

NATIONAL ANALGESIC NEPHROPATHY STUDY

PROTOCOL – PHASE II

Coordinating Investigator

William L. Henrich, M.D.

Principal Investigator

Joseph I. Shapiro, M.D.

Co-Investigators

David W. Kaufman, Sc.D.

Samuel Shapiro, M.B., FRCP(E)

Collaborating Investigators

Vardaman Buckalew, M.D.

Richard Clark, M.D.

William Finn, M.D.

Balakuntalam Kasinath, M.D.

George A. Porter, M.D.

Sharon Anderson, M.D.

Table of Contents

	<u>Page</u>
1. Introduction	2
2. Background and Objectives of NANS	2
3. Study Organization	3
4. Summary of Phase II Activities	5
5. Timetable	5
6. Summary of Activities Completed During Phase I	6
7. Study Network	7
8. Subject Enrollment	9
8.1 ESRD Patients	9
8.2 Control Subjects	11
9. CT Scans	12
9.1 Methods of CT Scan Reading	12
9.2 Technical Issues of CT Scanning	12
9.2.1 Type of Scanner	13
9.2.2 Patient Preparation	13
9.2.3 Scan Parameters	13
9.2.4 Display Parameters	13
9.3 Radiological Data to Be Analyzed from Non-Contrasted CT Scans	14
9.3.1 Renal Volume	14
9.3.2 Indentations	15
9.3.3 Calcifications	16
9.3.4 Other Observations	17
9.4 CT Scan Diagnostic Criteria for AN	17
9.5 Recording CT Information for Review	19
10. Interview	20
10.1 Questionnaire	20
10.2 Validation of the Questionnaire	25
11. Data Management	27
11.1 Quality Control	27
11.2 Data Handling Procedures	28
12. Data Analysis	30
12.1 General Issues	30
12.2 Relative Risk Analysis	32
12.3 Sensitivity and Specificity Analysis	35
13. Study Design Issues	37
13.1 Selection Bias	37
13.2 Information Bias	38
13.3 Misclassification	39
13.4 Confounding	40
14. Conclusion	41
15. References	42

1. Introduction

In October 1999, the contract for the National Analgesic Nephropathy Study (NANS) was awarded by NIDDK to the Medical College of Ohio. This document is the protocol for Phase II; the study methods and procedures were developed, tested, and finalized during Phase I. The proposed methods and their rationale are described in detail in the original Technical Proposal and further elaborated on in our response to questions of January 11, 1999. Although much of what appears here will therefore be a recapitulation and expansion of the proposal, reference is made throughout to modifications from the original and Phase I protocols.

2. Background and Objectives of NANS

Based mainly on clinical evidence, supported by experimental studies and correlational data, heavy analgesic use has been implicated in the etiology of a unique nephropathy characterized by renal papillary necrosis and interstitial nephritis (analgesic nephropathy, or AN). Excessive use of analgesics is also a suspected cause of other forms of chronic renal disease. There have been a number of formal epidemiologic studies of these topics, but they have been unable to delineate the precise relationship of analgesic intake to chronic renal insufficiency because of flaws in their design. The etiology of chronic renal disease is still poorly understood, largely because its onset is insidious, with early distinctive pathological changes often obscured as the disease progresses.

A new study is both necessary and timely because (1) the number and type of analgesics in the U.S., either as single ingredient products or as combination products, is large and increasing; (2) the incidence and prevalence of analgesic-related renal disease remains unknown in the U.S.; (3) the types of analgesics linked to the development of renal disease and the

requisite amount of consumption are still unclear; (4) recent evidence has suggested the possibility that the abdominal CT scan may be a useful diagnostic tool for the identification of AN; and (5) much of analgesic-associated renal disease may be preventable.

The specific objectives of the study are as follows:

1. To determine whether non-contrasted CT among patients with end stage renal disease (ESRD) can be used with adequate precision to identify cases of AN as a distinct entity.
2. If AN can be identified, to determine which analgesics, including combination analgesics, are associated with its development, and to determine the effects of duration of exposure and dosage.
3. If AN can be identified, to determine its incidence and the proportion of the total burden of ESRD in the U.S that it represents.
4. Whether or not it is possible to identify AN as a distinct entity by CT scan, to determine the relation of analgesic use (the effects of specific drugs, duration of exposure, and dosage) to the *de novo* onset of any chronic renal insufficiency leading to ESRD.

The study will be conducted in two phases, with a possible third phase depending on the outcome of the first two. The remainder of this protocol refers to Phase II, which will be a 22 month period devoted primarily to enrolling 200 ESRD cases and 200 matched control subjects, refining the CT scan parameters to be used as criteria for AN, and making a determination as to whether AN can be identified by CT scan.

3. Study Organization

The project is being conducted under the overall direction of William L. Henrich, M.D., Professor and Chairman of the Department of Medicine at the University of Maryland School of

Medicine, who is the Coordinating Investigator, and Joseph I. Shapiro, M.D., Professor of Medicine and Pharmacology and Chairman of the Department of Medicine at the Medical College of Ohio, who is the Principal Investigator. The Slone Epidemiology Unit (SEU) of Boston University School of Medicine is the Coordinating Center; David W. Kaufman, Sc.D., Associate Director, and Samuel Shapiro, M.B.,F.R.C.P.(E), Director Emeritus, are Coinvestigators.

In Phase I, the study was conducted under the direction of a nephrologist Collaborating Investigator in each of three centers: Vardaman M. Buckalew, Jr., M.D., Professor of Medicine at the Wake Forest University Baptist Medical Center in Winston-Salem, NC (WFU); William F. Finn, M.D., Professor of Medicine at the University of North Carolina School of Medicine in Chapel Hill, NC (UNC); and George Porter, M.D., Professor of Medicine and former Chairman of the Department of Medicine, Oregon Health Sciences University, Portland, OR (OHSU) and Sharon Anderson, M.D., Professor of Medicine, OHSU. For Phase II, the study will be expanded to Ohio under the direction of Joseph I. Shapiro, M.D. B.S. Kasinath, M.D. (University of Texas San Antonio Health Science Center) participated during Phase I as a consultant, and will continue in that role during Phase II. The radiology component of the study includes review of CT scans by a committee consisting of three radiologists from the University of North Carolina School of Medicine, Chapel Hill, NC, including David Warshauer, M.D., Associate Professor and Wendelin Hayes, M.D., Clinical Associate Professor, and directed by Richard Clark, M.D., F.A.C.R., Professor and Director of GU Radiology at the University of North Carolina School of Medicine, Chapel Hill, NC. The radiology departmental chair at UNC, Dr. Joseph K.T. Lee, will serve as a consultant. Dr. Clark also serves as a Collaborating Investigator.

The Steering Committee for the study consists of Drs. Henrich and J. Shapiro; Drs. Kaufman and S. Shapiro; and from NIDDK, Josephine Briggs, M.D., Lawrence Agodoa, M.D., Paul Kimmel, M.D., Mr. Robert Webber, and Ms. Diane Meeks. An External Advisory Committee with responsibility for general oversight of the study has also been appointed by NIDDK, headed by Vincent Torres, M.D. of the Mayo Clinic.

4. Summary of Phase II Activities

The main goal of Phase II is to evaluate the utility of the noncontrasted spiral CT scan in identifying AN as a unique diagnostic entity. To that end, study activities will involve a comparison of analgesic histories as determined by the questionnaire among patients identified from the CT scan criteria as AN or not AN. The study will be conducted in an area of the country where the putative incidence and prevalence of AN is high (North Carolina) and in two areas where the incidence is thought to be modest to low (Oregon and Ohio). A total of 200 incident ESRD patients beginning dialysis will be enrolled (50 from each of the four centers). We also propose to enroll 200 population control subjects without ESRD; controls will be necessary to fully interpret the results.

