PDAAJ-296
59867

TRIP REPORT
TECHNICAL ASSISTANCE TO THE "ORT GROWTH MONITORING AND EDUCATION" PROJECT AT THE INCAP

REVIEW OF "EPIDEMIOLOGY OF PERSISTENT AND CHRONIC DIARRHEA IN CHILDREN IN RURAL GUATEMALA" RESEARCH COMPONENT

A Report Prepared By PRITECH Consultants:
ROBERT E. BLACK, M.D., M.P.H.
GERALD T. KEUSCH, M.D.

During The Period:
JULY 17-21, 1988

TECHNOLOGIES FOR PRIMARY HEALTH CARE (PRITECH) PROJECT
Supported By The:
U.S. Agency For International Development
AID/DPE-5969-Z-00-7064-00

AUTHORIZATION:
AID/S&T/HEA: 11/30/88
ASSGN. NO: ACP-002-GU

7/26/88
REB:msl

TRIP REPORT

Technical Assistance to the "Oral Rehydration Therapy,

Growth Monitoring and Education" Project at the

Nutrition Institute for Central America and Panama (INCAP)

Review of "Epidemiology of Persistent and

Chronic Diarrhea in Children in Rural Guatemala" Research Component

July 17-21, 1988

Robert E. Black, M.D., M.P.H.
Department of International Health
Johns Hopkins School of Hygiene and Public Health

PRITECH Consultant

The research study "Epidemiology of Persistent and Chronic Diarrhea in Children of Rural Guatemala" is being conducted by Drs. Jose R. Cruz and Alfred Bartlett in the rural community of Santa Maria de Jesus. This is a longitudinal community-based study of the epidemiology of persistent diarrhea among children up to 35 months old. It is designed to determine the incidence of diarrhea of various durations and association with enteric pathogens, as well as to evaluate potentially important risk factors, such as nutritional status, cell-mediated immunocompetence, feeding and treatment practices, etc. The study was planned to include a panel of 150 children 0-30 months old at enrollment and a cohort of 150 newborns entered at 10 per month and was intended to continue for 3 years. As expected for this rural population, female literacy is low and infant mortality is high, with diarrhea and respiratory infection as predominant causes of death. Breastfeeding is nearly universal and continues for more than one year (table 1).

The study is at its midpoint with 1-1/2 years of surveillance completed, and I was asked to conduct a midterm review in July 1988. Since the project funds have been spent at a faster rate than anticipated due to INCAP salary increases, the available funds will be exhausted after 2 rather than 3 years of field work. Thus, after reviewing progress made in the study to date, the quality of information obtained and the plans for data management and analysis, I was asked to recommend a course of action regarding the project's future activities.

In anticipation of the visit, Drs. Cruz and Bartlett had initiated a preliminary analysis of the data from the first year of the study (February

1987 to February 1988). Planning began in January 1988 for this analysis; data processing and attempts at analysis continued up to (and throughout) my visit. Because of delays in data processing, many of the planned analyses were not accomplished; however, the results that were available are considered adequate to make the judgements needed at this time.

STUDY ACTIVITIES AND RESULTS IN THE FIRST YEAR OF SURVEILLANCE

Study Population

A panel of 147 children was enrolled at the outset and newborns began to be entered in surveillance during year 1. Approximately 100 newborns had been enrolled in year 1 and the remaining 50 will be entered by August 1988.

During the first year, 40 children were dropped from surveillance, including a sizable number who withdrew from the study shortly after the beginning of surveillance, and were replaced. Children leaving surveillance because of migration later in the study were not replaced, nor were the few children who died. Since this initial period, there has been a low rate of voluntary withdrawal or migration out from the study. By the end of 2 years, after entering all 150 newborns, it is anticipated that 300 children will have been in the study for at least 6 months (maximum 2 years). These 300 children should provide 4000+ child-months or 333+ child-years of observation. The age distribution of children in the study is appropriate.

Morbidity Surveillance:

The twice weekly household visits are collecting comprehensive information on the occurrence, clinical manifestations and treatment of

diarrhea. Limited additional information is collected on other serious illnesses. The information collected is appropriate to the study and is of high quality. It should also be noted that necessary medical care is provided for serious illnesses in accord with the highest ethical standards appropriate for such a community-based study.

Diarrhea Rates:

In the first year, 1,295 episodes of diarrhea (≥4 unformed stools in a 24-h period) were identified. This is an incidence of 7.7 episodes per child-year. Of these episodes, 140 (10.8%) lasted at least 14 days. The mean duration of the prolonged episodes was 26 days (range 15-229 days). The proportion of episodes lasting more than 2 weeks was estimated in the original protocol to be 20%, since this was the proportion found in an INCAP study done 20 years ago in Santa Maria Cauque. However, the 10% found so far is more consistent with other more recent studies in Peru, Bangladesh, and India (3-14%). Likewise, the incidence of prolonged diarrhea (83 per 100 child-years) in this study is not dissimilar to that found in Peru and Bangladesh (24-76 per 100 child-years).

In Guatemala the highest rates of more prolonged diarrheas appear to be in children 6-11 months and 12-17 months old (tables 2 and 3).

Future analyses should use the duration categories 1-7 d, 8-14 d, 15-21 d, 22+ d. This will permit comparison with other studies. An operational definition of prolonged diarrhea should then be diarrhea with duration of ≥15 days.

Recurrent Episodes of Persistent Diarrhea:

In the first year, 78 (29%) of 272 children participating had at least 1 episode of persistent diarrhea. Of these, 60% had only one episode and 40% had more than one (table 4).

Characteristics of Shorter vs Longer Episodes:

Prolonged episodes of diarrhea more frequently had reported mucus in the first week and vomiting in the second week than shorter episodes. They also had less frequently reported fever in the first week and anorexia in the second week than shorter episodes. None of the findings seem particularly useful to serve as predictive factors for episodes likely to become persistent. Prolonged episodes are as likely as shorter episodes to have been treated with ORS, antiparasitic drugs, aspirin and home remedies in the first week of illness, and less likely to have been treated with antibiotics. Collection of this basic information should continue, but does not suggest that therapies in the acute phase of illness, as used in this population, affects the duration of the illness (tables 5 and 6).

