Data Archive for
National Analgesic Nephropathy Study (NANS)

Note. Sections of this document are excerpted from a document prepared by the DCC. The citation for that document is: Slone Epidemiology Center of Boston University, 2004 (see below for details[1]).

The National Analgesic Nephropathy Study (NANS) was funded by NIDDK to: (1) study the relationship between analgesic use and end stage renal disease; and (2) to learn if non-contrasted CT scans are able to detect, as a unique entity, analgesic nephropathy.

The NANS archive contains:

1. Study Documentation
2. Data Collection Forms
3. Data
4. CT Scans --make special request for this data
5. Data Set Integrity Check

Below we describe the contents of the other NANS directories.

1. DOCUMENTATION DIRECTORY

NANS was conducted in two phases. A pilot phase to test procedures is referred to as Phase I. The full study implementation is referred to as Phase II.

The Documentation directory contains the Protocols and interviewer manuals for both Phases I and II. It also includes a 5-page document that includes summary information about the NANS study, e.g., investigators, sites, contact information, sample sizes, etc.

[1] Slone Epidemiology Center of Boston University, Description of data from the National Analgesic Nephropathy Study (NANS) included in the NIDDK repository. Boston: Slone Epidemiology Center of Boston University, September 2004.

2. DATA COLLECTION FORMS.

The Data_Collection_Forms directory contains "forms" for: (1) the interviews conducted in Phase I, (2) the Phase II interviews conducted with cases, (3) the Phase II interviews conducted with controls, and (4) the evaluation form used to obtain data from readings of the CT scans. It should be noted that the NANS data collection used computerized forms; the directory contains *printed* renderings of the interview questions appearing on these computerized forms. These are **not** images of the computer screens used in data collection.

3. DATA DIRECTORY[2]

General Information.

NANS interview data include information on 240 End-Stage Renal Disease (ESRD) cases and 206 matched controls. (34 ESRD cases did not have a matched control.) The variable IntvwID is the subject interview identification number. It can be used to link records across tables and files. The CT scan included 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 pilot phase of the NANS study is referred to as Phase I. The interview was revised after this initial data collection period. The subsequent full study is 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 and additions to study questionnaires between Phases I and II are detailed in Appendix A.

Access Database

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

[2] The following sections are excerpted from a document prepared by the NANS Data Coordinating Center; see title page.

Data Tables

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.

tblInterviewData_Drugs: Case/Control Interview Data Drugs
This is a 1:many table containing the detailed data on medication exposures (see below).

tblInterviewData_Occupations: Case/Control Interview Data Occupations
This is a 1:many table containing the information on occupation history.

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).

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.

Drug Exposure Measurements

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 to 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).[3]

ACCESS Database Documentation

The following files included in the directory "...\DATA\AccessDatabase\Documentation" contain detailed information about the ACCESS database:

- TablesRelationship_Diagram.pdf
- TableList_Descriptions.pdf
- TableFieldList_Descriptions.pdf
- LookupTables_AssociatedTables.pdf

SAS FILES

Five SAS files were 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.

Appendix B displays the number of variables, and number of cases included in these SAS files.

[3] In these data, the subject with IntvwID 10028 was classified as unknown for both bigdose1 and bigdens because the dose of phenacetin was unknown; NANS analysts have categorized this subject as having a bigdose1 > 3kg because his aspirin exposure alone exceeded that amount.

CT SCAN IMAGES

The scan images included in the directory CT_ScanImages are 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.[4]

DATA SET INTEGRITY CHECK

As a partial check of the integrity of NANS Study data set archived in the NIDDK data repository, a set of tabulations was performed to verify that published results can be reproduced using the archived data set. Analyses were performed to duplicate selected results for the data published by William Henrich et al. in Non-Contrast Enhanced Computerized Tomography and Analgesic-Related Kidney Disease: Report of the National Analgesic Nephropathy Study; Journal of the American Society of Nephrology May 2006 17: 1472-1480. Complete results are presented in the DSIC folder of the Official Archive.

[4] The DCC documentation says that scan images are available for 207 ESRD subjects and 26 Normal subjects. The data file has data for 282 unique subject IDs.

APPENDIX A[5]

Deletions and Additions to Study Questionnaires between Phase I and Phase II

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.

[5] This appendix is excerpted from a document prepared by the NANS Data Coordinating Center; see title page. This appendix is excerpted from a document prepared by the NANS Data Coordinating Center; see title page.

- 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.