is any edit procedure which modifies data in one base must effect modification of the same datum in other bases. This principle is particularly important when duplication of data bases (e.g. transfer of data for special analyses) is involved.

A fifth principle, discussed in section VII.C, is that at all steps there must be adequate protection of the data and of the interests of participating investigators.

The final principle, which should be obvious in a collaborative project is that the management system must provide for the monitoring of both data collection and transfer of collaborative data (i.e. monitoring and ensuring the effective operation of intracountry project systems) and for ongoing

intercountry data comparisons, (for example, the provision of periodic descriptive data analysis reports to SCB as well as the collection of frequent administrative reports on the status of data files).

Underlying all of this is the explicit assumption that, for the core data (data to be subjected to collaborative analysis across projects). There will be uniformity in i) the hard copy data forms, ii) the coding system, iii) the structure of organized data files. This, in turn requires a prior decision on the actual content of the "core data" and the expected nature of subsequent analyses. Both of these are expected to arise from discussions of the TAGs and of the SCB by June, 1982.

The goal of the data management system is to permit and to facilitate the comprehensive intra- and inter- project analyses that will come at a later stage. The mechanisms of these analyses are not a primary concern of the present consultation. Rather, at this stage it is assumed that provision will be made by the ME for interproject data analyses responsive to the recommendations of the SCB and that individual country projects will make provision for intracountry project analyses.

B. Data Transfer Requirements and Schedule

It is an implicit principle that all data should be made available within the country of origin. That is, although there may be more than one entry point into the data system, and varying levels of data reorganization and aggregation within the system, all primary data files not originating in the country should be copied back to the country. It is a scientific essentiality and contractual requirement that data pass onward to the central repository during the study on an ongoing basis during the study rather than after completion.

As an operating guideline it should be a reasonable expectation that primary data (NCR forms) will be transferred between data centers in the system (see flow diagrams) on a biweekly schedule and computer-entered data tapes on a monthly schedule, that restructured data flow to the core data base on a monthly schedule and that the time lapse between field data collection and entry of edited data into the core data base be no more than 4 months; this assumes that the time required between collection in the field to edited "good" data in the U.S. country project data base is 3 months or less and that final revisions of the core data base are then complete within 4 months. Transfer of data from country to the U.S. will be by hard copy NCR forms (biweekly) and computer tapes (biweekly or monthly). Transfer of data between the U.S. country project data base and the core data base might be by computer tape or by direct linkage of computers.

Operationally, the U.S. data manager and the country data manager must share primary responsibility for quality control of country project data. The role of the ME is to monitor the operation of this quality control system, to offer advisory support as requested and as feasible, and to directly intervene if it becomes apparent that a country project data management and/or quality control system is not operating effectively. Monitoring can be accomplished by review of appropriate administrative reports and by monitoring flow rates of central data bank entry and inter-project data consistency.

The system should provide back-up monitoring of data entry as well as control of data quality per se. Since it is envisaged that hard copy data (NCR forms) will be stored in two different repositories for safety, a simple procedure will be sample reentry of hard copy data at the second-level data base to confirm primary data entry. The sampling rate should be high in the initial phases of data collection but, based upon experience and confidence in reliable operation of the system, it can be reduced in later stages.

C. Idealized Data Management Systems

In the accompanying diagrams, an attempt is made to portray a system embodying the principles described above. It is recognized that this may not be the system finally implemented, and indeed that variants may exist between country projects. However, it is suggested that in considering variants, close attention be paid to the implications for the principles set forth above.

It is recognized by the present TAG that the cost of implementing the data management system portrayed, with its multiple levels of data control and storage will be high. It may be beyond the levels provided for in the existing budget plans. If this be so, there will be obvious need for reconsideration of both the idealized system and the existing budget divisions. Conversely, it must be reiterated that the scientific conclusions of any research can be no better than the quality of the data that enter analysis. It would be highly dangerous to compromise on an effective system of data management including adequate quality control measures.

C.1 The Normal Situation - Field Coded Data

The system portrayed in the first figure assumes primary quality control in the field and primary data entry in the country (some aspects of the operation of such a system, at the field and country level are portrayed, as an example, in Annex A). The diagram gives recognition also to the fact that not all data need be forwarded to all levels of the system, and that parallel data bases may be generated at the level of country projects if deemed necessary to investigators located in different institutions. By the direction of the arrows, it will be apparent that the Nutrition CRSP core data center receives data from three country projects but does not release country project data to other projects (except in summary comparative report form or in the form of interproject analyses).

Not shown on this scheme is the necessary flow of administrative reports describing current status of data banks and of editing processes.

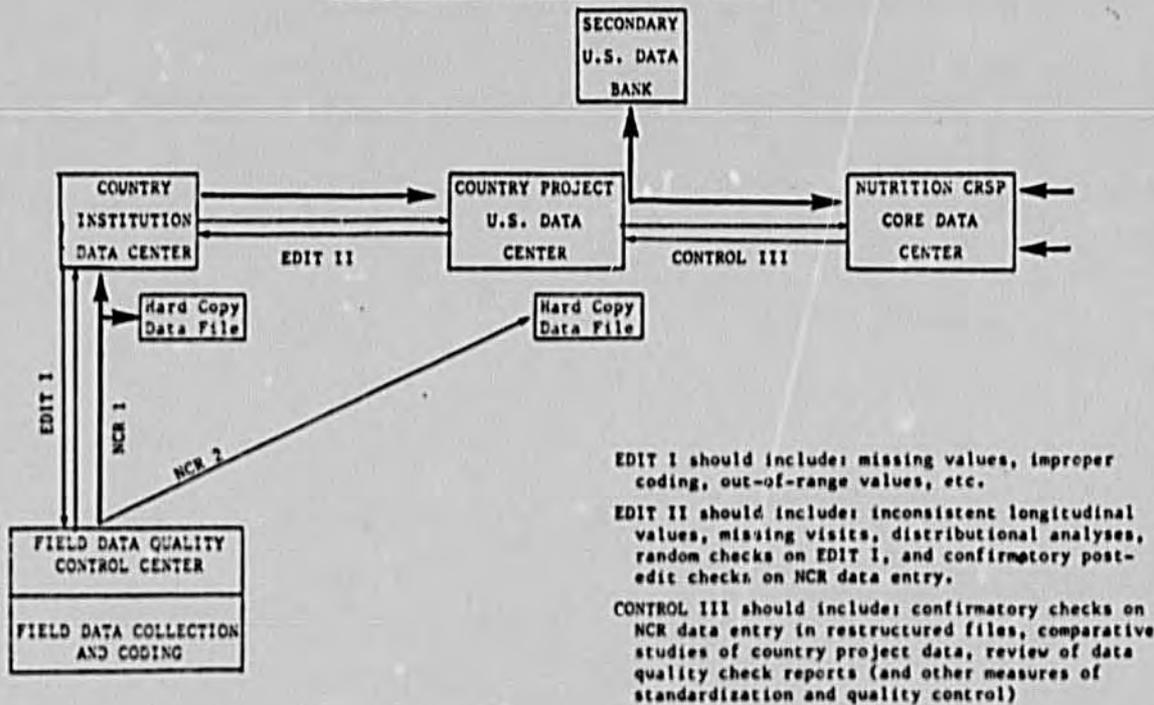
C.2 Special Data Analyses in U.S. or Third Country Institutions