5. Timetable

The final investigator meeting of Phase I, in conjunction with a third Executive Committee meeting, was held on November 30, 2000; the External Advisory Committee is scheduled to meeting regarding issues raised during Phase I on January 5, 2001. The timetable for Phase II is shown in Figure 1. The formal start date of this phase of the project is December 1, 2000. Enrollment of ESRD patients for interview and CT scan will begin on that date.

Interviewing will commence in January, after the modified questionnaire has been pilot tested (see Section 10). Interviewer training at the Ohio center is scheduled for early January. Dr. Henrich will visit the individual sites during the months of January or February 2001, meeting with local study staff and appropriate dialysis personnel. All interviewers will meet at the SEU in mid-2001 in order to restandardize study procedures. Members of the SEU study staff will visit the four centers early in 2002. Enrollment of ESRD cases will continue in the four centers through July 2002. Two meetings of the study investigators are allowed for, including one held in conjunction with the annual meeting of the American Society of Nephrology in San Francisco in October 2001. The Radiology Committee will meet at least four times to review data from the ESRD cases; Dr. Henrich will attend one of these meetings in order to observe the review procedures. The final review session will take place in August 2002, followed by an Executive Committee and Investigator meeting. Delivery of the results of Phase II and protocol for Phase III to NIDDK is planned for the end of September, at which time a decision will be made about whether to proceed; at the time the contract was initiated, plans for Phase III were deferred pending the results of Phase II.

6. Summary of Activities Completed During Phase I

The major work in Phase I involved the following: definitions of the subjects to be included were finalized; the logistical structure and methods for identifying and enrolling ESRD cases in three centers were developed; the process for using the CT scan in the study was established; data on normal renal dimensions from CT scans in subjects without ESRD was accumulated; the questionnaire was designed, field tested, and revised; methods of communication were developed, including transfer of interview data and CT scan results; and

procedures for enrolling controls were investigated. Computerized logs for tracking patient enrollment and other relevant software for monitoring the progress of the study were designed and communication links between the coordinating center and the three participating regional clinical centers were established.

A major activity of the SEU during Phase I was development of the interview and computer database. Input regarding modifications to the questionnaire was obtained from all of the investigators and the interviewers. The current questionnaire (see Appendix 2), described below, reflects the agreed upon modifications. Members of the SEU staff have completed the development of computerized database systems for both the interview and the CT scan aspects of the project. The latter is programmed in Microsoft Access. A manual of written instructions for administering the interview has been prepared (see Appendix 3).

Dr. Clark has discussed issues of CT machine compatibility with members of the radiology departments at WFU and OHSU regarding the possibility of digital transfer of CT data from those institutions to the Radiology Committee for their interpretation of the findings. The matter has also been explored with outside investigators currently using these techniques. Although this approach to the management of the radiology data has been recommended by the EAC, its implementation is complicated and expensive; additional resources beyond what is available in the NANS budget would need to be obtained. Furthermore, the use of films printed from the scans has proven to be satisfactory for study purposes.

7. Study Network

Due to lower than expected enrollment during Phase I, subjects will be enrolled during Phase II at the four clinical centers, including the three original centers from Phase I (WFU and

UNC in North Carolina and OHSU in Oregon) and an additional center at MCO in Ohio. Also, as described below, the dialysis networks will be expanded in the existing regions. ESRD patients will be recruited at each center, along with one matched control for each case.

New ESRD patients at WFU will be recruited primarily from seven dialysis units owned by WFU. These units draw patients from a six county area within a 60-mile radius of Winston Salem (Forsyth, Guilford, Davidson, Iredell, Surry, and Stokes). Two of these units are located in Forsyth County, two in Guilford County, and one each in Surry, Davidson, and Iredell Counties. Only three dialysis units in these six counties are not owned by WFU, all in Guilford, but investigators at WFU have a history of successful collaboration with the group of nephrologists who use all three. By recruiting patients from all ten units, the study will have access to essentially all new dialysis patients from a population base of approximately 1,023,570. With the expanded network, there should be sufficient ESRD patients who meet the study criteria to reach the numerical goals of Phase II.

At UNC, new ESRD patients will be recruited primarily from four dialysis units operated by UNC, one unit operated by Durham Dialysis in Durham, and four units in Raleigh, operated by Capital Nephrology and Wake Nephrology, one unit in Rocky Mount, one unit in Wilson, one unit in New Bern, one unit in Fayetteville and one in Wilmington. These units cover Orange, Durham, and Wake counties (the so-called "Research Triangle"), and the contiguous counties of Chatham and Alamance and the southeastern portion of North Carolina. The combined population exceeds one million, which should provide sufficient patients to meet the study goals.

In Oregon, in addition to ESRD patients recruited from the Pacific Northwest Renal Services and its network of ten dialysis units, the Fresenius units and a locally-owned unit, which provide services in the metropolitan Portland and Vancouver area, have been recruited to

expand patient enrollment. These units provide care for 90% of all hemodialysis patients in the area, which should be sufficient to meet the goals of Phase II.

In Ohio, the recruitment of patients will come from the dialysis units of the greater Toledo area, namely those of Dialysis Partners of Northwest Ohio, Innovative, CDC, Arrowhead and Gambro, Inc. We expect that these units will be sufficient to support the recruiting needs of Phase II. However, if recruitment is not sufficient from these units, we have made arrangement to recruit patients from the former Henry Ford Dialysis network (currently owned by Fresenius) which consists of 20 dialysis units in the areas surrounding Detroit and Ann Arbor, MI.

8. Subject Enrollment

8.1 ESRD Patients

In December 2000, enrollment of the 200 subjects with ESRD will begin. The case population will be male and female patients with ESRD who have newly begun dialysis. Noncontrasted CT scans and interviews will be obtained at the local study centers. Patients with acute renal failure will not be included; other exclusions will be patients with diabetes requiring treatment with insulin or oral antidiabetics (unless the diabetes is clearly not related to the initiation or progression of renal disease), polycystic kidney disease, renal transplant failure, biopsy-proven glomerulonephritis, amyloidosis, multiple myeloma, angiographically-proven renal artery stenosis, hereditary nephritis, AIDS nephropathy, acute renal failure that fails to recover, sickle cell disease, and toxic disease secondary to antineoplastic agents. These factors, when present, are overwhelming causes of ESRD. Patients who have had a nephrectomy or who were born with a solitary kidney will also be excluded because the increased renal size in such individuals will prevent the assessment of AN, since kidney dimensions are an important part of

the diagnostic criteria. There will be a lower age limit of 35 years because younger patients are unlikely to have used analgesics heavily enough over a long enough period to result in AN. Pregnant women will not be eligible because of the risks of the CT scan to the fetus. To ensure that such women are not inadvertently included, all women under the age of 55 years who deny being pregnant will be given a pregnancy test before the CT scan is conducted. Finally, eligible patients will be confined to those who reside within defined catchment areas of the participating clinics; the latter restriction is necessary for the valid estimation of incidence, and it will also increase the feasibility of obtaining appropriate controls. Enrollment is projected to continue for 20 months, until approximately the end of July 2002; analysis of the data will then be undertaken, concluding in September 2002.