Characteristics of Children With None or With One or More Persistent Episodes:

None of the characteristics of the environment or family analyzed so far seem to be different in children with or without persistent episodes (table 7).

Breastfeeding status was analyzed for diarrhea associated with Shigella or Campylobacter. Non-breastfed children had 4 (16%) of 25 episodes associated with one or both of these pathogens last ≥14 d. Partially breastfed children

had 17 (25%) of 69 episodes last ≥14 d (table 8). Future analyses will likely have episodes in the exclusive breastfeeding group and will need to be stratified by age as well. It should be noted that diarrhea associated with these pathogens appears to have a greater frequency (22%) of persistent episodes than that associated with other pathogens, or with no agent identified. Feeding patterns prior to or during the acute phase of illness have not otherwise be analyzed.

Risk Factors:

A review of tabulations of the raw data file suggest that some possible risk factors will be difficult or impossible to evaluate given the existing feeding patterns. Animal milk is given to children very infrequently in this community, and relatively little other animal protein is given. This will limit the ability of the study to evaluate lactose intolerance or protein sensitization as possible factors contributing to persistent diarrhea, although it would suggest that in these children who have a high rate of prolonged diarrhea, the possible risk factors mentioned are unlikely to be of importance. It appears that foods (e.g., frijol negro) high in lectins are infrequently given to young children, but are given in some frequency to older children. The very large amount of data coming from the 24-h recall will be of little additional use, indicating that data collection on dietary intake could be simplified substantially.

Children who developed at least one persistent episodes may have been more malnourished at the beginning of the period of analysis (April 1987). They may also have had a greater negative change in nutritional status between April 1987 and February 1988.

Cell Mediated Immunity:

Thirty children have been tested by the Multitest CMI system for delayed type hypersensitivity (table 9). Eleven children 6 months old and 11 children 17 months old were included, 6 children were 18-24 months old. DTH responses were as expected from experience with the Multitest CMI in Peru. No adverse reactions were noted. Only 2 (7%) of children had no DTH responses.

Association of Shorter or Longer Diarrheal Episodes with Pathogens:

A preliminary analysis was attempted of tests done to date on specimens from diarrhea and routine controls during the first year. Since some testing is incomplete, the denominator changes from pathogen to pathogen. A tabulation was only produced at the end of my visit, and therefore there was insufficient time to check the accuracy of the results. There was also some confusion in the tabulation labels for the varieties of adherent E. coli which could not be clarified prior to my departure. Nevertheless, some conclusions seem obvious. As seen in table 10, some agents are very infrequently identified. Others are not particularly associated with persistent diarrhea or even with diarrhea in comparison with routine cultures.

Of special interest are the adherent E. coli. Testing to date has included relatively few samples and has not identified separately auto-aggregative E. coli. Nevertheless, adherent E. coli were the most frequently found agents among the specimens tested and were associated with 59% of persistent episodes. Enteropathogenic E. coli identified by sero-grouping have been infrequent in the small number of samples tested, but suboptimal antisera have

been used for testing. It is clear that testing for these potentially important agents has been laborious and beset with problems, resulting in a tremendous backlog of E. coli for testing.

Other agents (not shown in table 10), e.g., Yersinia enterocolitica, Aeromonas hydrophila or Plesiomonas shigelloides have been very infrequent. There is no reason to continue to search for Yersinia which requires a separate plating medium (CIN agar).

Data Management, Processing, and Analysis:

The collection of data seems to be in good shape, but additional attention needs to be given to cleaning of data files, editing, and cross-checking. A system must be established that ensures the quality of information in the study. This includes verification of the accuracy of some types of information, e.g., the child number and date of laboratory specimens prior to coding of results and data entry. It includes entry of complete information in batches, e.g., monthly, which will make editing raw data files and updating master files easier than a continuous flow of data. The individual data files need better cross-checking, e.g., to look for morbidity-laboratory record matches and mismatches. Many data accuracy checks need to be done now for all data available, and should be done more quickly after data entry in the future since this may make work easier to correct.

The data are being handled entirely by microcomputer, which I fully support. However, the limitations of microcomputers (speed, available memory) and the advantages of using statistical and data base packages need to be more carefully considered. The raw data files for this type of longitudinal study with daily information are enormous. They must be condensed into analysis

files which create new variables and link relevant other variables. Of most importance now is to create a diarrheal episode file from the longitudinal morbidity file. This should include selected information on episodes, e.g., starting date, duration, clinical characteristics and treatment in the first week, etc., from the morbidity file and culture results (during and convalescent). Decisions need to be made now on how to summarize multiple cultures done during the episode. The episode file should probably also include data on pre- and during-illness dietary patterns from the 24-h recall dietary file.

It is probably best to keep other files, e.g., child and family data, and anthropometry separate for now. New condensed files on "out of episode" cultures, on calculated denominators (i.e., days of surveillance), and on routine feeding patterns need to be created.

Given the pressures of other responsibilities, the data management has not been given sufficient attention by the investigators or project staff. There should be frequent communication and planning meetings to ensure that the data are being handled optimally. It is also apparent that there is a need for a full-time data analyst working for the study from now until at least the end of 1989. This analyst must be skilled in the manipulation of large data sets, and in the use of data base and statistical software. In addition, to the full time analysist, occasional help in programming and statistics from more experienced staff will be needed.

Recalculation of Necessary Sample Size:

Since data from the first year are now available, it is best to recalculate the necessary and achievable sample size. The original estimates were based on certain assumptions about the diarrheal rates, pathogen recoveries, risk

factors and age groups of interest, and analysis methods. Some of these need to be reconsidered.

