Non–Contrast-Enhanced Computerized Tomography and Analgesic-Related Kidney Disease: Report of the National Analgesic Nephropathy Study

William L. Henrich,* Richard L. Clark,† Judith P. Kelly,‡ Vardaman M. Buckalew,§ Andrew Fenves,‖ William F. Finn,¶ Joseph I. Shapiro,** Paul L. Kimmel, †† Paul Eggers,†† Larry E. Agodoa,†† George A. Porter,‡‡ Samuel Shapiro,§§ Robert Toto,‖‖ Theresa Anderson,‡ L. Adrienne Cupples,¶¶ and David W. Kaufman‡

**Department of Medicine, University of Maryland School of Medicine, Baltimore, Maryland; Departments of †Radiology and ¶Medicine, University of North Carolina School of Medicine, Chapel Hill, North Carolina; ‡Slone Epidemiology Center, Boston University School of Medicine and ¶¶Boston University School of Public Health, Boston, Massachusetts; §Department of Medicine, Wake Forest University School of Medicine, Winston-Salem, North Carolina; ‖Dallas Nephrology Associates, Dallas, Texas; **Department of Medicine, Medical College of Ohio, Toledo, Ohio; ††National Institutes of Health, Bethesda, Maryland; ‡‡Department of Medicine, Oregon Health Sciences University, Portland, Oregon; §§Department of Epidemiology, Mailman School of Public Health, Columbia University, New York, New York; and ‖‖University of Texas Southwestern Medical Center Dallas, Dallas, Texas*

Previous studies suggested that the non–contrast-enhanced computerized tomography (CT) scan is a highly reliable tool for the diagnosis of analgesic-associated renal disease. However, this issue has not been addressed in the US population. A total of 221 incident patients with ESRD from different regions of the United States underwent a helical CT scan and detailed questioning about drug history. Specific renal anatomic criteria were developed to determine whether a constellation of CT findings (small indented calcified kidneys [SICK]) is linked to analgesic ingestion. For approximating use before the onset of renal disease, only analgesic ingestion at least 9 yr before starting dialysis was considered relevant. Fifteen patients met the criteria for SICK. This represented 7% of the enrolled patients and approximately 1% of the total ESRD population. There was a significant increase in the estimated risk among patients with a history of heavy aspirin ingestion (odds ratio [OR] 7.4 [95% confidence interval (CI) 1.2 to 43] for ≥1 kg lifetime; OR 8.8 [95% CI 1.2 to 66] for ≥0.3 kg/yr). Total analgesic ingestion of ≥0.3 kg/yr also was significantly associated with SICK (OR 8.2; 95% CI 1.5 to 45). These findings were accounted for largely by combination products that contained aspirin and phenacetin (used by three patients with SICK), which are no longer available. In addition, the CT finding of SICK was present only in a minority of heavy analgesic users, yielding a sensitivity of 5 to 26%. Findings of SICK are infrequent in the US ESRD population and do not occur among a sufficient proportion of heavy analgesic users to render the non–contrast-enhanced CT scan a sensitive tool to detect analgesic-associated kidney injury.

J Am Soc Nephrol 17: 1472–1480, 2006. doi: 10.1681/ASN.2005101096

The possibility that the chronic, long-term ingestion of analgesics leads to specific renal pathology and renal failure is a subject of interest and controversy. Interest is derived from observations in the United States and Europe that purport to show that heavy analgesic consumption is associated with significant risk for development of ESRD (1–9). Controversy has occurred as some have criticized the design and execution of previous positive studies (10), whereas other studies have been negative (11).

In the 1990s, Elseviers *et al.* (12,13) concluded from their work in Belgium that the non–contrast-enhanced computerized tomography (CT) scan could identify analgesic-related kidney disease with high sensitivity and specificity. If their observation were confirmed, then it would establish an anatomic entity of analgesic-related renal disease ("analgesic nephropathy") and provide a valuable diagnostic tool for clinical purposes. The National Analgesic Nephropathy Study (NANS) was undertaken to address the ability of the CT scan to detect heavy analgesic use in the US ESRD population.

Materials and Methods

NANS was an observational study that was conducted in five centers: University of North Carolina (UNC; Chapel Hill, NC), Wake Forest University (WFU; Winston-Salem, NC), Medical College of Ohio (MCO; Toledo, OH), Dallas Nephrology Associates and the University

Received October 21, 2005. Accepted February 15, 2006.

Published online ahead of print. Publication date available at www.jasn.org.

Address correspondence to: Dr. William L. Henrich, University of Texas Health Science Center, MSC 7790, 7703 Floyd Curl Drive, San Antonio, TX 78229. Phone: 210-567-0612; Fax: 210-567-4418; E-mail: henrich@uthscsa.edu

ISSN: 1046-6673/1705-1472

of Texas Southwestern Medical Center (DNA; Dallas, TX), and Oregon Health Sciences University (OHSU; Portland, OR). Data coordination was provided by the Slone Epidemiology Center (SEC) of Boston University. The intention in the selection of centers was to include a regional representation of analgesic use, with North Carolina being an area with putatively high use and consequently more analgesic nephropathy (14) and the other three regions being areas of relatively lower use. The study was initiated in April 2000 at UNC, WFU, and OHSU, and extended to MCO in December 2000. Low enrollment necessitated discontinuing the study at OHSU in June 2001; DNA was added in July 2001. Data collection was completed in February 2003.