C.3 Special Data Analyses in Country Institutions

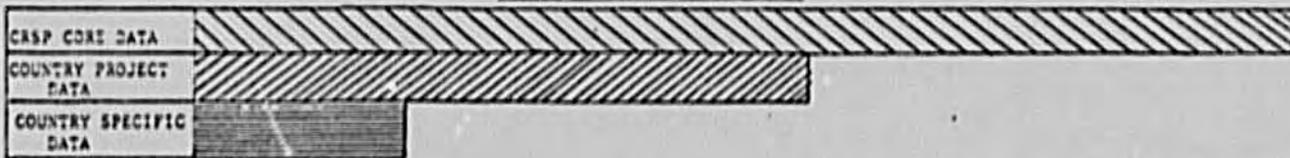
This pair of figures portray the possible situations with regard to data that cannot be field coded. The most obvious example relates to laboratory analyses with samples collected in the field but analysed in another location. A second, much less desirable, situation is that in which field notes are transferred to another location for preliminary analysis before coding is possible.

NUTRITION CRSP DATA MANAGEMENT SYSTEM

A. FIELD-CODED DATA: THE NORMAL SITUATION



MINIMAL CONTENT OF DATA BANKS



DATA MANAGEMENT SYSTEM. The figures portray the flow of field-collected data (heavy lines) and the data quality control and editing links (light lines). The system envisages three data banks, one in the study country, one in the US country project university, and a Nutrition CRSP core data center. Portrayed are pathways for any one country project; similar systems exist for each country project with the Nutrition CRSP center receiving core data from all of these. The schematic portrays also the different levels of data aggregation in these sequential data banks.

Fig. A is seen as the usual pattern of data flow. Fig. B and C portray variants that would hold when data are not coded and checked in the field.

The variation in these models is the logistics of point of entry into the data system, and hence direction of information transfer, and the location of quality control editing responsibility.

D. Nutrient Intake Data

Special mention is made of the management of food intake and derived nutrient intake data not only because this is the independent variable of the core research but also because it requires computerized processing and transformation of data (from food intake data to nutrient intake data). Three models are presented in which this transformation is undertaken (a) in the country institution data center, (b) in the U.S. country project data center, or (c) in the Nutrition CRSP data center (or in a subcontracting institution). To date it has not been decided which of these models will operate. The diagrams are relatively self-explanatory. Only a few comments need be offered.

If either (i) or (ii) are implemented then it seems mandatory that the ME establish a system of sending out mock food intake data records for processing at each project center and report on the comparative analysis of results (to generate data ensuring the comparability of interproject computations).

If (iii) is implemented, then it is mandatory that the original coding of food intake be on an identical basis across all projects. If either (i) or (ii) is implemented, then unique coding systems, compatible with that country project's analysis program would be implemented unless there is a plan and requirement to undertake interproject food intake (in contrast to nutrient intake) analyses. None has been identified so far. Unless such a requirement is adopted, the core data base would include only the desired nutrient intake data.

These matters should be referred to the TAG on food intake for consideration and recommendation.

E. Precoded Reporting Forms

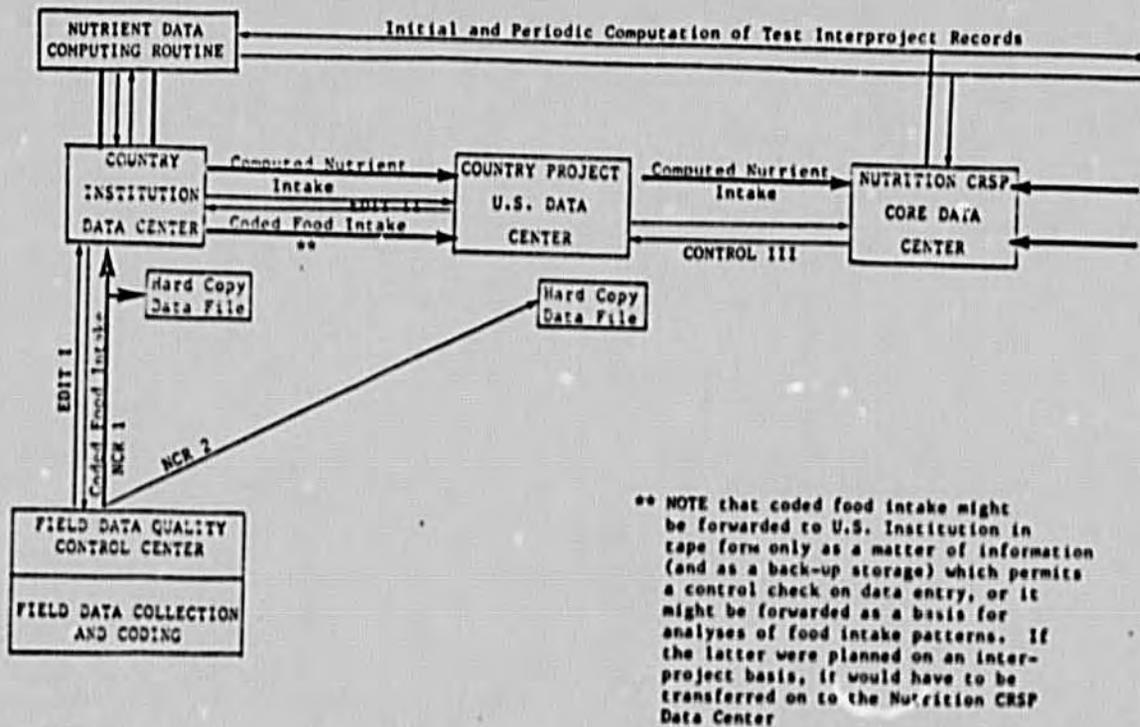
The system described above envisages the use of NCR forms that permit filing of duplicate copies in separate centers and which permit field coding of recorded data. (It also envisages adequate training of personnel in the use of these forms and in the coding of data - see Annex A). For core data it seems mandatory that the forms and coding be standardized across projects with as little project - unique features as possible. Annex B provides a discussion of a possible approach to these forms.