If we wish to detect a risk factor with a RR of 2 for persistent diarrhea with a power of 80% (α = 0.05), and if that factor affects at least 20% of episodes, and if the proportion of all diarrheas lasting more than 14d is 10%, the minimum sample size in the smallest group is 221 (or total episodes 2210). This might be reduced somewhat by less conservative estimates, but on the other hand the repeat episodes in the same child will require statistical procedures to control for the individual, thus increasing sample size. Thus, it is left that approximately 2200 episodes of which 220 are persistent would be needed.

If we wish to detect a RR of 2 for a factor affecting half of the children with a power of 80% (α = 0.05), and if the proportion of children with at least one episode is 0.3, then the minimum number of children needed is 227.

Another consideration is that diarrheal duration can be used as a dependent variable in a multivariate analysis providing much greater power to examine the determinants of duration. Thus, selected risk factors, pathogens and age can be evaluated individually and interacted providing perhaps the best analysis method.

At the end of two years of surveillance, there will be 300 children who will have at least 6 mo of surveillance. The children will provide approximately 2800 total episodes or at least 250 episodes per 6 month age group and at least 280 episodes of $\geq$ 14d duration. Thus, 2 years of surveillance should provide sufficient information for all of the analyses of interest.

Budget:

The study budget will be nearly exhausted by the end of 1988. This is because the study began later than expected, but had some early costs, and because of INCAP salary increases affecting all study personnel. It is anticipated that there would be $8000 of the original study budget remaining for 1989. The cost of continuing surveillance until February 15, completing laboratory work, cleaning data files and creating analysis files by May 1988 will be $20,000. Thus, an additional $12,000 would be needed to complete data collection and produce a clean data set. Additional support will be needed for analysis for at least 6 mo.

RECOMMENDATIONS

1. Continue surveillance until February 15, which will end 2 complete years and provide sufficient information for analyses of interest. This will also provide suitable information from all seasons of the year and avoid a seasonally biased data set.
2. Drop 24-h dietary history and ask only about specific types of food that bear on hypotheses of interest, e.g., animal milks, animal protein containing foods, high lectin foods.
3. Do Multitest CMI on all children in study who have not been tested yet, and repeat the test after 6 months on all children.
4. Simplify laboratory evaluation of stool specimens by:
   a. dropping PVA specimen and trichrome stain for second year of study,
   b. dropping CIN agar and attempts to isolate *Y. enterocolitica*.

5. Sample from stool specimens to select the ones most necessary to test for enterotoxigenic and especially enteroadherent E. coli. The E. coli isolated from all previous specimens have been stored. Testing for enterotoxin production is not so far behind, but evaluation of adherence properties in HEp 2 cells has been very slow. If personnel are available, the enterotoxin testing could be done on all specimens, but the adherence assays will definitely need to be done on a selected sample. This sample should consist of all E. coli isolated from episodes of diarrhea that lasted more than 14 days. There will often be a specimen from at least the first, second and third weeks of such episodes. The sample of specimens should also include diarrheal episodes that did not last longer than 7 days and healthy children. Both of the latter groups should be age- and date-matched to the persistent episodes. Thus, there would be an estimated 280 episodes of persistent diarrhea (with 3 cultures each), 280 episodes of short duration diarrhea and 280 cultures from healthy controls. The result of this sampling is that there would be a maximum of 1,400 cultures or 7,000 E. coli to test instead of 5,000 cultures with 25,000 E. coli. The reduced work load should be possible to complete by less than 2 months after the end of surveillance. The adherence assay is time consuming, but must be done with the highest quality, since the adherent E. coli are of special interest for this study. Care must be taken to differentiate the local, diffuse and auto-aggregative forms of adherence.

6. Improve the data management by establishing a more systematic process of data checking, cleaning, editing, file updating, file matching, etc. This will require more time from a skilled computer data manager and better

communication between the investigators and the data managers.

7. Begin to establish the variables of interest to evaluate the study hypotheses and to create condensed data files with these variables. Of most importance is to create a diarrheal episode file with the relevant clinical, dietary and laboratory information. This file will be much simpler to use than the longitudinal data files.

8. For adequate data management, the study needs the full time attention of a skilled data manager/analyst immediately and for at least the next year. In addition, the study would greatly benefit from the assistance of a higher level programmer when problems are encountered. The study will also have an increasing need for input from a statistician at INCAP.

9. The study cost will exceed the existing budget by a small amount if the recommendations above are followed. It is then suggested that an additional $12,000 will be needed to take the study to May 1989. It is understood that support for the data analyst and statistician will be available from other sources (INCAP funds).

CONCLUSIONS

The study will provide sufficient information after two years to achieve the original objectives and to evaluate the relevant hypotheses. The field and laboratory procedures are of high quality and should yield suitable data for the analyses of interest. Sampling of stool specimens to test for various forms of pathogenic E. coli will provide the necessary information without the delay and substantial extra expense that would result if all specimens were tested. This will not compromise the study. Data management needs to be improved to

ensure the integrity and accuracy of the data files and to create condensed secondary files that are easier to analyze. The study as a whole is going well and, with correction of the problems identified, should make an important contribution to our understanding of persistent diarrhea in this setting and the risk factors that may have both local and global importance.