The basic method of patient enrollment involves the study nurses visiting each dialysis unit in their regions at intervals of not more than every other week. Logbooks in the clinics and/or the head nurse of the dialysis unit are consulted to identify patients who have begun dialysis since the previous visit. Whether or not the initial criteria are met in terms of age and the reasons for exclusion is determined from the individual medical records or conversations with dialysis unit staff; patients who potentially meet the criteria are approached to participate in the study and informed consent is obtained (the consent forms have been designed according to the IRB requirements of each institution). The specific procedures for identifying new dialysis patients, obtaining medical records, obtaining informed consent, and scheduling CT scans and in-person interviews have been developed to fit in with the routine work in each of the dialysis clinics so as to cause minimal disruption. Patients will be compensated for their participation in the study; compensation will range from \$50-150, with the specific amount depending on local needs and patient circumstances.

8.2 Control Subjects

Two methods will be employed for identifying population control subjects. For subjects 65 years of age or older, we will apply to the Health Care Financing Administration (HCFA) for a listing of the names and addresses of individuals with Medicare coverage who live in the counties comprising the catchment areas of the four study regions. After the study protocol is approved, we will obtain a computerized file including the above information and gender. An individual of the appropriate age, sex, and region (determined by zipcode or telephone exchange) to be matched to a particular ESRD case will be selected using a computerized random procedure. The name and address will then be used to identify a telephone number for the targeted individual. An approved "Beneficiary Notification Letter", printed on the letterhead of the local institution will be mailed to the subject. Following a 7-10 day interval, the individual will be contacted by telephone by the local interviewer. The household will be screened for an individual of the desired age, sex, and race. If an appropriate individual cannot be identified at that household, or if the individual refuses to participate, the above process will be repeated, as necessary. Once the appropriate individual has been identified and agrees to participate in the interview (CT scans will not be required for control subjects), a mutually acceptable time and place will be determined (see Section 10.1 for a more detailed discussion of the interview procedures). Each control will be given \$25-50 as compensation for participating.

For individuals younger than 65 years of age, state drivers' license records will be used to identify potential control subjects. Although the available information varies somewhat according to state, each state can provide, at minimum, the name, sex, and address of individuals with valid drivers' licenses. In some states, e.g. North Carolina, these records also include individuals who do not drive, but have been issued a state identification card for proof of age and

identity. A computerized file will be obtained from the Department of Motor Vehicles in the states of North Carolina, Oregon, and Ohio. Following receipt of these files by the SEU, the same procedure as described above for the HCFA files will be used to identify potential control subjects. Again, each participating control will be given \$25-50 as compensation.

9. CT Scans

9.1 Methods of CT Scan Reading

For purposes of standardization, all CT scans obtained on patients from the regional centers will continue to be read in the Department of Radiology at the University of North Carolina. Three experienced GU/CT radiologists participate as "expert readers". They interpret the CT scans that have been submitted in a uniform format (see below). It is acknowledged that not every radiologist reads examinations in an identical manner, and thus variations in opinions among the individual readers are discussed as a group. All radiographic findings are agreed on by the three readers, and their final collective judgement is "the official interpretation" of whether AN is present or not. *The radiologists have no knowledge of the clinical status or analgesic history of any of the patients whose studies are submitted.*

9.2 Technical Issues of CT Scanning

It is likely that each of the participating institutions will continue to have different CT scanners with different technical capabilities. However, from the experience gained in Phase I, the interpretive methodology can be applied to any type of non-contrast renal CT provided that the recommended technical parameters are followed at each institution.

9.2.1 Type of Scanner

Because of the advantage of a true volumetric data set, scans are obtained with the "spiral" (helical) technique; each scanner must be capable of producing excellent images with the recommended standardized protocols. A CT technologist supervisor has been identified at each institution to help facilitate on-going image quality, and each scan is previewed for technical quality upon receipt, by Dr. Clark and his research assistant, Kathy Wilber, R.T.

9.2.2 Patient Preparation

No patient preparation was specified in Elseviers and Debroe's work¹ and none is required for this study.

9.2.3 Scan Parameters

A single breath-hold spiral sequence through the kidneys is required with a slice thickness of 5 mm, pitch of 1 and reconstruction interval of 5mm. This produces approximately 20-35 images per case depending upon renal size and position. Images are obtained from 1 cm above the top of the higher kidney to 1 cm below the bottom of the lower kidney. From the initial experience in Phase I, we have chosen an MAS of 210, a KVP of 120, a standard soft tissue algorithm, and a 35 cm field of view.

9.2.4 Display Parameters

For consistency, we print the images utilizing a 12 on 1 format (14" by 17" film). Two sets of images are printed. The first utilizes a "soft tissue" window width of approximately 250 HU (slightly more contrast than usual) with a center of 30 HU. The second set utilizes a narrow

window of 2 HU with the level set at the minimum value acceptable for diagnosing calcification. The latter level, in our experience from Phase I, is 110 HU.

Measurements of renal dimensions and parenchymal thickness (specified below) are made off the local workstation by the CT technologist and printed with the images of the case. This greatly facilitates the interpretation of the case by the central Radiology Committee. It is acknowledged that the raw digital data from each scan would be most desirable to review centrally but differing output formats from each local CT scanner and other technical issues currently preclude this.

9.3 Radiological Data to Be Analyzed from Non-Contrasted CT Scans

We will continue to look at Debroe's criteria (renal "volume"; cortical indentations; and presence of papillary calcification)¹ because of their potential usefulness and their lack of validation by others. We will also employ other more accepted radiological parameters to assess renal size, contour and calcifications, as described below. As mentioned in the Final Report from Phase I, the Radiology Committee has identified four of the 31 reviewed ESRD patients as having findings consistent with AN. Small kidneys with indentations and/or papillary calcifications were present in these four patients.

9.3.1 Renal Volume

In Debroe's work,¹ renal "volume" was taken as an arbitrary number defined as the sum of width and height at the level of the renal hila. We favor a methodology that takes into account true renal volume and parenchymal thickness rather than a sum of two renal dimensions taken from a single central slice. These parameters have been normalized in Phase I and the data are

summarized in the Final Report. Although renal volume data are easier to obtain, we believe that parenchymal thickness measurements should continue to be made during Phase II because so much of the renal volume in some ESRD patients consists of peri-pelvic fat.

It may be that the assessment of renal size as reported by Debroe will be just as helpful in discriminating patients with AN as the more complex methodologies suggested above. However, U.S. radiologists should be more comfortable with the techniques described above because they reflect true renal size (length and volume), which also has more clinical relevance, although other parameters are used more often. Renal length as determined from urography, ultrasound, and to a lesser extent, CT is the most commonly utilized and accepted measure of renal size. Based on a review of the literature, and from discussions with fellow urologists, renal volume, determined usually by ultrasound, is not used as frequently as renal length. The evaluation of parenchymal thickness has received even less attention in the literature and in daily clinical practice. However, assessing the two latter parameters will be a beneficial byproduct of this investigation.

To summarize, a number of potentially useful measurements reflecting renal size, parenchymal thickness, and volume will be used in Phase II. Data will be obtained to allow the evaluation of each of them; the parameter that ultimately proves to provide the best combination of discrimination between AN and other ESRD, along with wide applicability beyond the proposed study, will be recommended at the end of Phase II.

9.3.2 Indentations

Debroe treated indentations categorically, classifying them according to the single level that appeared to contain the most scars. As in Phase I, we will count all the indentations on the

renal surface as recorded on the series of CT slices and treat this as a continuous variable. Inaccuracies due to the presence of normal fetal lobations are taken into account by the Radiology Committee. Distinguishing fetal lobations from parenchymal scars due to AN or some other renal insult is often not straightforward. However, the indentations from fetal lobations usually are not associated with underlying parenchymal thinning and are usually “minor.” As mentioned above, indentations are classified as “minor” (<5mm deep from the normal contour) and “major” (\geq 5mm deep from the normal contour).

