DATA MANAGEMENT, COORDINATION AND ANALYSIS

One of the most demanding aspects of this investigation has been the development of a common protocol that is both scientifically pertinent and feasible for all sites. Steering Committee meetings, as well as Recruitment & Retention meeting, have been held to determine inclusion and exclusion criteria for the study, the data to collect at baseline and the data to collect in the follow-up phase. The protocol framework also includes the design for a computer system configuration for efficient management of the data.

COMPUTING AND DATA MANAGEMENT

Data Management at the Epidemiology Data Center

The Epidemiology Data Center has always had a strong commitment to the development, maintenance, and advancement of a sound data management methodology. The PoP Software System was designed at the Epidemiology Data Center to serve multi-disciplinary, multi-center clinical trials. PoP is used at the microcomputer level. In addition to PoP, the Epidemiology Data Center uses the SAS Software System (© SAS Institute Inc.) which is a powerful utility that encompasses a wide range of functions for uses in data management, as well as statistics. By combining PoP & SAS, a total data management solution responsive to a wide variety of study designs is developed. It can be configured for distributed or centralized data entry.

POP/SAS Combination

The PoP/SAS combination provides a complete software system that facilitates data collection/management processes from initial electronic form development through the archival of the final data set on a central network. PoP is a microcomputer-based system used to design the data entry system and provide the structure and security for the data entry process. Once the data are considered clean on the PoP level, control of the data shifts to a primary database in which data are then managed on the centralized network using SAS. SAS is used to append, update, edit, archive, schedule processes, and report on data collected for a study.

PoP

The microcomputer-based, menu-driven data entry component, PoP, has a study design and a data entry module. The study design module allows the workload for the development of the data entry system to be shared across the project team. Input on database design issues from all members of the project team is required from the start. Once paper data collection forms are created, the stand alone components of electronic study design are disseminated for the development of a data dictionary, coded lists (enumerated types), form screen drawing, and generic consistency and logical edit generation.

PoP has numerous built-in features that serve the quality control needs of a study including:

Data type and range checks, double entry verification, help messages, coded lists or enumerated types, default missing and not applicable codes, independent data triggers that allow for the automatic entry of the not applicable default value into data fields that are dependent upon the original question when the response "triggers" the dependent response, and case and form login to guarantee that the ID is for a valid participant and that duplicate records are not entered.

Once data are entered and verified, logical and consistency edits are performed. Edit reports describe the errors detected and these errors are corrected through an update subsystem and an audit trail is maintained. Records that have completed the defined entry cycle of entry, verification, and editing, are considered "clean". A Prepare module selects "clean" records, flags them, and copies the data in a readable format to a designated subdirectory for transmittal to the Coordinating Center and/or for use by the clinical site.

The data collected through PoP are verified, edited, and updated to provide clean data for analysis. Every effort is made to obtain complete data from the distributed sites and participants through data entry system features (missing values flagged in the clinical site editing process) and data management techniques (missing form reports).

SAS Software System
The SAS software system becomes operational with the transmittal of data to the Coordinating Center. All the data that comprise the common data set are managed centrally at the Coordinating Center. PoP provides data set (table) descriptors for each form and a study database is established. Data are then uploaded to a common PC for appending, duplicate checking, editing, and archived on a mainframe network

A "snapshot" or backup of the study database is "taken" to **archive** the database prior to the automatic data append process. This protects the study database against any hardware or software system failure during the append process. In the event of a system failure, the original state of the study database is restored.

The **Append Data Process** system subjects data to various record checks, such as duplicate record exclusion and missing record field values, prior to appending data to the central database. Questionable data are loaded into system holding tables so that inconsistencies can be resolved by data management personnel who are alerted by an Append Log Report. Only records that are error-free are appended to the study database.

Data updates at the microcomputer level at the clinical sites are applied to the central database through the **Update Data Process** system. Any updates to the original data entry are recorded as audit data and are transferred to the Coordinating Center simultaneously with new participant data. The audit record data are verified prior making the changes to the central database. The outcome of each audit record is listed on the Update Verification Report.

The study database is routinely subjected to a series of extensive edits through a **Data Edit and Development** systems. These edits are designed to compare values of new records with records previously appended to the study database. These edits also compare recently modified values of records to other records in the database. The inter- and intra-form edits have been designed to provide the highest possible level of data integrity.