TABLE 1

SANTA MARIA DE JESUS

CHARACTERISTICS OF STUDY POPULATION

OCCUPATION OF MALE HEAD OF HOUSEHOLD:

AGRICULTURE	82%
OTHER	18%

REPORTED LITERACY (AGE ≥ 10 YEARS):

MALES	77%
FEMALES	39%

PROPORTION OF POPULATION UNDER AGE 5 YEARS: 18%

MORTALITY - LIVEBORN CHILDREN UNDER AGE 5 YEARS: 146/1,000

REPORTED CAUSES (CIVIL RECORDS)

"DIARRHEA" OR "DEHYDRATION"	25%
"PNEUMONIA" OR "RESPIRATORY INFECTION"	22%

PROPORTION OF CHILDREN BREAST-FEEDING, BY AGE GROUP:

0 - 11 MONTHS	99%
12 - 17 MONTHS	95%
18 - 23 MONTHS	63%
24 - 29 MONTHS	54%
30 - 35 MONTHS	14%

TABLE 2

RESULTS - SANTA MARIA DE JESUS

(Feb. 16, 1987 - Feb. 29, 1988)

AGE DISTRIBUTION OF DIARRHEA , BY DURATION OF EPISODES

Episode duration	percentage of episodes occuring in age group					
	0-5 mo.	6-11 mo.	12-17 mo.	18-23 mo.	24-29 mo.	30-35 mo
< 14 days (n=1,155)	11%	18%	22%	23%	17%	9%
≥ 14 days (n=140)	11%	26%	21%	16%	14%	11%

(x^2 =8,489 , 5 D.F., 0.20 <p <0.10)

(all ages, 0-35 mo.)	percentage of all episodes in age group with duration ≥ 14 days					
	0-5 mo.	6-11 mo.	12-17 mo	18-23 mo	24-29 mo.	30-35 mo.
(10.8%)	10.9%	14.8%	10.6%	7.8%	9.3%	12.9%

TABLE 3

RESULTS - SANTA MARIA DE JESUS

(Feb. 16, 1987 - Feb. 29, 1988)

AGE DISTRIBUTION OF DIARRHEA, BY DURATION OF EPISODES

INCIDENCE-DENSITY RATES OF DIARRHEAL EPISODES
WITH DURATION ≥ 14 DAYS, BY AGE GROUP

	0-5 mo.	6-11 mo.	12-17 mo.	18-23 mo.	24-29 mo.	30-35 mo.	total
No. episodes ≥ 14 days	16	36	30	22	20	16	140
child-month of observation	325	354	341	423	367	212	2,024
Rate (episodes child-month)	0.049	0.102	0.088	0.052	0.054	0.075	0.069

Age group	6-17 months	vs.	others
	66 episodes ≥ 14 days		74 episodes ≥ 14 days
	695 child-months		1,329 child-months
	0.095 episodes/child-month		0.056 episodes/child-month

$Z = 3.28$
$p < 0.001$

TABLE 4

RESULTS - SANTA MARIA DE JESUS

(Feb. 16, 1987 - Feb. 29, 1988)

RISK OF RECURRENT EPISODES OF PERSISTENT DIARRHEA:

OF 272 CHILDREN PARTICIPATING AT SOME TIME IN THE STUDY, 78 (29%) HAD $\geq$ 1 EPISODE OF PERSISTENT DIARRHEA.

OF THESE 78 CHILDREN WITH EPISODES OF D.P.

47 (60%) HAD 1 EPISODE
16 (21%) HAD 2 EPISODES
8 (10%) HAD 3 EPISODES
3 (4%) HAD 4 EPISODES
4 (5%) HAD 5 OR MORE EPISODES

FIRST EPISODES OF D.P.

N = 78 IN 1502 CHILD-MONTHS OF OBSERVATION
(UNTIL FIRST EPISODE, OR END OF OBSERVATION)

= 0.052 FIRST EPISODES / CHILD - MONTH

ADDITIONAL EPISODES OF D.P. IN CHILDREN WITH FIRST EPISODE:

N = 62 IN 405 CHILD-MONTHS OF OBSERVATION
(FOLLOWING END OF FIRST EPISODE, NOT DURING ADDITIONAL EPISODES)

= 0.153 ADDITIONAL EPISODES / CHILD-MONTH

RELATIVE RISK (ADDITIONAL VS. FIRST EPISODES) = 2.9

$Z = 6.88$
$P < 0.001$

TABLE 5

RESULTS - SANTA MARÍA DE JESÚS

(FEB. 16, 1987 - FEB. 29, 1988)

SYMPTOMS AND TREATMENTS ASSOCIATED WITH FIRST WEEK

OF DIARRHEA EPISODES, BY EPISODE DURATION

Sx/Rx	EPISODES WITH DURATION < 14 DAYS (N = 1,155)	EPISODES WITH DURATION ≥ 14 DAYS (N = 140)	p*
MUCUS IN STOOL	37%	56%	<0.001
BLOOD IN STOOL	7%	9%	N.S.
VOMITING	14%	16%	N.S.
DECREASED ACTIVITY	23%	20%	N.S.
DECREASED APPETITE	42%	36%	N.S.
"FEVER"	30%	20%	< 0.02
DEHYDRATION	4%	2%	N.S.
ORS	21%	24%	N.S.
ANTIBIOTICS	39%	28%	< 0.02
ANTIPARASITICS	13%	14%	N.S.
ASPIRIN	20%	13%	0.075
HOME REMEDIES	5%	8%	N.S.

* UNIVARIATE ANALYSIS, CHI-SQUARED TEST, 1 D.F.

TABLE 6

RESULTS - SANTA MARIA DE JESUS

(FEB. 16, 1987 - FEB. 29, 1988)

SYMPTOMS AND TREATMENTS ASSOCIATED WITH SECOND WEEK

OF DIARRHEA EPISODES, BY EPISO DURATION

Sx/Rx	EPISODES WITH DURATION 7-14 DAYS (N = 199)	EPISODES WITH DURATION ≥ 14 DAYS (N = 140)	p*
MUCUS IN STOOL	35%	38%	N.S.
BLOOD IN STOOL	9%	11%	N.S.
VOMITING	10%	24%	< 0.001
DECREASED ACTIVITY	25%	29%	N.S.
DECREASED APPETITE	45%	31%	< 0.01
"FEVER"	30%	29%	N.S.
DEHYDRATION	4%	1%	N.S
ORS	20%	24%	N.S.
ANTIBIOTICS	4[illegible]%	31%	< 0.05
ANTIPARASITICS	17%	22%	N.S
ASPIRIN	33%	18%	< 0.01
HOME REMEDIES	2%	8%	N.S.