F. Data Management Committee

The preceding discussion illustrates the complexity of the data management system that must be established. The urgency will be recognized when it is realized that all parts of the system must be in place before the first item of data is collected - that is by September, 1982 according to present plans.

NUTRITION CRSP - ALTERNATIVE APPROACHES FOR NUTRIENT INTAKE DATA*

A. NUTRIENT COMPUTATION CARRIED OUT WITHIN THE COUNTRY INSTITUTION DATA CENTER



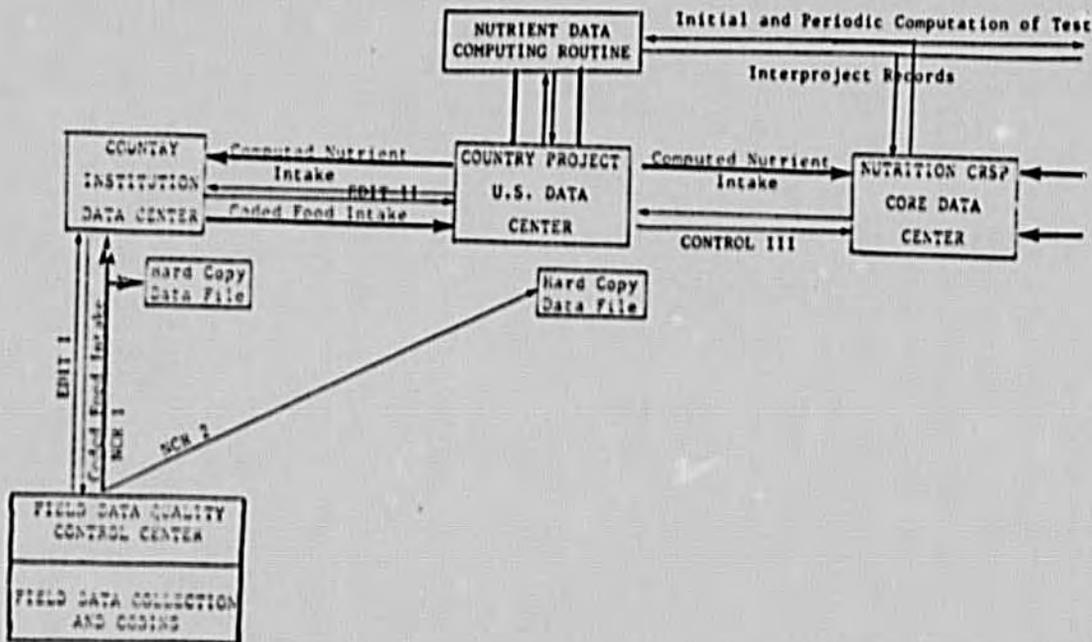
* See "Data Management System" for explanation of abbreviations.

NUTRIENT INTAKE DATA. For explanation, see "DATA MANAGEMENT SYSTEM" These schematics portray the flow of primary (food intake) data and secondary (computed nutrient intake) data and of quality control and edit processes including an interproject quality control nutrient computation check initiated by the Nutrition CRSP core center (to confirm comparability of peripheral food composition data banks and computing processes).

Fig. A, B, C portray different locations of the nutrient computations.

NUTRITION CRSP - ALTERNATIVE APPROACHES FOR NUTRIENT INTAKE DATA*

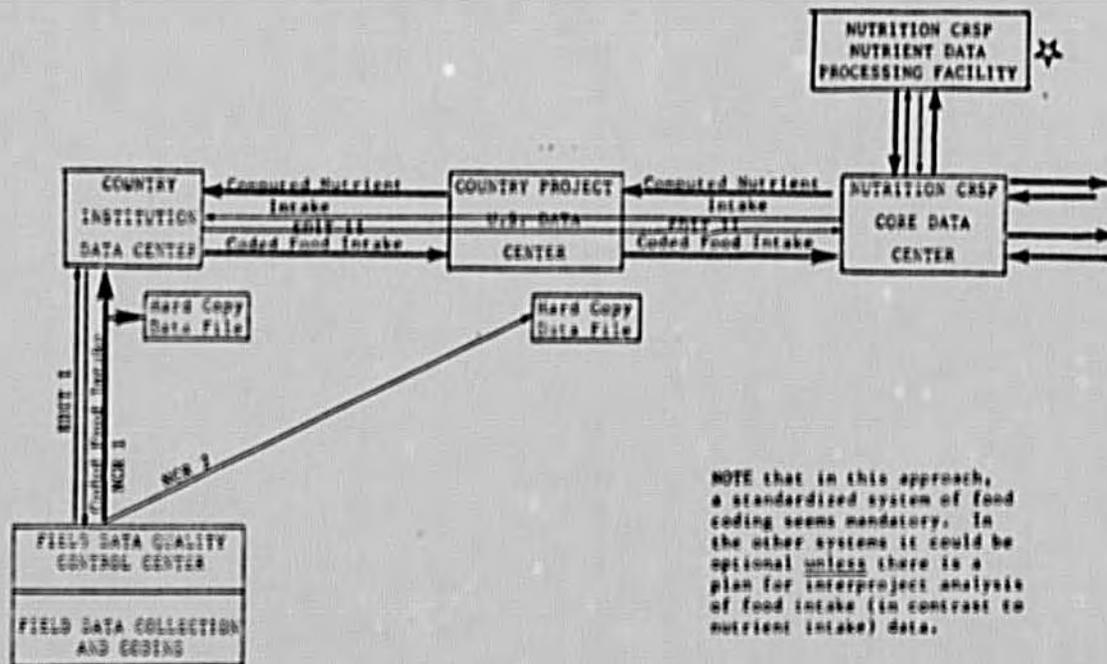
B. NUTRIENT COMPUTATIONS CARRIED OUT WITHIN THE U.S. PROJECT DATA CENTER (OR A SUBCONTRACTING INSTITUTION)



