

3.0 Data Management Preparations

This chapter describes our preparations for managing data for the 1999 NLTCs Functional and Health Changes of the Elderly Supplemental Studies. These preparations included programming the computer-assisted telephone interviews (CATI), electronic appointment scheduler, and appointment and reminder calls; implementing the control systems; and producing status reports for all data collection activities. In addition, a Geographic Information System (GIS) was developed for the Venipuncture Study in order to coordinate assignment of sample members to specific phlebotomists.

3.1 Preparation of CATI Interviews

3.1.1 Programming the Interview

The first step in any CATI development process is to design the instruments to be programmed. For the Kin and NOK surveys, Dr. Larry Corder of Duke provided the 1994 NLTCs instrument. The NOK document had been previously programmed by RTI for the conduct of the pretest in 1997. However, due to software updates, both documents had to be programmed from scratch. Prior to programming the Kin and NOK documents, the design specifications were edited to a format which expedited programming. During this process, we worked with Duke to resolve any remaining design issues and modified the design documents accordingly. Both instruments were programmed in CASES, Version 4.3.

A third instrument was also programmed for the Venipuncture study. Duke requested that RTI collect additional information from participants regarding their places of residence for the first twelve years of life as well as the birth and death dates of their parents. RTI drafted a document that was approved by Duke prior to programming. This questionnaire was also programmed in CASES and was administered during the phlebotomy appointment call.

3.1.2 Programming the Appointment Scheduler

RTI Research Computing Division (RCD) programmers developed the electronic phlebotomy appointment scheduler at the request of and with the input from the NLTCs project staff. This endeavor was a first for RCD. The appointment scheduler was developed and programmed in Visual Basic, and the scheduled appointments were output to an Access table. Phlebotomy appointments were scheduled by site, Monday through Thursday, in two-hour intervals to allow for driving time between appointments. During the interview, the Telephone Interviewer (TI) had to switch between running the CATI and the appointment scheduler, since CASES and Visual Basic are not compatible and do not interface with each other. The scheduler did not allow TIs to double-book appointments. For example, if an appointment was made for a nurse in site 110 at 9:00 a.m., the next available appointment would be at 11:00 a.m. TIs manually keyed in the respondent's zip code to access the schedule of the nurse for that area.

The scheduler output appointments to an Access table on a Telephone Survey Department (TSD) server. That table was then linked to the Venipuncture control system so that appointments could be viewed and modified by project staff outside of TSD. In that way, a respondent could call a member of the project staff to change an appointment, and the appointment scheduler would automatically be updated reflecting the change. TSD staff did not have to be notified that the appointment had changed.

An output page of the appointment scheduler is displayed in *Exhibit 3-1*. The scheduler contains the site number, appointment date and time, respondent ID number, name, address, phone number and comment column (not shown). The comment column allowed TIs to document information that would be helpful to the nurses, such as driving directions or a medical history of diabetes or rolling veins. Many respondents wanted the nurses to call them prior to the appointment, and this request was also noted in the comment column.

3.1.3 Testing

Using the design instruments as guides, the three CATI instruments and appointment scheduler were tested by both the programmers and project survey staff. Initially, the CATI and appointment scheduler were tested independently for the Venipuncture study. This was particularly important since this type of appointment scheduler had never been used before and because the two interview components ran separately but were accessed during the same interview.

During testing, two types of problems can be discovered: programming problems and design problems. Programming problems identified by the programmer in the NLTCs instruments were fixed by the programmer. Programming problems identified by the project staff were documented in email messages and fixed by the programmer. As design problems were identified, the problems were resolved, and all design changes were reprogrammed and tested.

3.1.4 Development of the Control System

The 1999 NLTCs Supplemental Studies project team worked with members of RCD to develop the control systems. These systems were developed using Visual Basic and Microsoft Access. Each study-specific control system was utilized to update the status code of sample members. In addition, the Venipuncture, Buccal Cell and Kin control systems were used to monitor the delivery of buccal cell kits to sample members, document the receipt of buccal specimens to the lab, and record the receipt of consent forms at RTI. The Venipuncture control system was also utilized to document the training of nurses and the receipt and return of phlebotomy supplies, set blood draw appointments, document the receipt of blood specimens by the lab, and record the receipt of the various phlebotomy forms, including the consent form and incentive receipt.

Exhibit 3-1. Phlebotomy Appointment Scheduler

CONFIDENTIAL DATA REMOVED

The 1999 NLTCS control systems were modified by the RTI Data Manager from the standard package control system developed by RCD. For each study, additional project-specific event codes were added in order to monitor sample members progress in the study. For example, we could track the date a buccal cell kit was sent from RTI in order to schedule the reminder call if the specimen had not been received at the lab. *Appendix C* contains the event codes for each control system. The tables list each event code by name and number.