There are many causes of renal scarring besides papillary necrosis (reflux nephropathy, infarcts, trauma, focal pyelonephritis). The ability to distinguish the causes of many cortical irregularities probably will not be possible from radiological findings alone.

9.3.3 Calcifications

Debroe also treats the presence of calcifications on the papillary line as a categorical variable.² What HU of CT density constitutes a calcification was not defined. From experience in Phase I, we have defined “calcification” to mean an HU density greater than 110, and have found that the total number of papillae with calcifications can usually be counted. From Debroe's work any papillary calcification suggests underlying AN with a relatively high sensitivity and specificity. As mentioned above, renal papillary calcifications or calcifications immediately adjacent to the renal papillae are likely to occur in many individuals not suffering from AN (patients with metabolic stone disease—14 of 61 normals in the Phase I data). Thus, it appears that a grading system for both number of papillae with calcifications (Class) and size (Grade) of papillary calcifications is useful. We assign a Class of 1 to kidneys with 1-3 papillae with calcifications, a Class of 2 to kidneys with 4-10 papillae with calcifications, and a Class of

3 to kidneys with greater than 10 papillae with calcifications. With regard to size of calcifications, Grade 1 indicates that the largest papillary calcification is less than approximately 2 mm; Grade 2 indicates that the largest papillary calcification is less than 5 mm and Grade 3 is assigned to kidneys with papillary stones greater than 5 mm in maximum diameter. We understand that different window and level settings may affect the apparent size of any renal calculi,³ but some estimate is important together with their location (either in the papilla, in the adjacent calyx, or deeper in the parenchyma).

9.3.4 Other Observations

Each CT scan will be evaluated for other findings of potential importance. For example, calcifications in locations other than the papilla (cortex, calyx, outer medulla, vascular) are recorded. Cysts or other masses are recorded and the amount of malrotation or other positioning abnormality of each kidney is assessed because of its direct bearing on measurements of renal length and volume (see above).

9.4 CT Scan Diagnostic Criteria for AN

A major focus of Phase I was gaining experience useful to the development of diagnostic criteria based on CT scan. The CT criteria must be precise enough to render a sensitive and specific diagnosis, and yet broad enough in the initial designation to include the possibility of detecting changes that may hold significance; the criteria must also yield reliable (i.e., reproducible) decisions from different reviewers. In previous studies, the specific diagnosis of AN has been made based on the demonstration of RPN in a setting of documented analgesic use.⁴ However, this approach negates a fundamental principle of epidemiology: the fact of

exposure *cannot* be used as a diagnostic criterion in a study undertaken to determine causation. To do so is to permit no possibility other than a 100% exposure rate in the cases, regardless of whether analgesics do or do not cause AN.

We can draw a number of conclusions from the radiology experience of Phase I that will affect our imaging approaches in Phase II. First, spiral CT is a suitable diagnostic modality to assess renal size and shape, and our chosen scanning parameters are valid. Second, our various renal measurements are reproducible, reliable and relatively convenient to obtain after a minimum of training. Third, our data regarding size and parenchymal thickness related to age, sex, BMI and the presence of ESRD appear clinically reasonable and are able to discriminate different groups appropriately. Fourth, there indeed appears to be a small subset of ESRD patients who, besides having small kidneys, also exhibit major contour irregularities (indentations) and/or grade 2 papillary calcifications similar to those described by DeBroe. These patients might therefore have AN. We think that the ability to distinguish size and a more exact location (not graded by DeBroe) of these calcifications will be important because of the relatively high incidence of tiny metabolically induced papillary calcifications in some regional U.S. populations. Likewise, the size of the indentations may help to identify patients with AN but with less accuracy than the presence of calcifications, since there are so many other causes of renal scarring.

We thus plan to continue to apply the basic DeBroe criteria to suggest the presence of AN, but will continue to refine these radiological observations. To date, with only four (out of 31) ESRD patients exhibiting findings suggestive of AN, it is not possible to set more rigid criteria. That will be a major focus of the radiological aspects of Phase II.

9.5 Recording CT Information for Review

A data form, enclosed in Appendix 1, has been developed for recording the information from the CT scans and the assessment of the Radiology Committee. The form has been programmed in Access for direct entry of the information into the computerized database. The evaluation of the CT scans will be conducted by the Radiology Committee in face-to-face meetings, with the participation of the Study Coordinator as a facilitator and recorder. This helps to ensure uniformity of the evaluations, and consistent, accurate recording of the information. Specifics of the information that are recorded for each kidney are summarized below:

Initial information (page 1), including basic patient details and technical information concerning the scan.

Renal length and volume (page 2). Length is calculated as the total number of slices in which renal tissue is seen $\times 0.5$ cm. Volume in cm^3 is calculated as length \times width \times AP dimension $\times 0.49$.

Average parenchymal thickness (page 3) of three contiguous central slices, calculated as anterior+posterior+lateral $\div 3$.

Kidney position and renal contour (page 4). Positional parameters recorded include malrotation, ectopia, and other aspects. For contour, the number of major (≥ 5 mm) and minor (< 5 mm) indentations are recorded.

Calcifications (page 5), including papillary calcifications, classified according to the number of papillae and graded according to size, and the location of any other calcifications.

Masses (page 6), including the number, size, and density, along with any extra-renal findings.

Decision of the reviewers (page 7). Subjects are classified according to the published Debroe criteria of decreased renal mass + bumpy contours and/or papillary calcifications, and according to possible new parameters for AN that will continue to be developed during Phase II.

10. Interview

10.1 Questionnaire

The questionnaire has been designed to obtain valid information on analgesic use and relevant covariate factors; it was revised according to the study experience during Phase I (see Appendix 2). It is programmed in Access, with the data entered directly into the computerized database. The questionnaire is the primary source of information that will be used to assess the performance of the CT scan. Particular attention has been paid to structuring the questions so that accurate and comparable information on regular analgesic use and the timing of early symptoms of renal disease (for cases) is obtained. This will allow the specification of exposure that occurred before the onset of ESRD. A summary of the information obtained follows. A detailed explanation of how to conduct the interview is contained in the study manual (the manual is presently being revised, and a copy will be sent to NIDDK as soon as it is available in January 2001—the original manual was submitted with the protocol for Phase I).

Initial information (page 1), including age, sex, marital status, race, years of education, height, current weight and weight 10 years ago, and current diagnoses (for cases, obtained from the medical record).

Exclusion criteria (page 2), including diabetes, nephrectomy, and current pregnancy (for women under 55 years of age). The interview is terminated if a subject answers affirmatively to either of the latter two factors. ESRD patients who deny being pregnant are also given a

pregnancy test before undergoing the CT scan. Although diabetics will ultimately be excluded unless it is clear that the disease is unrelated to the progression of their renal problems, the interview is completed for those who respond positively to the questions about diabetes, because patients sometimes have inaccurate information about that disease. The diabetic status of the latter individuals is confirmed by the local investigator, who then decides about exclusion.

Renal history (pages 3-5; for cases only). This section of the questionnaire is of major importance, since the information will be used to time the onset of renal problems. The date of a first diagnosis of renal disease is recorded, along with the date of first dialysis. Questions are asked on the occurrence and timing of conditions diagnosed by a physician that indicate renal disease (e.g., proteinuria, hematuria), elevated BUN or creatinine, nocturia (added in the revision) obstructed kidney and kidney infections, kidney stones, urinary tract blockage, urinary tract infection, and renal x-rays. For each of the above, timing is recorded only to the nearest year because patients very rarely remember the month of occurrence. The information obtained above will be used, where possible, to set an index date, or estimated onset of renal disease. The index date is important as a reference point for separating potentially etiologically relevant medication exposures from those that occurred after the illness had commenced. Symptoms that can occur with renal disease are no longer enquired about based on experience during Phase I: they occurred too recently to be useful in timing in the onset of renal problems. Hypertension is also enquired about, including a question about hospitalization for the condition, as a measurement of severity.