* See "Data Management System" for explanation of abbreviations.

NUTRITION CRSP - ALTERNATIVE APPROACHES FOR NUTRIENT INTAKE DATA*

C. CENTRAL NUTRITION CRSP NUTRIENT COMPUTATION FACILITY



NOTE that in this approach, a standardized system of food coding seems mandatory. In the other systems it could be optional unless there is a plan for interproject analysis of food intake (in contrast to nutrient intake) data.

* See "Data Management System" for explanation of abbreviations.

* This facility might be "in house" as a sub-routine computation or it might represent another institution operating under sub-contract for all country projects.

The present TAG recommends that the SCB and ME consider the general features and principles of the data management system as described in this report. Once this or a modification thereof has been approved, it is recommended that a Data Management Committee be established with delegated authority to design an operational system reporting to the ME, and to the SCB, as required for approval of budgetary and other implications, and for final approval of the system or system elements.

Desirably this committee will be composed of the U.S. based data manager of each country project and the Nutrition CRSP core data manager. The Program Director would serve ex officio. This committee might continue as a collaborative advisory group during the ongoing collaborative operation of the data management system.

Clearly a first task of this committee will be the determination, by intra-project discussion with the involved PI(s) of preexisting plans and budget allocations for data management. The second task will be to test these plans for consistency with the approved principles and required procedures. It will then be appropriate to begin considering necessary modifications of existing plans and budgetary implications.

G. Implications for Nutrition CRSP Core Staffing

There are two distinct aspects to the role that the ME must play in the data management system: a) a consultative and monitoring role in the organization and implementation of the total system from field to core, and b) an operational role in the "hands-on" management of core data forwarded from the country projects. At a later stage the ME will be heavily involved in the statistical analysis of core data.

The TAG recognizes the need for a full time person (or equivalent manhours) experienced with data management, at a somewhat later stage a half-time consultancy of an experienced statistician and, as needed supporting staff to work with data, computer entry, etc. The responsibility would not be to "manage" the country project data managers or statisticians, but to

- i) administer and coordinate those activities which need to be the same across the three projects (coding system, quality control, data flow checks, etc.)
- ii) work with the core data sets (structuring and maintaining files to facilitate analyses, conducting comparative analyses, preparing status reports and test analyses, etc.)
- iii) serve as a consultant and clearing house for data related problems and ideas
- iv) establish and implement quality control checks (e.g. see discussion of nutrient intake computation quality control; equivalent blind samples analyses for laboratory standardization, etc.)
- v) collaborative analysis of core data.

In the unlikely situation that operational problems are encountered in data management or quality control of one of the country projects, ME staff would be expected to first bring this to the attention of the U.S. institution and discuss problem identification and correction. If the problem is not satisfactorily rectified then the ME should be prepared to intervene more directly in the country project data management system for core data. It is expected that the sharing of experiences through the data management committee, with the core data manager serving as focal point or information transfer, will minimize problems and facilitate solutions.

All members of the TAG saw the need for a well experienced person whose primary responsibility was data management. Several members saw the need for one or more additional staff. Reservation was expressed about too great a reliance on data managers drawn from other sectors of the University (e.g., a data management unit) unless there was also a full-time person in the ME with assured access to the data management operations and the priority for handling Nutrition CRSP data.

It was seen as potentially advantageous to defer recruitment of permanent staff until the new Program Director had been identified. Conversely, it was seen that additional manpower may be needed to set the planning and implementation process in motion. It was urged that ME give careful consideration to the possibility of using consultant expertise during the next few months and until a more permanent arrangement was made.

VI. Logistical Considerations

The Nutrition CRSP has had a serious disruption in the flow of preparatory work consequent to the hiatus in a funding decision. Normally it would be ~~undertaken~~ from the time of submission of the project proposal until the implementation of the preliminary field work. This has not taken place to the extent that would have been desirable. The result is a major time pressure for a large amount of preparatory work that must precede actual study data collection.

Some tentative target dates provided by the Interim Program Director were:

Initiation of study data collection - September, 1982

"Finalization" of Design Decisions* - June, 1982

These target dates provide a framework for a timetable of preparatory work. If, in the opinion of the SCB, this timetable cannot be met, then it follows that plans for the initiation of data collection must be modified.

*Modifications may be made on basis of final stages of preliminary studies: see section IV. A major meeting of the SCB is anticipated in June, a preliminary meeting in April.

A. Development and Testing of Instruments

At present, meetings of all TAGs (except that proposed in this report) are scheduled by mid-April. It is crucial that these groups, through correspondence and discussion be prepared by early May to provide reports to the SCB on

- the desirable instruments
- the desirable schedule of measurements
- the draft of the data coding forms (see section V.E)
- the manner in which data collected would be used in analyses (e.g., how an "index" relating to the function would be generated and how this would be applied and interpreted in analyses)
- assurance that the methodology selected, when applied in the framework of the planned design and with the expected numbers of individuals/events would be sufficiently sensitive to measure reasonably expected differences (or alternatively, the required design changes to permit detection).