Data Reports are generated from the append process and are distributed to data management personnel at the Coordinating Center. Relevant reports are transmitted back to the clinical sites through the communications network. These reports are printed at the clinical sites to provide a copy of the errors generated during the correction process.

Data Management for REACH
The distributed data entry model was selected to allow initial processing (data entry, editing, and updating) activities to be performed at the clinical sites to generate clean data at the source of data collection and relieve the Coordinating Center staff of data entry tasks. Data entry software has been configured to prepare and send only the data necessary to contribute to the common data set.

"Clean" data at the clinical site are routinely transferred to the Coordinating Center using an unattended telecommunication process. Using specially configured communications software and properly configured hardware and a "power center", data can be transferred at night when telephone costs are lower. Security features are embedded in the telecommunication procedure to ensure that only the Coordinating Center has access to the data and that the data are transferred completely and accurately.

Data Management at the Clinical Sites

The microcomputer configuration for all the clinical sites are identical to the system configuration designed for the data management microcomputer at the Coordinating Center. Conformity among computing resources across all sites is essential for managing computing resources, human resources and responding to the computing needs and problems at the clinical sites. The Coordinating Center purchased the computing hardware and software for the clinical sites, and configured the systems with the appropriate software, testing the systems, training clinical site personnel at the Coordinating Center, and shipping the systems to the clinical sites for study use.

All participant data are entered at the clinical site using PoP.

Data Management at the Coordinating Center
All of the data from the clinical sites are managed at the Coordinating Center using SAS. SAS is used extensively for data validation through intra- and inter-form edits and other quality control measures built into the system. Standard reports are sent to the clinical sites that reflect the activity of the append, change database, and editing processes. Standard reports are also generated that list outstanding forms for participants that are due at the Coordinating Center for protocol compliance.

Telecommunications
Data are transferred and information exchanged between the Coordinating Center and the clinical sites through the unattended communications network. Clinical site personnel set their microcomputers so that transfer of data can take place at night. The Coordinating Center "hub" microcomputer initiates calls to all clinical site microcomputers.

For security, the clinical site microcomputers respond with a call-back prior to transmission. Clinical site data are held in a designated subdirectory until they are transmitted to the Coordinating Center for appending to the common data set.

Documentation
Project documentation is available for each clinical site. A Data Management Operations Manual has been provided to describe the computing environment and data entry and data management procedures. Commercial software documentation included with the configuration has also been provided.

Data Security
Several layers of security schemes have been employed to protect study data from inadvertent modification or access by non-project personnel. Security measures have been implemented on the operating system layer, at the file level, and even for specified data points. Login procedures for PoP have been established for access at the sites. Only specified study personnel at the Coordinating Center are given access to the network where the data are housed. If necessary, access to critical data points will be restricted to designated users through database software protection. Encryption schemes may be invoked at the Coordinating Center and the clinical sites to render any critical study files unreadable.

Disaster recovery procedures have been implemented to protect against potential crises. If critical hardware becomes inoperable, the Coordinating Center will have immediate access to University computing resources. The common central database will be archived to tape on a routine basis and taken to the Coordinating Center's off-site storage location. This will protect study data in the event of a disk failure or a catastrophe at the Coordinating Center or clinical sites. In the event of a disaster at a clinical site, a microcomputer system database will be restored from the Coordinating Center database. Additionally, the clinical center will be instructed to use the tape backup system supplied with their microcomputers to backup their system and data files on a regularly scheduled basis.

The clinical site microcomputers are dedicated to the REACH II project. Software not approved by the Coordinating Center is not allowed on any clinical site microcomputer to prevent contamination by a

computer virus. Virus protection software is run frequently on each of the clinical site microcomputers via the remote access software.

Participant Confidentiality
Participant confidentiality is preserved by encoding participant names into numeric IDs at the site level. Participant data sent to the Coordinating Center are identified by the numeric ID only.

PROCEDURAL TASKS AND QUALITY CONTROL

During the planning phase, the Coordinating Center designed the data collection forms and prepared the Manual of Operations. At the end of the planning phase, training sessions for the project staff were held. A quality assurance protocol is critical for a multi-site study to ensure uniformly high standards of data collection and reporting. Collection of high quality data is enhanced by well-defined data items and well-designed data collection forms administered by carefully-trained personnel. The Coordinating Center plays a key role in developing a quality assurance program that includes development of on-going study-wide quality assurance activities.