Other medical history (pages 6-7). Information on gout and peptic ulcer or hemorrhage is elicited from all subjects. In addition, the patient is asked about chronic pain conditions that

normally lead to the regular use of analgesics, including frequent headaches, arthritis, and other pain conditions.

Analgesic history (pages 8-11). A lifetime history of regular analgesic use (defined as use at least once a week for at least three months) is elicited by asking systematically about indications for use. The latter includes (prompt #1): pain, headache, backache, toothache, menstrual pain, muscle relaxant/spasms, arthritis, gout/high uric acid, and swelling/inflammation due to injury. To capture prophylactic use, analgesic abuse, and use due to vague symptoms of renal disease, the list also includes the following indications (prompt #2): “help sleep or relax”, “feel better or perk up”, prevent headaches or other pain, and prevent heart disease, thrombosis, or stroke. The indications are followed by a list of the trade names of commonly used products in all analgesic categories, i.e., aspirin, acetaminophen, combination products, NSAIDs (prompt #3). The current version of the list is included with the questionnaire (page 9). To further assist recall, subjects can be shown a book with pictures of analgesic products used in the U.S. in the past 20 years (the book will not be used if telephone interviews are adopted—it was rarely helpful in Phase I). Finally, subjects are asked about any other pain medications not covered by the specific indications or trade names (prompt #4).

Drug names are coded based on the SEU drug dictionary that cross links trade and generic names of drug products with their components. This sophisticated and powerful tool was developed in 1975 by SEU research pharmacists and has been continuously updated. It allows the evaluation of individual products (e.g., Tylenol), or groups of drugs containing the same active ingredient (e.g., all acetaminophen-containing products).

For cases, the reference point for enquiring about drug use (as well as most other information) is the date of initiation of dialysis, which provides a standard and recent event that

the interviewers and subjects can focus on. For controls, the reference point is the date of interview. The details of each episode of regular use before this point are recorded, including year started, duration, frequency of use, number of pills per day, and specific reason for use. It would not be possible to obtain a reliable lifetime history of occasional or very short-term analgesic use by interview, and we do not attempt to do so. This is not a problem in that the hypotheses are concerned with regular long-term use, which can be obtained with acceptable accuracy by the approach set forth here.

It is not feasible to obtain explicitly the milligram dose of each product used over an entire lifetime. However, the doses are fixed in combination products (although these may vary over time), and in single ingredient analgesics are usually inherent in the product name, which is recorded. Thus, in many instances it is possible with the names, dates of use, and number of pills per day to infer the daily dose, and hence the total lifetime dose. Of course, it is not always possible for the subjects to remember the specific trade names (the use of pictures can sometimes be helpful in obtaining the precise names) in which case the dose values are considered to be unknown.

Family history (page 12) of ESRD.

Habits (pages 13-14) of consumption of tobacco, coffee, tea, and alcohol. For beverages, consumption 10 years ago is recorded; questions regarding current consumption (i.e., before starting dialysis) have been discontinued because there is evidence that individuals change their consumption habits as their illness commences, so that current consumption does not reflect usual habits.

Occupational history (page 15), amended to include focused questions on particular occupations that could potentially involve exposure to heavy metals or solvents. Includes a brief description of job tasks, the year the job was started, and its duration.

Income (page 16). This information is placed late in the interview because of its sensitive nature.

The questionnaire has thus far been administered to all study subjects in person. However, telephone interviews are thought to be a more feasible method for obtaining participation of controls. During the month of December 2000, we will explore the appropriateness of telephone interviews for ESRD cases by randomizing them to interview method (telephone or in person) unless an in person interview is judged to be necessary (e.g., because of short attention span). A comparison of the completeness of responses for the two methods will then be made. If telephone interviews are shown to yield good information, then that method will be used for each case-control pair where possible. If it is necessary that a particular case be interviewed in person, then the control for that subject will be interviewed in person to maintain a comparable interview setting. In person control interviews will be conducted at a local medical facility (e.g., the dialysis unit where the case was identified).

Routine quality control procedures include detailed review of each interview by the Study Coordinator in Boston, automatic computerized checking of the data for consistency, and prompt correction of errors that are discovered, with feedback to the interviewers on an ongoing basis.

Considerable attention has been, and will be, paid to training the interviewers in the clinical centers so that the information obtained is standardized and comparable. At the existing centers in North Carolina and Oregon, the interviewers from Phase I will continue their efforts

throughout Phase II. The interviewer hired for the Ohio center will be trained in Toledo by the SEU staff in early January 2001. All study methods and procedures will be covered in the training, including developing the mechanisms for identifying potential cases, obtaining information on possible exclusions, obtaining informed consent, scheduling the CT scan and the interview, maintaining a log of study subjects, including reasons for nonparticipation, administering the interview, and transmitting the interview data to the SEU.

10.2 Validation of the Questionnaire

It is always difficult to validate interview information, because there is seldom an absolutely reliable “gold standard” against which the questionnaire data can be measured. This is particularly true for use of nonprescription medications. While that situation applies in the present instance, an indirect approach to validation was used during Phase I to gain insight into the accuracy of the histories. Twelve ESRD patients interviewed during Phase I were selected at random and reinterviewed two to three months after the initial contact. We then conducted a detailed comparison of each individual’s responses on the two questionnaires. Good to excellent agreement was observed for demographic factors and for most dichotomous variables. In general, dates and durations of exposures were less well remembered. The analysis of analgesic exposure that occurred before the index date and lasted for at least five years revealed concordance for 7 of the 11 subjects who provided this information on both interviews. There was discordance on this measure of exposure for two subjects due to different information on the timing or duration of use on the two interviews; one subject’s information was discordant because of unknown information on timing and duration of use on one of the interviews. Finally

one subject neglected to report exposure to one drug that had been mentioned at the first interview.

As discussed in greater detail in the final report from Phase I, there are several issues to consider in interpreting the reinterview data. First, certain external influences could result in poor agreement between the two interviews. These include the stimulation of the patient's memory at the first interview, resulting in better recall on the second; improved physical condition of the patient at the second interview, resulting in more thoughtful responses; patient boredom at the second interview, with knowledge that briefer responses will shorten the time required; and increased rapport between the interviewer and subject, resulting in more thoughtful responses at the second interview. Alternatively, the patient could remember his responses to the first interview and repeat them, creating apparent agreement regardless of the correctness of the responses. Finally, it must be remembered that at best the reinterviews provide a measure, albeit imperfect, of the reproducibility of the questionnaire data, but not the validity, which is really what is desired.

A further indirect measure of validity was obtained by examining the questionnaire data from four enrolled subjects who had been identified by their physicians as being analgesic abusers or having analgesic nephropathy. A substantial analgesic history was obtained from all four patients in the study interview. While this is reassuring, it must be remembered that the physicians' diagnoses were also based on histories provided by the patients.

Based in part on the analysis of the reinterview data, the questionnaire has been redesigned for use in Phase II, with the goal of shortening and simplification. We judge that the new questionnaire will yield more reproducible responses, with the caveat that an unavoidable level of imprecision will remain due to the shortcomings of human memory.

11. Data Management

11.1 Quality Control

Attention to quality control of the data is a major emphasis of the SEU in its role as Coordinating Center of the current project. In addition to on site training of the interviewers, the following procedures are employed.