B. Preliminary Field Study

The preliminary examination of communities will begin by late February. It is essential that by June 1 this be completed to a sufficient extent to (a) permit confirmation of community selection; (b) provide information about expected demographic structure, birth rates, morbidity rates, nutritional status, and their range of variation and other descriptors that may be expected to affect design decisions; (c) provide information about expected attrition rates for households and about extent and patterns of temporary out-migration of individual household members. (See section II)

It is expected that preliminary community data collection will continue (along with instrument development and testing) such that by August a sampling frame will be available for final household selection and study implementation. Community consent will have been obtained at an early stage of the preliminary work. Household and individual consent must be arranged as the sampling process is initiated. Completion of the general community description should be completed before or shortly after study initiation. More detailed descriptions of household, and, by aggregation, the community will be obtained during year 1 of data collection.

An important question to be addressed by each project in the preliminary survey is the operational definition of a "household" and the size and composition of households so defined. Both aspects of this have major significance in study design (and logistics within the constraints of a fixed budget).

C. Data Management Capability

A major and crucially important gap in planning to date is the development of an effective data management system. The features of such a system are discussed in section V. What is important to appreciate is the urgency of action in this regard if the system, including coding forms, is to be in place when data collection begins in September.

Toward this objective the tasks may be divided into two major categories: (a) the mechanisms and processes of data coding, quality control and transmission (the system) and (b) the generation of appropriate interproject data recording and coding forms. Both aspects require field testing with real or mock data.

In establishment of the system there needs to be (i) interproject concurrence on the features of the desired system, (ii) establishment of an interproject data handling committee (which would probably be a standing committee rather than a TAG but which, at appropriate times, might need to draw upon external consultation resources) and (iii) a statement of intra-project plans for the location(s) of project-specific data banks and description of the operational characteristics of the data bank storage system (computer, language, interface capabilities, etc.). It would seem to be a matter of considerable urgency that the ME recruit existing (in UCB) or new human resources to play a major role in the facilitation of the development of this system and in the operation of the system.

Given SCB policy decisions on the desirable operating features of the Data Management System, the standing committee should be charged with responsibility for the design of the system. Emphasis should be given to the urgency of action. A target for final design decisions should be mid-May (with information to SCB in June) and operational testing during the summer of 1982.

To develop the data recording and coding forms, the standing committee again should be given responsibility but it must be given support of the TAGs. Specifically it should be an assigned responsibility to the TAGs that as well as deciding on the instruments of data collection the TAGs should also provide a consensus decision on the data recording and coding forms to be used. As soon as the instruments are approved as a part of the final design (June), the draft forms should be transferred to the standing committee with the charge that they be assembled in a manner that is compatible with both research design and data coding approach decisions. The English version may then be translated into appropriate languages. Some forms can be finalized immediately; others will require field testing before finalization by the standing committee and printing of the multiple copy forms. The target date for completion of this task is the date of initiation of data collection (September 1, 1982).

Immediately after the initiation of data collection there will be urgent need for intraproject close monitoring of data quality control mechanisms to ensure they are operationally effective. Interproject sharing of experiences in problem solving will be invaluable at this time. Progressively the close monitoring can be replaced by secondary data quality monitoring (e.g., screening for outlying values, examination of distributions, screening for missing data, etc.) on an intraproject basis and interproject comparisons (see section IV).

The importance and urgency of this area of preparatory work cannot be overemphasized. Without data recording and coding forms, and without a system for the assurance of reliable data editing and storage, data collection cannot be initiated. It may prove that the time required for development and implementation of this system is the operational constraint to implementation of data collection.

D. Dietary Data Management Capability

Decisions and action must be taken on the system(s) that will be used to convert food intake data to estimated nutrient intake for the 50,000 or more observations in each project called for in the present design (this may be modified pursuant to discussions and approaches suggested in section II). It may be that interim processes and composition data banks will be used in the initial phases but if so, precautions will have to be taken to avoid the possibility of multiple data sets generated with different methods at different times.

VII. Other Organizational Considerations

In the process of the TAG's deliberations, certain other matters arose. These are identified and briefly discussed below. All are matters that will require the attention of the SCB and ME although some may have a lesser time priority than the issues discussed above.

A. Inter- and Intra-project Communication and Collaboration

The Nutrition CRSP is a very large program involving investigators in widely separated locations. The program is designed to be collaborative in nature--collaborative in the planning, conduct, and interpretation of the research.

In the absence of specific and continuing investment of time and effort there is serious risk that individual investigators will become isolated, that the strength and advantage of collaborative inputs will be forfeited, and that the research will be uncoordinated.

The TAG notes that some aspects of this are addressed, at least with regard to the core research, by arrangements of the SCB, the structure of interproject TAGs concerned with development of design detail, the decision

to apply standardized methodology (identical or approved analogous methods in each project applied in identical protocols) and the system of central data monitoring discussed elsewhere in this report. Important as these steps are, in themselves they are not enough.

It is recommended that each of the country projects take steps to ensure intraproject communication and, importantly, intraproject definitions of responsibility. It is recommended that approaches be shared between projects and that information concerning assignments of responsibility be shared.

It is recommended that serious consideration be given to the question of routes of communication between country projects and between country projects and the ME. That is, it is essential that Principal Investigators be kept aware of all project-related information flow. It is equally important that individuals within country projects be kept aware of at least those aspects of the information flow that influence the operation of the country project and their roles in it. If this is to be a two-stage operation with all inter-project communication routed through the PI, then it is essential that the PIs establish mechanisms for rapid follow-through.

It is recommended that the ME consider the feasibility and desirability of establishing a periodic "Nutrition CRSP Information Bulletin" that could be distributed to all participants with the intent of keeping them generally informed about the overall progress of the Nutrition CRSP.