Design of Data Collection Forms
Data collection forms were developed in accordance with protocol design. Brief instructions and definitions are given on the forms themselves. More comprehensive instructions and definitions are documented in the Manual of Operations. Forms were pilot tested so that unclear, difficult, and nonessential items could be modified or eliminated, and missing essential items could be added. Despite careful planning, the forms may need to be revised during the first months of data collection. Thus, each page of the paper data collection forms has a version number and date and the corresponding electronic data forms have been given version numbers.

The information recorded and collected on standard study forms is the database for the analyses. The format of answers is kept consistent within any one form and across all forms to ensure their interpretability. Forms are laid out logically and clearly to promote
accurate completion. Also, the electronic data forms have been created to mimic the paper forms for easy, accurate data entry

Site Visits
Site visits will serve as important quality control checks. The Coordinating Center will send two staff members to each of the five intervention sites twice in the course of the study. The quality of the site's data will be assessed and the certification status of the interviewers, data manager, and data entry personnel will be reviewed. Also, the site's administrative/ organizational structure, the interview procedure, and if possible, an intervention will be reviewed.

Reporting
During the course of this study several types of technical and statistical reports will be prepared. Monthly reports will be generated on recruitment activities at each clinical site. More comprehensive reports will be provided to the Data Safety and Monitoring Board (DSMB) and Steering Committee. These reports will be presented at the periodic meetings to include: 1) Patient recruitment and follow-up by site, 2) Quality control reports that include the timeliness, completeness, and quality of data received from the sites, and 3) Participant follow-up adherence data that include missed visits and attrition 4) Descriptions of the study sample overall and by race/ethnic identity 5) Safety reports (adverse events, serious adverse events).

Annual progress reports and a final summary will be delivered to NIH. At the end of the project, computer disks and/or tapes with well-organized files and documentation of all study data will be prepared and delivered to the NIH (pending approval of NIH, data analysis will be made available to other investigators outside of the REACH project).

STATISTICAL ANALYSIS

The REACH II intervention is designed to target 5 domains. Consequently, a multivariate outcome will be employed to assess the efficacy of the intervention. The domains targeted by the intervention are described here:

1. **Depression**. Using 10-items from the original Centers for Epidemiologic Studies Depression Scale (CES-D, Radloff, 1977). Scores on the 10-item CES-D range from 0 through 30, with higher scores indicating increased presence of depressive symptoms.

2. **Caregiver Burden**. Using a modified version of the Burden Interview (BI, the description of the original version can be found in Zarit et al., 1985; the modified version is described in Bedard, et al., 2001). The modified BI consists of 12 questions, however since one of the questions is not appropriate for caregivers of care recipients who are institutionalized [Do you feel that you don't have as much privacy as you would like because of (CR)?], the REACH BI score will be based on 11 questions. Scores for the REACH BI measure range from 0 through 44, with higher values suggesting greater levels of caregiver burden.

3. **Self-Care**. The caregiver's diligence in looking after his/her health will be assessed through 11 questions. Examples of the self care questions include: whether the caregiver has had an eye examination in the past year and whether the caregiver has had his/her blood pressure checked. Self-care scores range from 0 through 11. Higher scores are suggestive of increased attention to one's health and well-being.

4. **Social Support**. Social support data will be collected using a hybrid instrument containing items from several established instruments [Inventory of Socially Supportive Behaviors (Barrera, Sandler, & Ramsey, 1981); Lubben Social Network Index (Lubben, 1988); Satisfaction with Support and Negative Interactions (Krause, 1995; Krause & Markides, 1990)]. Social support questions can be categorized into 4 specific domains: 1) Received support, 2) Satisfaction with support, 3) Social network, and 4) Negative interactions. Total social support scores will be computed by summing scores across the categories (after recoding responses so that all scores are in the same direction). Social support scores will range from 0 through 40, with higher scores indicating increased social support.