* UNIVARIATE ANALYSIS, CHI-SQUARED TEST, 1 D.F.

TABLE 7

RESULTS - SANTA MARIA DE JESUS

(FEB . 16, 1987 - FEB . 29, 1988)

CHARACTERISTICS OF CHILDREN HAVING Ø VS. ≥1 EPISODE OF PERSISTENT DIARRHEA

CHARACTERISTIC	CHILDREN WITH Ø EPISODES ≥14 DAYS (N=189)	CHILDREN WITH ≥1 EPISODE ≥14 DAYS (N=76)
MALE	50%	55%
FAMILY SIZE > 5 PERSONS	59%	60%
SIBLING AGE < 3 YEARS	32%	24%
SIBLING AGE 3-5 YEARS	67%	68%
EARTHEN FLOOR	77%	79%
3 PERSONS SHARE SLEEPING ROOM WITH STUDY CHILD	58%	60%
WATER SOURCE = PUBLIC WELL	98%	100%
REPORT LETRINE USE BY NON-DIAPERED CHILDREN	48%	58%
ANIMALS IN HOUSEHOLD COMPOUND:		
DOGS	22%	21%
PIGS	5%	3%
CHICKENS	66%	60%

TABLE 8

RESULTS - SANTA MARIA DE JESUS

(Feb. 16, 1987 - Feb. 29, 1988)

RELATION OF BREAST-FEEDING STATUS TO OCCURRENCE AND DURATION OF DIARRHEA IN CHILDREN INFECTED BY CAMPYLOBACTER JEJUNI, SHIGELLA, OR BOTH

	BREAST-FEEDING STATUS		
	COMPLETE	PARTIAL	NONE
SHIGELLA ONLY			
> 14 DAYS	Ø	8 (22%	Ø
≤ 14 DAYS	Ø	23 (64%)	6 (86%)
ASYMPTOMATIC	Ø	5 (14%)	1 (14%)
TOTAL	Ø	36 (100%)	7 (100%)
CAMPYLOBACTER ONLY			
> 14 DAYS	Ø	8 (18%)	1 (6%)
≤ 14 DAYS	Ø	28 (64%)	9 (56%)
ASYMPTOMATIC	Ø	8 (18%)	6 (38%)
TOTAL	Ø	44 (100%)	16 (100%)
SHIGELLA PLUS CAMPYLOBACTER			
> 14 DAYS	Ø	1 (25%)	3 (23%)
≤ 14 DAYS	Ø	1 (25%)	6 (46%)
ASYMPTOMATIC	Ø	2 (50%)	4 (31%)
TOTAL	Ø	4 (100%)	13 (100%)

TABLE 9

Delayed-type Hypersensitivity Reponse to Multitest CMI

Antigen	Percent Responding (n = 30)
Tetanus Toxoid	70
Diphtheria	57
Streptococcus	13
Tuberculin	50
Candida	18
Trichophyton	10
Proteus	33

TABLE 10

Enteropathogens Identified During the First Week of the Diarrheal Episode by Final Episode Duration and In Routine (Non Diarrheal) Specimens

Enteropathogen	Episode Duration						Routine	
	1-7 d		7-13 d		≥14 d			
Giardia lamblia	56/500*	(11.2)	20/207	(9.7)	21/137	(15.3)	131/1288	(10.2)
Cryptosporidium	14/423	(3.3)	3/183	(1.6)	4/120	(3.3)	10/942	(1.1)
Entameba histolytica	3/258	(1.2)	1/113	(0.9)	1/82	(1.2)	2/365	(0.5)
Enteropathogenic E. coli	2/70	(2.9)	0/36	(0.0)	0/120	(0.0)	0/92	(0.0)
Adherent E. coli	47/156	(30.1)	10/62	(16.1)	17/29	(58.6)	27/147	(18.4)
Enterotoxigenic E. coli								
LT	41/495	(8.3)	11/196	(5.6)	8/136	(5.9)	50/1181	(4.2)
ST	19/495	(3.8)	6/196	(3.1)	8/136	(5.9)	24/1181	(2.0)
LT/ST	8/495	(1.6)	3/196	(1.5)	0/136	(0.0)	3/1181	(0.3)
Shigella sp	49/495	(9.9)	23/207	(11.1)	11/139	(7.9)	57/1151	(5.0)
Salmonella sp	2/478	(0.4)	3/205	(1.5)	5/134	(3.7)	7/1165	(0.6)
Campylobacter jejuni	63/497	(12.7)	10/200	(5.0)	14/138	(10.1)	68/1205	(5.6)
Rotavirus	18/482	(3.7)	2/191	(1.0)	3/131	(2.3)	6/1160	(0.5)
Clostridium difficile	5/146	(3.4)	2/57	(3.5)	2/28	(7.1)	15/569	(2.6)

Number positive/number of specimens tested (%).

TRIP REPORT

Technical Assistance to the "Oral Rehydration Therapy, Growth Monitoring and Education" Project

Instituto de Nutricion de Centro America y Panama

Guatemala City, Guatemala

Review of Research Component

"Epidemiology of Persistent and Chronic Diarrhea in Children in Rural Guatemala"

August 7-10, 1988

Gerald T. Keusch, M.D.
Professor of Medicine
Chief, Division of Geographic Medicine
and Infectious Diseases
New England Medical Center
Boston, MA 02111

(PRITECH Consultant)

SCOPE OF WORK:

This site visit/project review is part of the midterm peer review for the above noted research project at INCAP. The stated objectives of the review were to:

1. evaluate data collection methodologies and quality control
2. review preliminary results from year 01
3. evaluate progress in meeting research goals
4. advise on future activities
5. suggest analytical approaches

The activities proposed for this reviewer included:

1. review analysis with PI and statistician
2. review data and project likely outcomes of study
3. meet with PI and prioritize study activities
4. meet with PI and statistician to develop analytic approaches