B. Publications and Shared Data Use

In the process of considering approaches to data management and analyses, the TAG was repeatedly reminded of the important issues surrounding the control of use of data. These might be summarized as:

- i) protection of rights to data
 - the rights of individual investigators
 - the rights of intraproject US and country institutions
 - the rights of country projects
 - the rights of the CRSP Nutrition and Function program
- ii) assurance of acknowledgement of participation of individuals in generation of data in any publications arising from the data
- iii) control of intra- and inter-project interpretations of data analysis (minimization of potentially contradictory interpretations of data)

The TAG sees it as the responsibility of the SCB to address the matter (in keeping with the assertions in the original grant proposal) and urges that this be done at an early stage before any problems arise. It is noted that the issues identified are characteristic of any multiple investigator collaborative

project and experience has shown that mutually satisfactory arrangements can be made. It is noted also that without prior development and application of a program policy on publication and other data use, there can be very serious internal misunderstanding and avoidable intraprogram conflict.

The TAG points out that in developing such a policy, the SCB should take into account the various types of data sets that may arise in this program.

- i) core data consolidated and analyzed in an inter-project data base
- ii) country project data which includes a segment of (a) but may be expected to include also additional data collected for country project-specific purposes (i.e., supplementary studies)
- iii) specific supplementary data sets generated between two of the country projects in collaborative substudies
- iv) country institution data collected for specific use by the country institution and not included in either the country project or interproject data basis
- v) data bases generated by individual investigators and incorporated into one or more of the above.

Some of the various levels of requests for data are,

- i) collateral, intraproject data for specific purpose analysis by investigators on their students
- ii) collaborative interproject data
- iii) intra-project single country analysis
- iv) collaborative analyses of core data

and some of the potential levels and forms of "publication"

- i) interim project reports
- ii) full publications of papers in refereed journals
- iii) inclusions in chapters or monographs authored by individual investigators
- iv) "private communications" to other investigators
- v) contracted release to other supporting agencies (e.g., if additional funding is solicited from other sources)
- vi) abstracts for oral communication and oral presentations at scientific meetings (with or without further publication)
- vii) graduate student theses (which are published in the form of microfilm)

C. Household Socio-Economic Characteristics

As noted in section II.B, the present TAG has questioned the feasibility and desirability of including socioeconomic status as a household selection variable. At the same time this is recognized as a potentially important analysis variable. The following comments are offered to urge further consideration and discussion of the matter, possibly in the setting of a new TAG.

Assumptions

1) There are a series of "background variables" (independent variables/contaminating variables/etc.) that will be important to measure in order to:

- a) control for unforeseen effects on relationships among other prime variables of food intake, nutritional status and outcome variables.
- b) explain and elaborate manipulable variables or factors affecting food intake differentials and related observations.

2) Most if not all these background variables can and should be conceptualized as household qualities or characteristics which affect the lives and situations of persons who are members of the households. This is the reason for specifying households as the units being sampled.

3) Ideally all these "background" household variables should be quantified, usually as ordinal scales.

Household Variables

A. Household economic/material resources constitute a major cluster of variables that have demonstrated effects on food intakes, food use, and other individual level variables (including health/morbidity). In our experience, based on field projects in North America, Latin America and elsewhere, the following are important to consider (and to include, where possible in a household interview schedule, plus related observations):

- a) "Material style of life," defined as ownership of range of physical equipment and resources including house (style, vehicles, cooking equipment, furnishing, modern equipment (radio, TV, lamps, etc.)
- b) "cash income" (admittedly hard to measure) (and degree of cash reliance)
- c) amount of land owned/controlled
- d) animals owned/controlled
- e) perceived wealth status as defined by other community members
- f) occupation/economic pursuits

B. Household composition, "dependency ratio" and related qualities have some hidden problems, particularly in the case of persons who are temporarily or periodically absent, etc....

C. Household "cosmopolitiness" or "social articulatism" is usually a powerful variable — made up of contacts with mass media (magazines, newspapers, radio, TV, etc.), travel, and knowledge of region/national affairs.

Other variables that deserve careful scrutiny:

1. education/literacy of household heads (\bar{q} , σ^2)
2. religiosity (as well as denomination)
3. ethnicity (including sub-ethnicity or "remnant" ethnicity)
4. migration history
5. present "migration index" or periodic "labor migration pattern"
6. social memberships
7. social network
Social network (kinship and non-kinship relations) are frequently of great importance in affecting food use, food availability, etc.
8. male/female role performance (task allocations, activities) in household. (Note that this relates closely to the problem of measuring "social competence".)

Measurement and control of all these variables poses a large problem for the projects. Some of the variables may be manageable with moderate effort — other variables discussed here may require very special attention.

Consideration of these variables should be closely related to details of operation aligning the key dependent variables.

An Example of the Operation of Data Management
at the Field Level

This discussion is based on a number of assumptions:

1. The community(ies) has been selected
2. The census has been completed and households for study identified (priority list because of possible refusals) - (Census protocol discussed below)
3. Data collection forms have been developed, translated and precoded as possible (code development discussed below)
4. Human subjects protocol and consent forms have been developed and approved for field use.

The discussion is divided into four major sections. Study personnel study subjects, data collection and data handling.

Study Personnel

It is assumed that study personnel involved in on site interaction with study subject recruitment, data collection and data handling will be indigenous, familiar with the language, culture, customs and area. (It will be necessary to select these persons with attention to local mores to avoid problems of invasion of privacy and breach of confidentiality.)

Study staff will be needed to visit households, explain the study, do interviews, make measurements, collect other needed data, maintain contact with the households, review and code collected data, package and ship material for off site analysis (e.g. blood specimens) and prepare and forward collected, reviewed, coded data forms to country and US center data facilities. The division of labor among study staff and number employed will depend on the specific country and area. The only exclusive tasks are those of data collection and review; one person should not review work done by him-or herself.

Study personnel will have to be trained in the tasks for which they were employed. Interviewer training manuals, measurement directions and data handling manuals will have to be developed for use in the training and in regular refresher courses for study staff. Practice sessions will be arranged, along with supervised and "staged" interviews and other procedures to assure conformity with protocol.