For each status code, a data quality check was implemented to ensure that sample members moved between status codes in the intended fashion. The third column of the tables in *Appendix C* (mustprecede) demonstrates that mechanism. The status code listed in the third column must have preceded the new status code in order for the update to be successful. If a respondent was moved from one status code to another such that the mustprecede rule was broken, an error message was generated alerting the user to the problem.

For each study, the Status Frequency Report and the Status Percentages Report were utilized to provide progress reports to the Duke Senior Investigator. The Status Frequency Report was a standard field status report developed by RCD that contained a list of the total number of sample members by event code. The Status Percentages Report was developed by the RTI Data Manager. The report listed the percentage of sample members in each status code. The Status Percentages Report was written in Visual Basic and was denominator driven. Denominators were determined based on eligibility criteria. Sample members were deemed eligible or ineligible based on data obtained from the history and preload tables. A sample member was ineligible for participation if he/she could not be located or was deceased, incapacitated, hearing impaired or if he/she spoke a language other than English. The final Status Frequency Reports and final Status Percentages Reports are included in **Chapter 6 Results and Recommendations**.

3.2 Development of the GIS System

A Geographic Information System (GIS) was utilized by the Venipuncture project management staff to analyze the national and regional distribution of primary sampling units (PSUs), examine the distribution of patients and phlebotomists, assign phlebotomists appropriately, and evaluate response rates at the local, state and regional levels. A GIS is a geographic database management system that provides users with a powerful set of tools for viewing and analyzing geographic data and performing spatial analysis. The geographic data consist of a series of spatially-referenced map layers that contain information about features that are located in specific locations. These features might include state, county, zip code and PSU boundaries; locations of survey respondents and field offices; demographic and socioeconomic data; and survey response rates by PSU or other geographic unit.

GIS functions can be extremely useful to the decision-making and program administration responsibilities of survey and epidemiology projects. The most commonly used function of a GIS is map production. Maps provide the viewer with a visual picture of the distribution of features across a geographic area and the relationships among them. Other commonly-used GIS functions include geocoding, distance computations, spatial queries, buffering techniques, and overlay analysis. Geocoding is the process of linking a record in a non-spatial database to a geographic feature or location in the map layer through a geographic identifier such as state, county, zip code, census tract, or street address.

Three major tasks were involved in the implementation of an Internet Map Server (IMS) for the Venipuncture Project: 1) acquire and develop spatial data for mapping and analysis; 2) develop and deploy the IMS application for program managers and analysts; and 3) maintain the IMS application and update GIS databases with information from surveys and other sources.

Task 1: Acquire and develop spatial data for mapping and analysis. This was the most time-consuming phase of the IMS development. Geographic (i.e. map) data were developed for all features to be represented in the IMS. Most of these map data were available from existing sources, but some entailed geocoding and additional development. **Exhibit 3-2** lists all feature types used in the Venipuncture Project Management IMS and their availability, source, and required development processes.

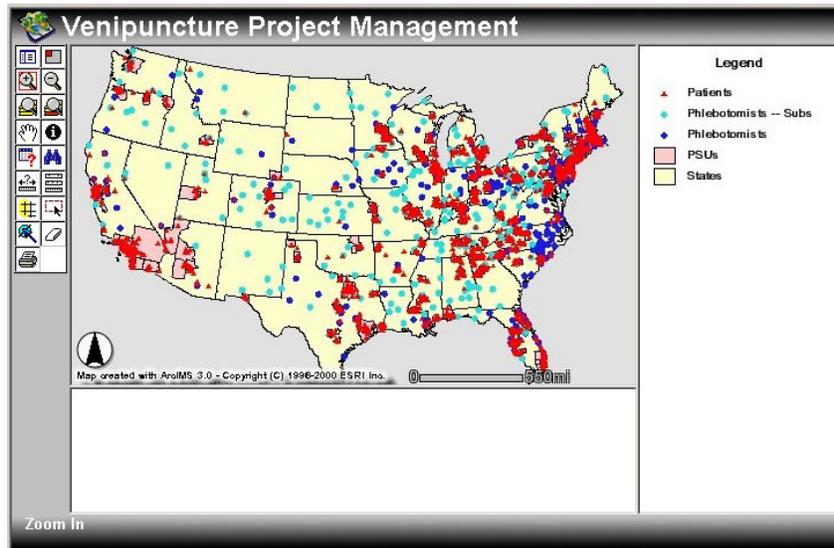
Exhibit 3-2. Venipuncture IMS Features

Feature	Available	Source OR Development Process
State boundaries	YES	ESRI Data & Maps 2000*
County boundaries	YES	ESRI Data & Maps 2000
Zip code boundaries	YES	ESRI Data & Maps 2000
Primary Sampling Units (PSUs)	NO	Compiled from list of county-PSU associations. Most PSUs were aggregates of counties.
Survey respondents	NO	Point locations were obtained by address geocoding to street network.
Field offices and phlebotomist locations	NO	Point locations were obtained by address geocoding to street network.