ACTIVITIES:

August 7: travel to Guatemala, preliminary discussions with Dr. Jose Cruz, INCAP, and Elena Brineman, ROCAP
review of project documents, including original proposal, prior trip reports by Drs. Black and Bartlett, Dr. Blacks trip report dated July 26, 1988

August 8: am meeting with Dr. Cruz and Bartlett

pm meeting with Dr. Cruz and Bartlett
meeting with Dr. Gilda Pareja, Biostatistician
evening meeting with Dr. Cruz

August 9: am visit to field site with Dr. Bartlett
discussion of data collection with field supervisor
observation of data handling and processing of laboratory specimens

pm meeting with Dr. Cruz
meeting with Administration (Armando Paz)
meeting with Division Chief, Dr. Hernan Delgado
debriefing meeting with Elena Brineman, Drs. Cruz Bartlett, Delgado, Ms. Parejo, and project advisors
evening meeting with Elena Brineman

August 10: travel to Boston
review of new data analyses
preparation of trip report

GENERIC ISSUES:

1. Caveats: This review was limited by the short time available for the visit to Guatemala, which was partly compensated for by conducting evening meetings during my 2 1/2 days in country. An even more significant limitation, however, was the lack of sufficient preliminary data analyses for my review. This can be attributed primarily to problems in the data management process, including entry, cleaning, and preparation of analytical files, but most of all of computer time and problem solving, which is severely handicapped at present (see point 2 below). In spite of this, the PI was able to produce some important cross tabulations I requested during my visit. On the other hand, the limited time was balanced by my familiarity with INCAP, the Division and its programs and with the investigators conducting the research, which has been gained over 19 years of repeated visits to INCAP as a program reviewer and for collaborative research with investigators in the Division.

2. Data Management: The recent trip and report by Dr. Black points out many of the structural issues related to delayed preliminary analysis of first year data. There are two reasons, insufficient personnel (including supervisors) devoted to this projects data analysis needs, and a lack of enough hardware to make the most efficient use of the available personnel. Both of these issues are INCAP management problems that require immediate resolution. It is my understanding from discussions with both ROCAP and INCAP personnel from the Division that solving problems of data analysis are INCAP's responsibility for which Institute

funds have been allocated. I strongly urge the donor agency to insist that this be attended to immediately. Moreover, technical data management expertise at the higher levels in the Division is competent but stretched too thin, and the lower level personnel lack experience and training. I am aware from my talks with Dr. Delgado that steps will be taken to correct both problems. The responsible project advisors must now insist that this be done.

3. Research Fund Accounting: The flow of accounting information to the PI is grossly inadequate, which precludes effective and efficient management of funds. This is directly traceable to computer programming problems at the Institute, however I do not consider this an acceptable explanation. While the problems are apparently being dealt with, I did not find any sense of urgency to correct the situtation. Thus, funds are encumbered based on estimates but the report on actual expenditures is so late that timely corrections are not made and the investigators never know how much they have spent and how much they have remaining. While this is in part due to prolonged turn around time for documents at PAHO, and not necessarily under the total control of INCAP itself, one clear result is that an unspent balance remains from year 01 which cannot be spent this year without permission from USAID. I have been informed that ROCAP agrees to this but the information has not reached the accounting office. A second problem is that funds encumbered for personnel remain on the books long after personnel turnover occurs so that funds are obligated for both slots. This results in an apparent deficit in available money, and at the time of my visit the project was

actually prohibited from ordering necessary supplies such as paper towels for the lab and clinic because of this problem. While the error will be corrected as soon as the "system" catches up, I saw no evidence that this was an urgent priority except for the PI.

4. Budget Planning: This project is currently in fiscal trouble for two important reasons. First the original budget did not plan for the inflationary pressures in Guatemala, which reduced the buying power of the dollar nor did it plan for the across the board salary increases mandated by INCAP with no provisions for the funds needed, which increased the actual dollar costs to the project (amounting to about 1/3 of the present shortfall). Second, laboratory work and personnel needs estimates were based on previous data from Santa Maria Cauque which were used to calculated the expected diarrhea rates and numbers of specimens to be processed. This turned out to be an underestimate, resulting in an increase in the work load and a need for additional personnel in the field and in the laboratory (accounting for around half of the shortfall), and leading to a delay in the generation of laboaratory data. It is important to note two things here: one, the supplies categories of the budget was on target, and two, that 20 years after Santa Maria Cauque the problem of diarrhea in highland Indian villages remains of equal importance in Guatemala. The latter cannot be overstated in the perspective of the data-based intervention orientation of this project. The remainder of the budget shortfall is due to the purchase of a steam autoclave for the use of the project.

While this is clearly essential equipment for this project, I did not investigate the reasons why it was necessary to purchase it now and I reserve judgement as to where the funding should have come from. This basic facility is absolutely necessary for INCAP's infrastructure to enable it not only to carry out this project, but also its funded mandates as a regional reference microbiology laboratory for WHO.

At any rate, I consider that at least 30% and possibly up to 50% of the budget shortfall is due to inadequate management and/or planning at the institute level. For example, as soon as the budgetary pressures became apparent as a result of inflation and mandated salary increases, I think it was a responsibility of INCAP's administration to immediately begin working and planning with the investigators to help them to identify new sources of funding, to rebudget available funds to fullfill as many of the original goals of the projects as possible, and/or to renegotiate funding with the donor agency. To put the onus for the deficits on the investigators alone is, in my opinion, blaming the victim.

PROJECT EVALUATION:

First of all, this is a very important project, directed at a problem that remains as much, if not more prominent, a health problem in highland Indian villages as it was 20 years ago at the time of the longitudinal study in Santa Maria Cauque.