Work will be reviewed by the supervisor regularly to assure conformity, to identify those in need of further training, and to reassign or dismiss those not able to work effectively in the particular assigned tasks.

Study Subjects

A priority list of households for recruitment will be provided study staff members. Every reasonable effort will be made to enroll the selected households into the study. (Household definitions? - a group of persons the majority of

whom are related by blood, ethical or legal ties who live in some approximation to each other and who typically eat their major communal meals together or from the same food resources.) All family members will have the study explained to them according to protocol. All adults will be asked to consent individually adolescents (12+) not of voting age will be asked to consent for themselves in addition to consent from the legally responsible adult, children under 12 will be subject to consent from the legally responsible adult.

After consent is obtained, a unique identifying number will be assigned the household and each of its members (see identification system below). Appropriate arrangements will be made for completion of initial interviews, food intake observation, measurements, etc. as specified for the initial data collection contact.

At this time, the schedule of follow-up for the full study duration will be developed for each household member. This is necessary to schedule contacts, especially as different study staff members may have the expertise necessary for different aspects of the data collection. This schedule will be subject to modification if pregnancy is reported on a follow-up visit, if a birth occurs, if an individual or household permanently migrates from the community or if a study subject dies.

Data Collection

One study staff member will be responsible for overall data collection from each household. Other study staff may be involved in specific measurements or holiday fill-in or for other reasons, but the relationship and interresponsibility of study staff member and household will be maintained as fully as practicable in the community selected. Completeness of collection will be monitored at four points:

- 1) A check list of follow-up dates will be maintained for each household and individuals. The on-site supervisor will check for timely completion as noted by the responsible study staff member. Reasons for non-compliance must be recorded and reviewed.
- 2) The on-site data reviewer will note on the check list the completion and sending of data collection forms.
- 3) The country data processing facility will query the site supervisor for household or individual data collection forms due but not sent. (see below)
- 4) The US center data processing will do the same (see below)

Data Handling

Completed data collection forms will be turned over to on-site study staff for coding and review.

Coding-in general, the data collector will be responsible for completing all "self coding" and all coding for which information is available on the forms, measurement results, etc. Non data collector coders will be responsible for all complex coding requiring the use of code-books,

transformations, combinations or other manipulations of the information collected.

Review-a staff member not involved in the data collection or "in office" coding will review all of the data collection form and codes editing for completeness fo data collection, readability of information accuracy of recording and correctness of coding. Corrections will be made in colored ink and will be reviewed with data collector or coder. The supervisor will review forms before sending both to assure effectiveness of data handling and to assess the completness of data collecting, coding and reviewing (editing) staff.

After data forms are completely processed on site, the originals (top pages) of NCR forms will be separated from second pages. Shipping lists will be made up in triplicate, including form designator and household and individual identifiers. Originals and one shipping list will be sent to Country data processing center, 2nd pages and one shipping list will be sent to US Center. The third list will be retained at the site until receipt of documents is acknowledged by Country and US centers.

For materials to be sent to another site for processing (e.g. blood specimens), appropriately identified data forms will be made up and shipped, along with a shipping list, to the intended site. Copies of this list will be sent to country and US centers to advise them of shipment and that appropriate data can be expected from the other site. A copy will be retained at the on-site office until receipt is acknowledged (The same general procedure will be followed for completed data collection instruments sent to some other place for review or work. In that case, the 2nd copy will be retained at the site for safety; the receiving center will be responsible for forwarding completed copies to both country and US centers.)

Queries on incorrect or questionable data will be sent to the site supervisor by the country data center. Any queries from the US center will be processed through the country center.

Identification System

In an ongoing data collection effort, it is necessary to identify individuals as individuals and as members of households in such a way that confidentiality is preserved and yet linkage of data from different times and sources can be completed. For a study such as this, done using indigenous personnel, with data collection at a fair remove from the processing and analysis, we recommend the use of 3 or 4 separate identifiers. The first would be an arbitrary, possibly sequential or geographically connected household number with an additional two digit code for the individual within the family. This latter number could be arbitrary or could contain information relative to the individual status within the family (e.g. 01 head of household, 11 eldest male child, etc.). The second number would be approximate birthdate- if birthdate is a generally known fact in the community, it would be used directly; otherwise some personally meaningful date which could be reported by the individual on query would be substituted. The third identifier would be an initial set, using Roman letters for two names chosen in a predetermined fashion (e.g. own and mother's first given name) The fourth would be an arbitrarily coded (numeric) name for the name by which the individual is known within the household.

The use of four identifiers in this manner will help to overcome errors in identifiers, and will help to identify an individual who may be lost from one household and only to surface in another, or other problem of this sort.

Identification codes for household and name codes will be maintained in confidential form by senior project study staff on site, by the senior data management personnel in the country data facility and, for safety, in sealed form by the country and US principal investigators.

Some Types of Considerations in Data Collection
Forms

Data Collection Forms

Original data collection forms, or modifications of extant forms, should be developed by and in collaboration with the appropriate TAGs and recognizing final decisions of the SCB and ME on research design. All forms should include provision for direct coding on the forms. All forms should make provision for the separation of a duplicate copy.

In the writing, approval, establishment of coding, and preparation of specifications and operational manuals. English should be the language used. Pretesting of forms might be conducted on appropriate U.S. groups using the English language. After completion of the English language forms, translation of the data collection instrument and supporting documentation into the appropriate native languages should take place.

Translated forms should be pretested among subjects similar to, but not in the same community as, the study subjects. Results of these pretests should be reviewed by the country project field director to identify any problems requiring correction. Recommended changes, with explanation, should be presented to the Nutrition CRSP Program Director who will approve, disapprove, or refer to the appropriate TAG, SCB or consultants for advice.

Before final approval and printing, a retranslation to English of the final document should be performed independently of the first translation. This translation should be sent to the US country project center and to the ME and to an appropriate TAG member or consultant for approval.