Automatic consistency checks of the data have been built in to the computerized interview. Some examples of automatic error checking and rejection of inappropriate values include: providing predefined choices for categorical variables with the use of drop-down lists; warnings when questions are not answered; warnings on outlying values (e.g., more than 50 alcoholic drinks a day); checking for valid dates; checking for valid ages (e.g, the age at diagnosis of a medical condition cannot exceed the current age); checking that the year analgesic use started is consistent with the duration of use (e.g., the year started plus duration must not be later than the current year); automatic skipping of questions that should only be answered if there is an affirmative answer to a previous question; online coding of drugs with the SEU drug dictionary, which will prevent misspelled drug names and coding errors.

Each interview is individually reviewed by the Study Coordinator in Boston for completeness and consistency; errors are corrected promptly, in consultation with the local interviewer as necessary. The Study Coordinator maintains regular telephone and e-mail communication with the interviewers to ascertain problems, clarify questions about the data, and provide feedback on performance. During Phase II, a site visit will be conducted to each of the study centers by the Study Coordinator and Co-Investigator or Epidemiologist. There will also be one meeting of the data collection staff in Boston during Phase II; the latter is important to assure the consistency and comparability of the data. In addition to the ongoing quality control

activities, the first step in analyzing the data will be for the epidemiologist to check the data for outlying values and computing errors. Any errors in the raw data that are discovered during this process will be corrected.

Quality of the radiological data is ensured by the use of a rigorous and standard protocol for conducting and printing the scans, as described above. Relevant data from the scans, along with the assessment of the Radiology Committee, is recorded on standard forms during the review sessions. To maximize the consistency of that process, the Study Coordinator participates in the meetings as a recorder and facilitator.

11.2 Data Handling Procedures

Secure master data files are maintained at the Coordinating Center in Boston. To assure confidentiality, names and other specific individual subject identifiers are not included in the main data files. The records are identified only with an anonymous study number. Where individual identifiers are required (e.g., if it is necessary for the local center to contact a patient concerning results of their CT scan; control addresses and telephone numbers used for recruitment), these are kept in separate files in the regional study centers or SEU, that are not linked to the master data files by computer.

Data are electronically transmitted to the SEU every week over a virtual private network (VPN) with an enhanced firewall; the VPN allows a connection to the Unit's computer system through a local internet service provider or academic computing center with internet access while still maintaining security. An automatic upload procedure developed by the SEU computing staff transmits an encrypted database containing the interviews that have not previously been sent. These are then added to the master data file for quality control. At regular

intervals, a copy of the quality-controlled data is transmitted back to each center. Previously uncorrected interviews on the local database are automatically replaced with the “clean” data, so that each center has a copy of the interviews conducted at that site that is as up to date as possible.

Interview data are backed up daily onto Iomega Zipdrives, using a procedure developed by the SEU computing staff. Since each interview computer contains a copy of the Access database, it is straightforward to simply copy it from the hard drive to the Zipdrive. Ten date-stamped backup copies are maintained on a rotating basis, with each copy being overwritten every tenth day. The main database on the SEU server is backed up daily to a tape drive. All data on the servers are automatically backed up in this way, with procedures that run automatically during the night. Server backups are maintained on a rotating basis for seven days. Friday night backups are rotated offsite weekly, and rotated back in four weeks, with the exception of the backup from the first Friday of the month, which is permanently archived offsite. Analytical data files are also kept on the server, and thus automatically backed up.

An audit trail is maintained by a number of steps. First, data transmitted from the local centers are copied to the active database, and also to an inactive copy to which no changes are made. The latter provides a record of the data as originally collected; the interviewers are identified, and an automatic date and time stamp for each interview is included. As part of the quality control procedure, the Study Coordinator or any other person making changes to the data (e.g., the Epidemiologist) first logs their name in to a quality control startup dialog in the main active database. Any changes that are made during that session are then automatically recorded in a table which contains the subject ID number (an internal study number to which all data for that individual are linked), field changed, old value, new value, and the date, time, and name of

the person making the change. The procedures provide a comprehensive archive of the original raw data and a change log where all differences between the original and working dataset are specifically identified. These measures allow the reconstruction of the sequence of events for any and all changes made, with considerable ease and flexibility (e.g., the log could be sorted by ID number, date, person making the change, field changed, or some combination of factors). A particular advantage of the automatic logging of changes is that it is minimally burdensome to study personnel in the ongoing performance of their regular duties; it is therefore not only efficient, but also cost-effective.

12. Data Analysis

12.1 General Issues

The analytical plan for Phase II has been discussed in detail in the original proposal (especially Section 5.3) and the response to technical proposal questions 8g-8i of January 1999. Here we will cover several of the most important points. As noted in Section 2, the main analytical objective of Phase II is to determine whether AN can be identified by a non-contrasted CT scan. This requires first a classification of the ESRD patients according to whether or not they meet CT criteria for AN. We have explained in Section 9 of this protocol that both the published criteria of Debroe and alternative criteria developed specifically for this study will be utilized for that purpose. Second, patients must be categorized according to their analgesic use prior to the onset of renal disease. Then the question in its most simplified form is whether patients with positive CT criteria for AN are more often heavy users of analgesics than those whose CT criteria are negative, and if so, how much more often.

A major challenge that has bedeviled all studies of AN is to establish the temporal sequence between analgesic use and renal disease, because the main hypothesis is that analgesic use is the initiating factor in renal insufficiency that progresses to ESRD. A separate hypothesis concerns the effect of more recent analgesic use as an exacerbating factor among individuals who already have renal disease. The evaluation of that hypothesis is problematic because of the likelihood that persons with renal disease modify their analgesic intake in response to their symptoms (confounding by indication), thereby distorting comparisons. The primary focus of Phase II, and hence this protocol, will be on the de novo causation of renal disease by prior analgesic use.

In order to examine analgesic use prior to the illness, it is necessary to set an *index date* that approximates the onset of renal disease. Any such date is necessarily imprecise because of the insidious onset of renal insufficiency and the imprecision in patients' recall of renal signs and symptoms in the distant past. There is probably no single definition of the index date that can be considered absolute. One basic approach is to take the earliest reported date of an agreed upon group of renal signs and symptoms as reported by each patient, and then add an additional interval (e.g., two years) to allow for the earlier actual onset of the disease (referred to in the proposal as "lagged exposure analysis"). Another approach is to apply a fixed interval to all subjects, which avoids inaccurate reporting but inevitably results in misclassification for individuals; the ability of the analysis to detect associations is thus reduced. The questionnaire has been designed to obtain the information needed to set an index date using the first approach. However, our initial experience with the Phase I data has shown that this may not be feasible for many patients because dates are unknown, or reported to be so recent as to not provide credible timing of onset of renal disease (e.g., within one or two years before dialysis). Under the

circumstances, another possibility is a hybrid of the two approaches, in which a specific index date is set for those subjects whose information permits it, and a fixed interval is used for the remaining subjects. The hybrid approach was adopted for the analysis of Phase I data. Although we plan to use a similar strategy for Phase II, it should be emphasized that any approach will be subject to the limitations of the data. The key point is that the necessary data are being obtained to allow various definitions of the index date, each of which will have advantages and drawbacks.

There will be two basic types of analysis to evaluate analgesic use *before* the onset of renal disease in relation to the CT findings: a relative risk analysis and a “sensitivity and specificity” analysis. Each is described below. In addition, the data will permit the evaluation of use *after* the index day. As noted above, the latter analysis will address the question of exacerbation of pre-existing renal disease by ongoing analgesic use. We do plan to look at the data in that way, with the caveat that interpretation will be difficult because of confounding by indication. Since the exacerbation question, although clinically important, is not germane to the main thrust of the study, it will not be referred to further in this document.