* Data provided by Environmental Systems Research Institute (ESRI), Redlands, CA, producers of Arc/Info, ArcGIS, ArcView and ArcIMS software.

ESRI's Arc/Info and ArcView software were used to develop initial geographic data over a four-week time period during Fall 2000. **Exhibit 3-3** shows several of the map layers that were acquired or developed for the project. Address data for patients and field offices were sent to Geographic Data Technology, Inc. (GDT) for processing. GDT matched these addresses to street network databases and provided latitude and longitude coordinates for each patient and field office. GIS software was then used to create patient and field office map layers. Each sample member was represented by a red triangle, an Interim franchise office was indicated with a blue dot, and a contract office was represented by an aqua dot.

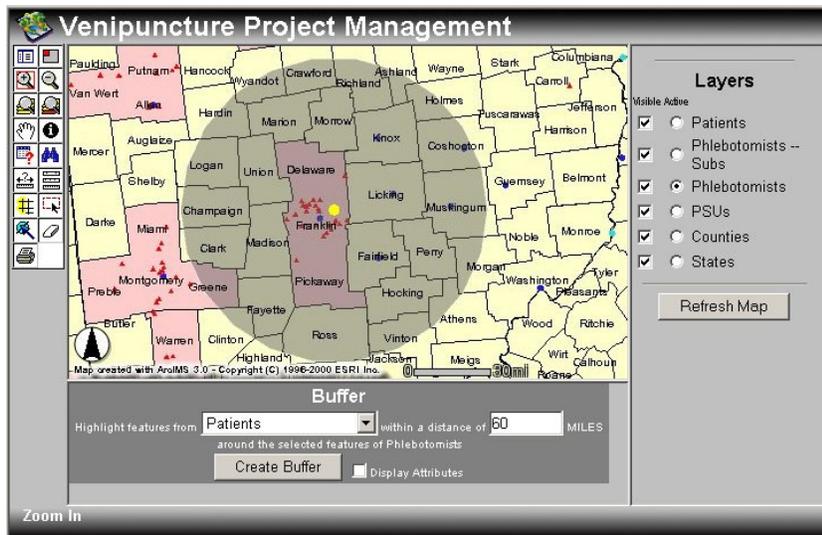
Exhibit 3-3. Map layers developed for the Venipuncture Project Management IMS



Task 2: Develop and deploy the IMS application for program administrators and analysts. After the development of feature map layers, an internet map server application was developed with Arc/IMS software (ESRI, Inc., Redlands, CA). Out-of-the-box IMS functions include mapping; display functions such as zooming in or out, panning, and symbol selection; spatial queries; feature selection; distance measurement; and buffering.

The Venipuncture Project Management IMS application was developed by RTI's GIS staff and installed on RTI's GIS map server during Fall 2000. Logins and passwords were provided to program administrators and analysts, who were then able to access the application via their web browser. Users were able to turn map layers on and off, perform spatial queries, identify statistical attributes of geographic features, and determine distances among features (e.g. from field offices to patients). *Exhibit 3-4* shows how a buffer function in the IMS application was used to locate all patients within 60 miles of a selected field office. This feature was important for the allocation of phlebotomists to patients.

Exhibit 3-4. Study Participants within 60 Miles of Field Office



Task 3: Maintain IMS application and update GIS databases with information from surveys and other sources. The Venipuncture Project Management IMS has been maintained continuously since Fall 2000. For the most part, maintenance consists of updating the ArcIMS software when new versions are released (the current version is 3.1) and restarting the map server after operating system upgrades and routine maintenance. During Fall 2001, however, significant upgrades were made to geographic feature data. A zip code-based map layer of field office services areas was developed to show completion rates and refusals for the venipuncture study and three of the National Long Term Care Survey (NLTCS) Supplemental Studies combined: Venipuncture, Buccal Cell and Kin. **Exhibit 3-5** shows completion rates for the Venipuncture study in the Upper Midwest.

Exhibit 3-5. Completion Rates for the Venipuncture Study in the Upper Midwest