Second of all, this project is being scientifically conducted in an exemplary fashion in both field and laboratory components. It is uncommon to find a developing country setting

in which high quality is achieved in both field and lab, with the major supervision by local investigators, and where communication between field and lab is so clear, direct, and continuous. In both activities, a highly competent team has been assembled, with sustained motivation and individual involvement in the project, and with quality control and a system for monitoring completeness and accuracy of collected data. In contrast, the significant limitations in the flow of these data for analysis have been noted above. This is an issue requiring immediate attention by the investigators, the Division chief, and INCAP administration itself. There is a real credibility issue at stake here.

Third of all, I believe this project has the potential to fullfill its original objectives given sufficient time to carry out the studies, and that the resulting information will be of immediate relevance in planning, implementation and evaluation of interventions. At the same time, it is my opinion that this project has the potential to provide important new insights into pathogens, epidemiology and consequences of prolonged diarrhea that will permit new understanding of the problem of diarrhea in the community leading to the developoment of new priorities for programs already in action or those to be developed on the basis of the new data. This potential is, in fact, one of the most cogent arguments for the inclusion of more fundamental research components into operational/interventional projects as a general rule. Never the less, it is apparent that some mid-course corrections are both desirable and necessary.

Original Goals:

1. Determine the age-specific incidence and nutritional impact of prolonged diarrhea

2. Evaluate potential risk factors, including etiologic agents, nutritional status, previous diarrhea and other illnessess, feeding practices, treatment, and adequacy of cell-mediated immune mechanisms of host defense.

Current Study Population:

Approximately 225 children are currently being followed, including the nearly complete cohort of 150 newborns, enlisted at a rate of 10 per month beginning in May, 1987. If the project were to continue for the originally projected 3 years, about 150 cohort children would be followed for at least 18 months. Since the current data indicate that persistent diarrhea is not a major problem in the first 12 months of life, and because the incidence rate of persistent diarrhea has been about half of the estimated rate, it will be necessary to follow all the newborns into the period of maximum risk in order to determine the age specific incidence of persistent diarrhea and to identify risk factors that preceed and/or predict the episodes. Second, since the risk of persistent diarrhea seems to be greatest in the age group 12-18 months, it is imperative to observe as many infants during this high risk period in order to evaluate the causes and risk factors of these episodes. The lab data already show that most episodes are associated with several potential pathogens, and less than 1/3 are associated with a single agent. It is only the latter group that can be cleanly evaluated as far as pathogen

specific clinical characteristics and risk factors are concerned, however the former may be of particular importance to evaluate carefully since multiple pathogens may represent a critical risk factor. Thirdly, since one of the risk factors for persistent diarrhea that comes out of the preliminary assessment of data is persistent diarrhea itself, it is clearly of importance to follow those infants with one or multiple episodes during the high risk period during their subsequent years if the impact of persistent diarrhea on growth and morbidity is to be determined. Thus, only infants under 12 months of age when enrolled at the start of the study can be used for this analysis which requires a period of surveillance during and after the period of highest risk. The cross sectional data and limited surveillance of older children are interesting, providing preliminary data on the age-specific risk, but they are not of use for the definitive evaluations.

Logistical Considerations: The team is functioning smoothly at the present time, and the transportation of personnel and materials to and from the field is working well. The fecal samples are being transported in transport media, and when appropriate are being kept in a cold chest during transport. The study headquarters is equipped with a refrigerator/freezer to keep the media in. Clean, but not sterile technique is being used for inoculation as soon as samples reach the headquarters. This is acceptable considering the nature of the sample and the organisms being sought.

Data Collection and Evaluation in the Field: Data forms for the field are filled in during the home visit, and checked over by

the field worker when she returns to headquarters. Missing data are identified, and it is the practice to go back to the homes in the afternoon to complete missing data if necessary and to find study children who might not have been at home during the morning visit. There is regular evaluation by the field supervisor, and a weekly summary is made by each field worker and a total summary sheet is filled in by the supervisor. This system is working. Laboratory samples are accompanied by a data sheet, which contains the subject identifying information as well as the sample identifiers. A log of samples sent to Guatemala City is maintained in the field headquarters.

Laboratory work is under close supervision of Dr. Cruz, who is consistently present in the lab. Appropriate controls are included in tests being performed. The resulting data are considered to be reliable.

<u>Assessment of Progress and Status of Data Collection:</u> It is likely that the age specific incidence of acute and persistent diarrhea will be reliably evaluable by the end of two years of surveillance. It is unlikely that etiology specific incidence will be known, since only those cases with a single potential pathogen isolated can be employed for this analysis. It is my belief that the latter data are of importance (since specific etiology carries implications about transmission routes and interventions). This constitutes an important rationale for continuing surveillance and laboratory studies to the end of the originally planned three years.

Since the number of episodes of persistent diarrhea is less

than anticipated, and there appears to be a clear peak in the age group 12-18 months, it is imperative to maximize surveillance among this age group in order to meet study goals for cases. This will also eliminate seasonal factors as a confounder in the interpretation of age-specific data and will enable the study to determine the etiology-specific consequences of an episode on the subsequent growth and nutritional status of the infant. These indicators are critical criteria of the need to intervene. In order to accomplish this goal, all cohort children should be followed for the original project period of three years.

Risk factors collected with great frequency during the study, such as anthropometry, can most likely be assessed by the end of two years.

I am unable to assess the potential importance of overall morbidity as a risk factor nor the likelihood of assessing this factor within the original 3 year project period (or less) since no data except for diarrheal disease have been evaluated as of yet. The data on diarrhea morbidity have, however, suggested that an initial episode of persistent diarrhea is itself a risk factor. Since the peak incidence of persistent diarrhea is in the second year of life, infants need to be observed during this period in order to measure the importance of persistent diarrhea as a risk factor and the same subjects must be followed over sufficient time to determine any subsequent effects on health and growth of the individual subject. These goals will not be accomplished by the end of year two of study.