12.2 Relative Risk Analysis

Odds ratios (and 95% confidence intervals) will be estimated as measures of the relative risk for various categories of analgesic use. There will be three outcome groups: ESRD patients with AN according to the CT scan criteria, other ESRD patients, and normal controls. Odds ratios will be estimated firstly by comparing rates of analgesic use in AN with those in other ESRD; this will evaluate the question as to whether there is an entity of AN identifiable by CT scan that is associated with heavy analgesic use, and that can be reliably distinguished from other

ESRD (i.e., whether the CT criteria are useful in identifying AN). Secondly, rates of analgesic use will be compared between non-AN ESRD patients and the controls in order to evaluate the question of whether heavy analgesic use is associated with ESRD other than AN. “Heavy use” will be defined according to duration, frequency, dose, or some combination of those parameters. More than one definition will be used in order to evaluate different aspects of exposure (e.g., long duration vs. large total lifetime dose). The precise specifications cannot be made in advance because they will depend on what is feasible based on the distributions in the total ESRD population (without regard to AN status) and controls—i.e., for informative analysis, it will be necessary to define categories that contain a reasonable number of subjects. Ordinarily, however, heavy use will be considered as *not less than* two or more days a week for at least five years.

An index date will be set by one or more of the approaches described in the preceding section, and analgesic use will be classified with reference to that date. For comparisons between ESRD patients and normal controls, an index date will be specified for each control that is equivalent to the date for the matched case, and exposure will then be defined in the same manner. Regular users of various specific analgesics (e.g., acetaminophen) before the index date will be further classified according to duration of use (e.g., $<5\text{yr}$, $\geq 5\text{yr}$), total dose consumed (based on number of days per week, number of pills per day, and product names), and timing (e.g., first use $<10\text{yr}$, $\geq 10\text{yr}$ before the index date). Users will also be subdivided according to single ingredient and combination products (e.g., aspirin/acetaminophen/caffeine). As noted in the preceding paragraph, the optimal duration, timing, frequency, and dosage categories that identify heavy and more moderate analgesic users will be determined during interim analyses, based on the distribution of subjects from the total ESRD population and without regard to CT

findings. In the results, the greatest credibility in terms of valid exposure will be given to use of several years' duration (e.g. ≥ 5 yr), since such use should be well-remembered by all subjects, and there is a greater probability that at least some of the use preceded the true onset of renal disease.

Odds ratios and 95% confidence intervals will be estimated for the various exposure categories, compared with non-use of the drug at issue. In addition to the evaluation of categories of heavy use, the possibility of dose-response and duration-response relationships will be evaluated in detail by estimating relative risks for categories of increasing dose and duration, compared to never use. Apparent trends in such analyses will be tested for statistical significance by standard methods (for example, by entering ordinal terms into logistic regression models, with increasing levels indicating increasing analgesic use defined according to duration, total dose, etc.).

To allow for overlapping use of multiple drugs, and the confounding effects of other factors such as age, sex, and predisposing conditions, multiple logistic regression will be used. Conditional logistic models will be employed in the comparisons with population controls, since it will be necessary to take matching of cases and controls into account. In comparisons between the groups of ESRD patients, which will not be matched, unconditional logistic regression will be used. Separate analyses of the data from each of the regions will also be conducted.

If there proves to be a strong and relatively precise measurement of an association between "AN" as defined by the CT criteria and analgesic use among ESRD patients, as indicated by a large odds ratio and narrow confidence interval, it will constitute evidence supporting the existence of AN as a discrete entity that can be distinguished from other ESRD by CT scan. Assuming that AN exists, a modest association could reflect imprecision either in the

CT scan criteria or in the questionnaire, or it could reflect some level of association between analgesic use and ESRD other than AN. In the latter situation, the prevalence of analgesic use in the CT negative ESRD patients would be relatively high, and the difference between that and the prevalence in the CT positive patients would be reduced. The normal controls will provide a measure of the prevalence of heavy analgesic use in the base population that will serve as a standard against which analgesic use in the AN and other ESRD patients can be compared, thereby enabling the interpretation of a modest association.

Little or no association would constitute evidence that AN may not be a discrete form of ESRD that can be identified by CT scan, but it would leave unanswered the question as to whether analgesics increase the risk of ESRD in general. That question can be evaluated in a traditional case-control analysis, in which analgesic use in all ESRD patients is compared with controls. The data from Phase II will be used for such an evaluation if the results indicate it. Although this part of the study was not powered to provide a definitive answer to that question, there will be sufficient data to give an indication as to the direction of the answer. The considerably larger amount of data collected in Phase III would be more definitive, if that phase is conducted as we have proposed.

12.3 Sensitivity and Specificity Analysis

Ordinarily, the assessment of a diagnostic test such as the CT scan is done by measuring its sensitivity and specificity, and such an evaluation was requested in the original RFP for this contract. Sensitivity and specificity are determined by applying a diagnostic test to two groups of patients: those who are known by some other measure to have the disease in question (e.g., biopsy-proven cervical cancer), and those who do not. The sensitivity is then the proportion of

diseased patients who test positive, which can be restated as the likelihood that the test will diagnose the disease if it is present; the specificity is the proportion of nondiseased patients who test negative, or the likelihood that the test will correctly identify subjects who do not have the disease.

Elseviers and De Broe have published a number of papers on the sensitivity and specificity of the CT scan in diagnosing AN in Europe.^{1,2,5,6} There is, however, an important conceptual difficulty with their use of the terms: there is no “gold standard” for determining the presence or absence of the disease against which the CT scan can be evaluated. Until now, a history of heavy analgesic use determined by interview has always been a requirement for the diagnosis. That requirement contravenes one of the first principles of causal research in epidemiology--that the disease be defined independently of the exposure which is under study. Thus, Elseviers and De Broe did not measure sensitivity and specificity in the classical sense, because there was no absolute determination of the presence of AN without reference to the analgesic history.

With those caveats in mind, we will nonetheless conduct a sensitivity/specificity analysis as requested. For this analysis, exposure will be defined as described above for the relative risk analysis. Subjects will be divided into heavy users of each relevant drug (e.g., use four or more days a week for ≥ 5 yr), other users, and non-users. CT scan results will be tabulated for the groups, and “sensitivity” and “specificity” will then be estimated, both with non-users and “other” users combined, and with the latter subjects excluded. The specific exposure categories to be evaluated will be limited to those that are identified in the relative risk analysis as strongly associated with CT positivity.

13. Study Design Issues

In this section, a summary of the methodological issues that have guided the study design is provided.

13.1 Selection Bias

In an unbiased study, the reason for selecting any subject should be independent of the exposure. If it is not independent, there is selection bias. In the current study, we have avoided such selection bias by obtaining exposure information subsequent to identification for the study.

Selection bias may also arise if a large proportion of eligible subjects targeted for inclusion are not successfully enrolled, and if the rates of analgesic use differ between the subjects who are enrolled and those who are not. In the context of AN, selection bias could arise if otherwise occult uremia were selectively diagnosed among analgesic users (e.g., if they are under active medical care because of arthritis, and thus more likely to have diagnostic tests performed). Similarly, bias could arise if there was a selective tendency to perform CT scans on uremic patients (and thus to identify features compatible with a diagnosis of AN) if they had used analgesics, but not otherwise.