5. **Problem Behaviors**. Three questions, developed by REACH II investigators and corresponding to the primary factors from the Revised Memory and Behavior Problem Checklist (i.e., memory, depression, and disruption; Teri et al., 1992), will be used to assess change in problem behaviors exhibited by the care recipient in the time from baseline to the 6-month follow-up interview. These questions will be scored on 5-point scales ranging from 1 (substantial improvement) through 5 (substantial decline). Thus, total scores for the three questions range from 3 to 15, with higher scores indicating increased decline.

We will use a Generalized Least Squares (GLS) test (O'Brien,1984) to compare the outcomes between the two treatment groups. The GLS test takes into account the correlations among the domains comprising the multivariate outcome and tests the alternative that the active treatment is uniformly superior to the control condition. GLS can therefore be distinguished from other multivariate methods, such as Hotelling's T^2, which only indicate whether a difference exists between the study groups, and not the direction of difference. If the overall statistic from the GLS test is significant (indicating evidence that the treatment is superior to the control), post-hoc analyses will be performed to examine the individual outcome measures comprising the endpoint. The post-hoc analyses will be carried out using a step-down approach (Lechmacher et al., 1991).

Separate GLS tests will be performed for each racial/ethnic identity group. We do not intend to adjust for covariates in these analyses since we are randomizing participants to equalize the groups with respect to important measured and unmeasured covariates.

We will perform subset analyses to determine the efficacy of the intervention within various sub-groups. Examples of sub-groups that will be examined include groups defined by relationship (i.e., spouse and non-spouse), by income level, and by baseline level of the 5 domains of interest (e.g., we will examine efficacy in caregivers who had high levels of depressive symptoms at baseline).

INTERIM MONITORING

In a clinical trial such as REACH II data may be examined during the course of the study to help determine whether the study should continue or be terminated early. One reason for ending a trial prematurely occurs when one treatment arm is out-performing the other by such a large degree that it would be unethical to enroll participants in the weaker arm. For example, if it is found in REACH II that the active intervention is vastly superior to the control condition, then it would be inappropriate to assign more caregivers to the control condition. Another reason for stopping a trial early occurs when the treatment arms are equally efficacious (or equally non-efficacious) and continuing the trial to completion would not change the ultimate findings and interpretation of the study. For example, the REACH II study might be terminated early if interim analyses reveal that the active intervention and control group produce comparable results, both statistically and clinically, and there is no chance that the results would change if the study were carried out to completion.

A question to answer regarding interim analyses is the number of times during the course of the study that the data should be examined. Given that we do not anticipate a huge difference in outcomes in the two groups, and that the active intervention is relatively benign, we propose one interim analysis within each racial/ethnic group. This interim look at the data will take place when one-half of the expected number of patients have their outcomes assessed.

The goal of REACH II is to enroll 200 persons within each racial/ethnic group. We assume that 15% of those who are enrolled in the study will not be included in the final analyses due to drop-outs, caregivers lost-to-follow-up, etc. This estimate is based upon data from the REACH I study. After removing these 15% from the 200 persons whom we will be recruiting, 170 participants in each racial/ethnic group remain for inclusion in the final analyses.

The interim monitoring will therefore occur when the outcomes measures are available for approximately 85 participants (actually, 86 participants or approximately 43 per treatment arm). Assuming that the rate of recruitment is constant over the 19-month period, we expect to have 86 participants in an ethnic group between the 15^{th} and 16^{th} months of recruitment (corresponding to calendar months of August and September 2003, respectively). If it is decided to terminate the trial early after reviewing the results of the interim analysis, then roughly 168 total caregivers would have been recruited within each racial/ethnic group, as opposed to the original projection of 200 caregivers.

Ideally, recruitment will be similar across the racial/ethnic group. This will permit us to perform all 3 interim analyses (one for each racial/ethnic group) at the same calendar time. However, if recruitment varies across racial/ethnic groups, the interim analyses will be performed at separate calendar times.

Adjusting for multiple examinations of the data: When the data from a trial are analyzed multiple times, the alpha error rate, or probability of erroneously rejecting the null hypothesis that the outcomes in the two treatment groups are the same, increases. *A priori*, we specified an alpha error rate of 0.0167 for each racial/ethnic group in REACH II. Note that this alpha level is based on a Bonferroni adjustment of the traditional 0.05 alpha level [i.e., 0.05 divided by 3 (the # of racial/ethnic groups) = 0.0167]. Since we are doing two examinations of the data, one interim and one final look, we also need to distribute the error rate across the two analysis time points. In addition, we want to give more weight to the final analysis which is based on all patients.

