DESCRIPTION OF DATA FROM THE NATIONAL ANALGESIC NEPHROPATHY STUDY (NANS) INCLUDED IN THE NIDDK REPOSITORY

Prepared by the Slone Epidemiology Center of Boston University

September 2004

The NANS data in the NIDDK Repository consists of three parts: an ACCESS database, SAS save files, and CT scan images. Each of these is located in a separate folder on the DVD, and is described below. More detailed documentation about the study, including progress reports and manuals, is posted separately in the Repository.

ACCESS DATABASE

The database is named **NANS_NIH.mdb.** It contains 5 data tables, 5 description tables, and 30 lookup tables.

Data Tables

1. tblInterviewData: Case/Control Interview Data

This is a 1:1 table including data from the subject interview, including demographic information, medical history, beverage consumption history, and a limited set of computed drug exposure variables. The complete medication exposure information and occupation history are contained in separate tables (see below).

2. tblInterviewData_Drugs: Case/Control Interview Data Drugs

This is a 1:many table containing the detailed data on medication exposures.

3. tblInterviewData_Occupations: Case/Control Interview Data Occupations

This is a 1:many table containing the information on occupation history.

4. tblCTScanData: CT Scan Measurements and assessments

This is a 1:1 table containing the data from the CT scan, except information on renal masses (see below).

5. tblCTScanData_Masses: CT Scan details of masses

This is a 1:many table including information on renal masses. The size of renal masses is given in centimeters.

Description Tables

For each of the above tables there is a corresponding table that begins with the name tblVarDesc… and includes detailed information about the fields in the table. For example, tblVarDescInterviewData_Drugs contains the order of the field in the table, the name of the field, a description of the field, the name of the lookup table for that field, the interview question that produced the response, the section of the interview that contained the question, and the variable type.

Lookup Tables

Tables beginning with the prefix tlkp... include the variable codes and their values.

Database Documentation

The following separate files contain detailed information about the database:

- TablesRelationship_Diagram.pdf
- TableList_Descriptions.pdf

- TableFieldList_Descriptions.pdf
- LookupTables_AssociatedTables.pdf

General Information

The interview data tables include information on 240 ESRD cases and 206 matched controls included (34 ESRD cases did not have a matched control). The variable **IntvwID** is the subject interview identification number. It can be used to link the above tables. The CT scan tables include 221 ESRD cases and 61 healthy subjects. The variable **StudyIDN** is a unique identification number assigned to each subject; ESRD subjects also have a value in the field **IntvwID**. The healthy subjects were volunteers on whom CT scans were performed in order to get baseline renal size information; they were not matched to the ESRD cases, nor were they interviewed. The controls were matched to the cases and interviewed; they did not have CT scans.

The beginning of the study was referred to as Phase I. The interview was revised after this initial data collection period. The remainder of the study was referred to as Phase II. The variable **phase** in tblInterviewData is coded as 1 or 2 to denote the period during which the subject was enrolled.

Deletions from the interview after Phase I include:

- A question about whether diabetes was treated with insulin or pills (**InsOrPil**) had been followed by specifying whether it was pills (**DiabPill**) or insulin (**Insulin**). The specification was dropped.
- A global question about any immediate family member with diabetes, hypertension or end stage renal disease (**FamDis**) was dropped.

- Consumption of decaffeinated coffee was dropped.

Additions to the interview after Phase I include:

- A follow-up question added after swelling of the legs: due to kidney disease (**SwelKid**)?
- For subjects younger than 50 years, questions about nocturia before starting dialysis (**Noct**) and year (or age) or first occurrence (**NoctAge**).
- Questions about prior renal disease were found to be poorly understood in Phase I. The questions were simplified in Phase II to include a question about history of obstructed kidney (**ObstrKid**) and kidney infection (**Pyelonep**).
- A follow-up question added after bladder infection: did you have more than 5 episodes (**UTIEpis**)?
- For history of diabetes, gout, and hypertension, the subject was given the option of providing age or year of diagnosis.
- For beverage consumption history, the period of interest was changed from before dialysis (Phase I) to ten years ago (Phase II). Consumption was recorded separately for hot tea and iced tea, and consumption of colas and Dr. Pepper was added.
- For occupation history, instead of obtaining a complete list of occupations (Phase I), yes/no questions were asked for selected occupations that may involve exposures potentially related to renal disease (Phase II), e.g. furniture, textile. If an occupation response from Phase I fit one of the Phase II prompted categories, it was assigned the appropriate value for the question.

A limited set of drug exposure variables can be found in tblInterviewData. These include an indicator for any exposure before the index year (nine years prior to first dialysis for ESRD cases; nine years prior to date of interview for controls) for acetaminophen, aspirin, ibuprofen,

and a composite variable (i.e. exposure to acetaminophen or phenacetin or aspirin or other salicylates), and a number of measures of dosage for the same groups. For the latter, in instances where the patient did not specify the dose per pill of a single ingredient product, we assumed standard doses of 325 mg for aspirin and acetaminophen, and 200 mg for ibuprofen. Aspirin reported as "low dose" was assumed to be 81 mg. When a patient took multiple drugs in an exposure group (e.g. acetaminophen-containing drugs), the kilograms for all drugs were summed. For the composite group, the total kilograms of all relevant ingredients (aspirin, other salicylates, acetaminophen, phenacetin) were summed. For multicomponent drugs, dose was assumed based on the amounts of each component in the product during the relevant years of use (e.g., according the OTC Handbook, the amount of acetaminophen in Excedrin was 97 mg in 1979 and 250 mg in 1982). Cumulative kilograms of use before the index year are provided for acetaminophen (**apapkgb1**), aspirin (**asakgb1**), other salicylates (**saldose**), phenacetin (**pctdose**), and the composite (**bigdose1**). Average kilograms per year of use is provided for acetaminophen (**tdensity**), aspirin (**adensity**), and the composite (**bigdens**). In these data, the subject with IntvwID 10028 is classified as unknown for both **bigdose1** and **bigdens** because the dose of phenacetin was unknown; in our analysis, we categorized this subject as having a bigdose >3kg because his aspirin exposure alone exceeded that amount.

SAS FILES

Five SAS files have been created by exporting the five ACCESS data tables:

- tblCTScanData.sas7bdat
- tblCTScanData_Masses.sas7bdat,
- tblInterviewData.sas7bdat,

- tblInterviewData_Drugs.sas7bdat
- tblInterviewData_Occupations.sas7bdat

A SAS program (Procformat.sas) includes the format, variable names and value labels for each of the files.

CT SCAN IMAGES

The NANS CT Scan images on the DVD are the anonymized versions of the kidney scans collected for the NANS study. All patient identifiers and hospital names have been removed for confidentiality. The images are segregated into two folders representing the two patient types, ESRD and Normal. These folders contain subfolders named after the patient's CT Scan ID. In these subfolders are each patient's image files in dicom format. There are 207 ESRD image sets and 26 Normal image sets (8,245 files, 3.28GB). Digital images were not available for 14 ESRD patients and 35 normal subjects.