Feeding pattern data are being obtained in a qualitative

fashion, that the question is being asked if certain foods are being eaten or not by the subject at all visits. It is already apparent that sources of animal protein in the diet other than mothers milk are infrequent and there is the suggestion of a somewhat limited and common diet among the community. This is difficult to assess with assurance because of the lack of any preliminary analysis. It is my guess, however, that this will prove to be the case and that it is likely that all the data needed for the assessment of diet as a risk factor will have been collected by the end of year two and that this field activity may be terminated at that time. While this will not save much in the way of project funds for personnel (eg field workers), it will for sure save considerable time in data entry, processing, and analysis, and in this sense will save some money.

Repetitive skin test reactivity for assessment of cell-mediated immunity is unlikely to yield important or interpretable information in my opinion. This is primarily because skin tests are a measure of the overall process of cell-mediated immunity and tend to be read in dichotomous fashion, positive or negative. They do not closely reflect the continuous changes in the complex immune reactions leading to the visible skin test result, and they are influenced by multiple factors, including exposure to specific antigens, genetic capacity to respond to the antigen, nutritional status of the host, and recent infections. There is no careful analysis of risk prediction based on anergy, except for the increased post-surgical infection rate and mortality rate in some studies, which can be reduced by a period of nutrtitional

rehabilitation. Moreover, the planned periodic skin tests are to be done too infrequently to permit an analysis of any association with persistent diarrhea as either a cause or an effect. This portion of the study can be dropped now, with a small savings in purchase price of reagents.

Treatment is an important variable for analysis in the original study design. During the implementation of the project, a systematic antibiotic intervention has been included because of both clinical and ethical concerns, involving the treatment of all cases of clinical dysentery and of all patients with Shigella identified in stool cultures. It is likely that this algorithm will curtail episodes of diarrhea due to Shigella, already known to be associated with chronic diarrhea and deterioration in nutritional status. Thus, some of the potential cases of persistent diarrhea are likely being prevented. At the same time, some 10% of persistent cases detected in this study have been associated with positive cultures for Shigella. Since all of these patients should have received antibiotics according to the treatment algorithm, it is unclear whether or not these cases are treatment failures. It is essential to analyze this, because the major strategy to reduce diarrhea morbidity/mortality due to shigellosis is based on the use of appropriate antibiotics, since ample data already show that ORT does not reduce morbidity or mortality. As there has been no preliminary analysis of the data performed to date it is not possible to provide insights into the efficacy of antibiotics in this community or to judge how many cases will need to be studied in order to come to a conclusion.

For example, I am unable to determine whether the organism was identified at the outset of the episode, when treatment might be most effective, or only after 1 or 2 weeks of continuing symptoms when treatment might be too late. Similarly, a sizeable proportion of the study population (by impression of the field staff and not by analysis of data as of yet) have received antibiotics prior to being picked up by the study as an incident case. The information as to which drug was given has been collected but not coded, and of course not analyzed. This apparent promiscuous use of antibiotic in the community, much of it likely in homeopathic doses, could be an important determinant for selection of antibiotic resistant strains, contributing to the poor response to presumptive or early treatment. Since the prevalence of persistent cases is about 10%, of which about 10% will be culture positive for Shigella, the number of analyzable cases are clearly too small at the present time, since this bug is uncommon in the first year of life and only older infants are at risk. In view of the possibility that polymicrobial episodes (eg more than one potential pathogen detected in one sample or during one episode) will prove to be a real risk factor for persistence, selection of resistant strains, either from use of drugs by study personnel, by self administration or via local medical care systems, could well represent a significant risk. The inital data should be tabulated as soon as possible and an attempt made to calculate a sample size needed for the analysis. I suspect this will require a full three years of study.

ADDITIONAL BENEFITS OF MAINTAINING THIS STUDY:

This study also represents a unique opportunity to evaluate the impact of introduction of running water into households. At the present time a gravity water system is being introduced into the community, which should become operational in the coming year. Continuing the field operation for the originally planned project period will provide the baseline needed for assessing the impact of the new water supply. While new sources of funding will be necessary to accomplish this aim (grant proposals are being developed), termination of the study and disbanding the field team will make it both more difficult and expensive to start up again, with gaps in surveillance and delays in start up of a new project. Similarly, a project is currently engrafted on to this study which is evaluating the effect of a particular antibiotic intervention in established cases of persistent diarrhea. Again, termination of this study will adversely affect the completion of this intervention. These are both addtional rationales to seek ways to maintain the study for the originally planned period, although these are not the obligation of the donor agency.

PROJECT RESTRUCTURING:

There are several things than can be done to reduce the work load and save funds to continue the remainder of the project beyond the 2 year period now contemplated.

1. There is no need to continue to determine the presence of *E. histolytica* or *Yersinia enterocolitica*, since these pathogens

have not been found in a significant number of infants.

2. Similarly, I find no compelling reason to continue to evaluate toxigenic E. coli. The current data show that these are a cause of acute diarrhea but not persistent diarrhea. The lab work to evaluate LT and ST are substantial in costs and personnel time.

3. It is unclear whether or not the study should continue to look for Pleisiomonas or Aeromonas, since I was unable to evaluate the "other pathogens" category because of the lack of data analysis.

4. It may be possible to stop the dietary questionnaires soon, as it is possible that evaluation of the data already collected will show a consistent pattern of feeding before, during and after diarrhea in this population. This will not be known without the preliminary analysis of data. While stopping this activity will not reduce field personnel needs, and hence not save much money, it will significantly reduce the amount of data entry, checking and analysis, and hence reduce data analysis costs and computer time somewhat.

RECOMMENDATIONS:

1. Revise work load as suggested above, and in particular reduce laboratory evaluations of certain pathogens, redirect efforts to identification and characterization of adherent E. coli and Campylobacter, rebudget time and activities of personnel to these ends and rebudget funds.

2. Evaluate whether or not dietary component can be terminated now.

3. Rebudget program funds or find additional outside funds to

keep this study in continuous activity for the orginially proposed time in order to reach major goals.

4. Develop new proposals for prospective study of the impact of changes in water supply on water use and morbidity and mortality.