We propose to use the O'Brien & Fleming approach to handle adjustment for multiple examinations of the data during the course of the study (O'Brien & Fleming, 1979). In this method, the significance level required to reject the null hypothesis at each look at the data increases as the study progresses. With a single interim look, the type I error at the initial look, according to the O'Brien-Fleming approach is 0.0007. Thus, if the p-value corresponding to the GLS test is less than 0.0007 for the interim analysis, it will be concluded that the treatment arms are significantly different and a decision should be made in regards to continuing the trial. Note that with a significance level as small as 0.0007, the differences

between the treatment arms would have to be very large to find a significant difference. This is precisely what we want – that is, to adopt a conservative approach for the interim analyses since not all the data are available.

Conversely, we want to be relatively less conservative for the final analysis. The significance level for the final analysis is 0.0165. So despite the fact that we adjusted the type I error for the interim analysis, the alpha value for the final examination is close to the initial 0.0167 level.

EFFECT SIZE DETERMINATION

As the sample size is fixed for this study, we sought to determine the effect size, or the detectable difference, between the treatment and control groups. To calculate the effect size, we performed a series of simulations by creating 1,000 samples, containing 86 (43 control, 43 active participants) and 170 persons (85 control, 85 active participants) for the interim and final analyses, respectively. We assumed for these simulations that the 5 measures comprising the outcome follow a multivariate normal distribution and the variance-covariance matrix was the same for the active intervention and control groups. The variance-covariance matrix was estimated using data from the control groups from the REACH I study. None of the correlations between the measures that were used in REACH I were greater than 0.25 so we assumed that the correlations between the measures would all be 0.25. Then, we found the effect size that would yield approximately 80% of the simulated experiments, rejecting the null hypothesis at $\alpha=0.0007$ and $\alpha=0.0165$ for the interim and final analysis, respectively. These simulations showed that we would be able to detect an effect size of 0.50 and 0.25 for the interim and final analyses, respectively, assuming that the intervention had an equal effect on all components of the multivariate outcome.

REFERENCES

Barrera, M., Sandler, I., & Ramsey, T. Preliminary development of a scale of social support: Studies on college students. American Journal of Community Psychology 1981, 9, 435-447.

Bedard M, Molloy DW, Squire L, et al. The Zarit Burden Interview: a short version and screening version. The Gerontologist 2001; 41:652-657.

Krause, N. Negative interaction and satisfaction with social support among older adults. Journal of Gerontology: Psychological Sciences 1995, 50B, 59-73.

Krause, N., & Markides, K. Measuring social support among older adults. International Journal of Aging and Human Development 1990, 30, 37-53.

Jennison C, Turnbull BW. Group sequential methods with applications to clinical trials. (Chapter 15). Chapman & Hall/CRC, 2000.

Lehmacher W., Wassmer G., Reitmeir P. Procedures for two-sample comparisons with multiple endpoints controlling the experimentwise error rate. Biometrics 1991; 47:511-521.

Lubben, J.E. Assessing social neetworks among elderly populations. Family Community Health 1988, 11, 42-52.

O'Brien PC. Procedures for comparing samples with multiple endpoints. Biometrics 1984; 40:1079-1087.

O'Brien PC, Fleming TR. A multiple testing procedure for clinical trials. Biometrics 1979; 35:549-556.

Pocock SJ, Geller NL, Tsiatis AA. The analysis of multiple endpoints in clinical trials. Biometrics 1987; 42:487-498.

Radloff LS. The CES-D scale: a self-report depression scale for research in the general population. Applied Psychological Measurement 1977; 1:385-401.

Tang DI, Gnecco C, Geller NL. Design of group sequential clinical trials with multiple endpoints. Journal of the American Statistical Association 1989; 84:776-779.

Teri L, Truax P, Logsdon R, Uomoto J, Zarit S, Vitaliano PP. Assessment of behavioral problems in dementia: the revised memory and behavior problems checklist. Psychology and Aging. 1992;7:622-31.

Zarit SH, Orr NK, Zarit JM. The hidden victims of Alzheimer's disease: families under stress. New York: New York University Press, 1985.