In this study, selection bias is minimized by identifying *all* ESRD subjects in the participating centers who meet the definition criteria of the study, and enrolling as many as possible. A number of approaches are taken to minimize losses. The requirement that patients be identified at the time of commencement of dialysis will reduce the number that cannot be located, or cannot be interviewed because they are too sick. It also ensures that few have yet begun home dialysis, increasing the feasibility of conducting the interviews in the clinic. Based on experience during Phase I indicating difficulty in scheduling CT scans during the allotted

time, the acceptable time interval has been increased to from one month prior to the start of dialysis (for patients who already have had access inserted, have a creatinine clearance of 10 ml/min or less, and have scheduled a dialysis appointment) to three months after the initiation of dialysis. The latter extension allows time for the patient's physical condition to improve following the beginning of dialysis treatment. If possible, the interviews are conducted while the patients are actually being dialyzed, at a time selected by the interviewer when their mental state is optimal. That arrangement further enhances participation, as well as the quality of the data because the patients are available for interview over a number of hours. As noted in Section 10.1, we are also proposing to use telephone interviews where feasible, which should also enhance participation, especially for controls.

13.2 Information Bias

Information bias, or differential reporting of exposure histories by cases and controls, is a major methodologic concern in the current study. Because of general concerns in the medical community about AN, patients with otherwise unexplained renal failure are likely to have been questioned carefully about their analgesic use by medical personnel in advance of their interview for the study. Controls do not have this prior stimulation of the recall of their analgesic use, and are probably less motivated to probe their memories.

In the current study, stringent measures have been taken to ensure the completeness and comparability of the information from all subjects. We use a highly structured questionnaire, designed to maximize the completeness of recall, and interviewers are carefully trained to conduct it in the same manner with all subjects.

In the analysis, information bias can be minimized by placing the greatest reliance on evaluation of the most heavy analgesic use categories. For example, if there was a strong association between ESRD and daily use of a particular analgesic for five or more years, with the reference category being subjects who have never been exposed, we would judge the association to be unlikely to be affected by information bias.

13.3 Misclassification

In any causal research an absolute requirement for validity is that an exposure to an agent must antedate the onset of the disease it is alleged to cause. This requirement, although obvious, is difficult to satisfy because of the insidious and chronic nature of the changes that follow the onset of renal insufficiency, and the impossibility of pinpointing exactly when it began. In addition, as already noted, symptoms due to the underlying renal insufficiency (e.g., fatigue, lassitude, nausea, and headaches) may “cause” analgesic use (confounding by indication), rather than the reverse. To evaluate the hypothesis that analgesics cause renal disease *de novo*, it is necessary to restrict the evaluation to analgesic use that took place before the onset of disease, and that date can seldom be established with precision.

The issues concerning the definition of the index date have been discussed in some detail in Section 12.1. We acknowledge that any definition will be an approximation of the onset of renal disease, and that to the degree that misclassification of exposure is introduced, associations will be attenuated. This must be kept in mind particularly in interpreting observed modest associations. Despite the uncertainty about the onset of renal disease, however, it is reasonable to assume that regular analgesic use that lasted for many years before the best estimate of disease onset would have commenced before the illness began. It will not be possible to evaluate short

term or occasional analgesic use in this study because reliable lifetime histories of such use cannot be obtained retrospectively by interview.

The potential for misclassification of outcome also exists. Inasmuch as the CT scan criteria are subjective, with room for variation in observer judgment, there could be some misclassification in the assignment of AN status. Such misclassification would reduce the precision of the CT scan as a diagnostic tool. However, as noted above, it will not be biased according to exposure status, since that information will not be known to the reviewers. The effect of individual variation is minimized because the decisions are made collectively by a three member Radiology Committee.

13.4 Confounding

Confounding arises when an apparent association (or lack thereof) is explained by one or more factors related both to the exposure and the outcome. The principle in controlling confounding in case-control studies is to compare cases and controls within similar categories of the confounding factor. In the present context, with the emphasis on analgesic use prior to the onset of renal insufficiency, the main confounding is likely to come from the use of multiple analgesics at different times or at the same time. Such confounding can be controlled by multivariate analytic techniques, since information on such exposure is obtained, in addition to information on other factors that might influence both analgesic use and the likelihood of ESRD, such as socioeconomic status, race, and prior medical history.

14. Conclusion

It is expected that the 22 months of Phase II, which have been carefully planned, will allow for the full enrollment of 200 ESRD cases and 200 matched control subjects and the evaluation of whether AN can be identified by CT scan. Although the original RFP and proposal called for a third phase in which a larger number of patients would be enrolled to determine incidence and more fully evaluate other questions such as the relationship of analgesic use to ESRD in general, NIDDK decided to defer a decision on whether to proceed with that. We continue to believe that Phase III as originally projected is worthwhile, regardless of the direction of the results of Phase II. Thus if indicated, a modified protocol for the next phase of the project will be developed based on the experience during Phases I and II, which will allow for completion of all of the originally proposed objectives.

15. References

1. Elseviers MM, De Schepper A, Corthouts R, et al. High diagnostic performance of CT scan for analgesic nephropathy in patients with incipient to severe renal failure. *Kidney Int* 1995; 48: 1316-23.
2. Elseviers MM, Bosteels V, Cambier P, et al. Diagnostic criteria of analgesic nephropathy in patients with end-stage renal failure: results of the Belgian study. *Nephrol Dial Transplant* 1992; 7: 479-86.
3. Olcott E, et al. Accuracy of Detection and Measurement of Renal Calculi: In Vitro Comparison of 3D Spiral CT, Radiography, and Nephrotomography. *Radiology*. 1997; 204:19-25.
4. McCredie M, Stewart JH, Mathew TH, et al. The effect of withdrawal of phenacetin-containing analgesics on the incidence of kidney and urothelial cancer and renal failure. *Clin Nephrol* 1989;31:35-39.
5. Elseviers MM, Waller I, Nenov D, et al. Evaluation of diagnostic criteria for analgesic nephropathy in patients with end-stage renal failure: results of the ANNE study. *Nephrol Dial Transplant* 1995;10: 808-14.
6. Elseviers MM, De Broe ME. A long-term prospective controlled study of analgesic abuse in Belgium. *Kidney Int* 1995;48:1912-19.

Figure 1

Proposed Timeline for NANS Phase II*

November 30, 2000	Meeting of investigators and NIDDK to review Phase I
December 1, 2000	Begin Phase II ESRD patient enrollment and CT scans; include Ohio site
December 20, 2000	Delivery of final Phase I report and Phase II protocol
January 5, 2001	EAC meeting at NIDDK
January 2001	SEU to visit Ohio for interviewer training and review of network and enrollment procedures
January-February 2001	Dr. Henrich to visit individual sites, meet appropriate dialysis unit personnel
February-March 2001	Begin enrolling controls for the study
Mid-year 2001	Dr. Henrich to meet with Slone Epidemiology Unit (SEU) to review progress and address ongoing issues Radiology meeting at University of North Carolina Interviewers to meet at SEU
October 2001	Investigator meeting at ASN
January 2002	Radiology meeting; Dr. Henrich to visit SEU and to attend a radiology meeting
January-March 2002	SEU to visit study sites
Mid-year 2002	Radiology meeting; Dr. Henrich to visit SEU
August 2002	Final radiology meeting
September 2002	Final investigator meeting Phase II complete

*Dr. Henrich to meet with SEU and investigators as needed. One or two additional radiology meetings may be needed.

APPENDIX 1
RADIOLOGY DATA FORMS
(paper version)

APPENDIX 2
INTERVIEW QUESTIONNAIRES
ESRD Patients and Controls
(paper version)

APPENDIX 3
INTERVIEW INSTRUCTION MANUAL
(for computerized version)