Preliminary considerations suggest that at least the following data forms will be required:

- Census
- Human Subjects: explanation and consent
- Study Intake Interviews
- Data Forms as developed/proposed by TAGs
 - questionnaires
 - observations
 - measurements
 - on site
 - laboratory

"Meta" Data Forms

- Household)
Individual) follow up schedules
- Shipment of Material (e.g. bloods) and Data Capture Forms
- Shipment of Completed Data Forms
- Correction of Data/Coding Forms

Work Manuals

- Guide to study performance
- Data collection instructions
- Data collection specifications - why, how, etc question or measurement
- Coding manuals
- Data management manuals

Checklist of Questions
for Technical Advisory Groups

A primary purpose of the Nutrition CRSP TAGs is to provide expert and collaborative advice on the data to be collected and procedures to be used in the core research of the Nutrition CRSP. In considering overall research design, the TAG on Statistical Design and Data Management recognized that the TAGs must play an essential role in overall design decisions and indeed that only the TAGs could play this role. Specifically, in recommending methodologies to be used in the research it is essential that they consider and report on the implications and applications of the methodology in research design and in data analysis. They will have selected methodologies with certain research goals in mind; these goals will carry certain design and analysis requirements.

The following questions seem appropriate to all TAGs. Answers to these questions, with reference to recommended methodologies, will provide the necessary input for final research design decisions.

- A. What are the research questions being addressed?
- B. What data are needed to answer A?
- C. How should these data be collected (from what subjects with what frequency, for how long)?
- D. What are the measurements to be made?
- E. What are the data capture instruments?
- F. Who should be responsible for making the measurements or applying the instruments? What training requirements does this pose?
- G. How should the information be coded? Can this be done in the field?
- H. What sort of tables or statistical analyses are anticipated to use these data to address the questions in A? What ancillary data will be needed?
- I. What is the minimal number of subjects required, using the proposed methodologies and collection procedures, to answer the questions in A? (This includes reasonable assurance that failure to see an effect can be interpreted as absence of an effect rather than inability to detect a biologically or socially significant effect.)
- J. What approaches to interproject standardization are suggested for implementation before data collection begins (training period) and for implementation as a part of an ongoing quality control system?



INSTITUTE OF INTERNATIONAL STUDIES
Nutrition CRSP

BERKELEY, CALIFORNIA 94720
215 Moses Hall

Institutional Council
Summary Minutes of Conference Call
February 1, 1982

- Represented: University of California (Berkeley) - Robert Edwards (Chairman)
University of Connecticut - Clarke (representing
Alexandra van Gelder)
Purdue University - D. Woods Thomas
University of Arizona - William A. Matlock
University of Kansas - Billy G. Hudson
Program Administrator - Doris Calloway
Interim Program Director - George H. Beaton
- Not Represented: University of California (Los Angeles)

1. Voting Procedures:

It was agreed that most matters coming before the Council would be resolved by consensus without the need for formal vote. In the event that a vote was required, each institution would vote in its own right except when dealing with matters relating to country projects as such (e.g. inter project budget allocations). In this situation voting would be weighted to give equal vote to each country project. The Chairman would declare the type of vote applicable to any particular question; if there was doubt, the IC might decide by vote of members.

2. Nutrition CRSP Policy on support of Dependents

It was agreed that Nutrition CRSP funds could be expended for the support of dependents of faculty members (defined as persons holding tenure track or tenured professorial appointments) provided that the dependent(s) would reside with the faculty member in the foreign country for at least 12 months.

It was recognized that exceptions to this rule might be requested and suggested that this should be considered by the IC on a case by case basis. It was suggested also that the exceptions should be granted only when it seemed clear that the scientific aspects of the project would be jeopardized by refusal of dependent support and when the scientific conduct of the project would not be compromised by the diversion of the country project funds for this purpose.

3. Budget Matters

The status and basis of funding by AID of the CRSP and collaborating institutions was reviewed and clarified. Specifically although there has an AID-approved budget through Mar. 31, 1983, present funding was until Mar. 31, 1982. California has requested release of onward funds and in June expects to prepare budgets for AID approval and commitment for the period to March 31, 1984.

Dr. Woods Thomas pointed out that this arrangement was different from the original intent and AID practice with regard to other CRSPs. Specifically in other CRSPs, institutions received two year forward funding.

There was agreement that the IC should explore this matter with the intent of having the arrangements reopened with AID. Dr. Woods Thomas agreed to draft a letter which would be sent to California and circulated to all members of IC. It was agreed that this should be an agenda item for the June meeting.

4. Institutional - Country Problems

The question was raised concerning common features of all country agreements and required inclusions. It was noted that agreements with the U.S. Universities stipulated that lower Tier agreements be modelled on the California - U.S. institution agreement (and be reviewed by California). All obligations of the first Tier agreements that involve actions of the foreign institution would have to be passed on in lower Tier agreements. Two specific examples were cited: the audit requirement and the requirement to send data to the U.S. institution for transfer to California.

5. Next Meeting

There was agreement that an IC meeting should be held in proximity to the SCB probably in late June, 1982 and that this be the first annual meeting of the IC.

6. Selection of Permanent Chairman

It was agreed that the chairmanship of the IC should be the first agenda item for the annual meeting.

SCHEDULE OF NUTRITION CRSP MEETINGS

New Orleans, Louisiana

Time	Sunday April 18	Monday April 19	Tuesday April 20	Wednesday April 21	Thursday April 22	Friday April 23
9 a.m. to 6 p.m.	TAG ON SOCIAL COMPETENCE (Reception Room 8)*		TAG ON FOOD INTAKE (Reception Room 8)* TAG ON REPRODUCTION (Reception Room 9)* TAG ON DISEASE (Reception Room 10)*		Scientific Coordination Board (Reception Room 8)*	
Evenings	TAG ON NUTRITIONAL STATUS (To be held in hotel room)		AVAILABLE FOR COUNTRY PROJECT MEETINGS**			

February 26, 1982

- * All Reception Rooms are located at the Louisiana Superdome
- ** Not scheduled; to be arranged by country project discussion.