Patients

Incident patients who had ESRD without a clear nonanalgesic cause for renal failure and had begun dialysis within the previous 3 mo were potentially eligible for inclusion. The Institutional Review Board of all institutions approved the protocol, and informed consent was obtained from all patients. A list of exclusion criteria is provided in Table 1; hypertension was not a reason for exclusion. The strategy was to enroll patients without known causes of ESRD to maximize detection of analgesic-associated changes in kidney anatomy. A non–contrast-enhanced CT scan was performed, along with an interview that was conducted in the clinic to obtain information about analgesic use and other factors.

The initial data collection from April through November 2000 was considered a "pilot" period for refining the study methods; 32 eligible patients who met the final inclusion criteria were identified during that phase. To shorten the interview, we dropped some items that did not yield useful responses; there were no other material changes to the protocol, and, in particular, the inclusion criteria for patients with ESRD remained the same. In the main phase that started in December 2000, 2344 new dialysis patients were identified, 338 (14%) of whom were eligible for inclusion. Of the 402 eligible patients from both phases, 221 completed both the CT scan and the interview, giving a participation rate of 54%.

Radiologic Methods

Non–contrast-enhanced CT scans were performed using a standardized protocol. Only helical CT machines were used. Details of the methods are in the Appendix. Digital images ($n = 195$) and films ($n = 26$) were sent to the NANS radiology center at UNC. In addition to the patients with ESRD, 61 healthy volunteers were recruited from three age categories (35 to 49 yr [$n = 20$] 50 to 65 [$n = 20$], and >65 [$n = 21$]) and scanned at UNC and WFU to provide normal values of renal size. All of the healthy volunteers had a negative history for renal disease, normal blood urea nitrogen and creatinine concentrations, and a normal urinalysis.

Measurements were made centrally from digital images by a radiologic technologist. When necessary, films were scanned for digital measurement. Three methods of determining renal size/volume were used: Ellipse, voxel count, and average parenchymal thickness; details of these methods are described in the Appendix. A committee of three radiologists who had no knowledge of analgesic exposure reviewed the films of scans of all healthy volunteers and patients with ESRD. Clinically relevant findings (*e.g.*, solid renal mass, liver lesion, abdominal aortic aneurysm) were communicated promptly to collaborating nephrologists.

Three parameters were used to identify patients who might have "analgesic-induced" kidney disease: Size, indentations, and calcifications. Our aim was to develop a more quantitative and stringent definition of the criteria of Elseviers *et al.* (12,13) using the higher resolution helical CT scanning technique and improved methods of renal volume analysis. The "DeBroe" criteria included bilaterally small kidneys (based on an unusual method of renal size determination that consisted of summing the anterior–posterior and lateral dimensions as seen on the axial CT slice closest to the hilum) plus three or more indentations on the CT slice with the most contour changes and/or calcifications seen along the papillary line. Four (2%) of the 221 patients met these criteria.

For the new criteria (Table 2), size measurements of the healthy volunteers were used to determine size cut points for "small" kidneys in the patients with ESRD, set as 2 SD below the mean values; these were determined separately for men and women. Definitions for calci-

Table 1. Exclusion criteria for incident dialysis patients[a]

Acute renal failure
Diabetes requiring medication
Polycystic kidney disease
Renal failure after a kidney transplantation
Biopsy-proven glomerulonephritis
Amyloidosis
Multiple myeloma
Angiographically proven renal artery stenosis
Obstructive uropathy
Hereditary nephritis
AIDS nephropathy
Sickle cell disease
Toxic nephropathy from antineoplastic agents

[a]CMS form 2728 was used to screen incident dialysis patients with respect to inclusion/exclusion criteria.

Table 2. Criteria for bilateral SICK[a]

Criteria
small: bilaterally decreased kidney size[b] by at least one of three methods
ellipse volume: men <93.2 cm^3; women <73.9 cm^3
voxel volume: men <102 cm^3; women <77.5 cm^3
parenchymal thickness: men <1.44 cm; women <1.27 cm
indented: at least three indentations on each kidney, with at least one major indentation (≥5 mm deep) on one side
calcified: papillary calcifications in each kidney involving at least four papillae or at least 2 mm in size
Classification
grade 1: both kidneys are small and indented
grade 2: both kidneys are small and calcified
grade 3: both kidneys meet all three criteria

[a]SICK, small indented calcified kidneys.

[b]Cut points were calculated using the healthy volunteers ($n = 61$ for parenchymal thickness and ellipse volume, $n = 30$ for voxel volume) reviewed early in the study and are based on 2 SD from the mean value derived from an average of the right and left kidney values. Renal size measurements in healthy individuals were normally distributed.

fications and indentations are as shown. Consonant with the general approach of the DeBroe group, patients with ESRD were classified as having some degree of small indented calcified kidneys (SICK) when they had bilaterally small kidneys and at least one of the other characteristics. Among the 221 patients, 11 were classified as grade 1, three as grade 2, and one as grade 3 SICK, for a total of 15 (7%). Two of the patients who met the DeBroe criteria were classified as grade 2 SICK, and two did not meet the SICK criteria for indentations (although they satisfied the DeBroe criteria in that regard).

Interview

Interviews were conducted by trained medical personnel. Information was obtained on demographic factors; occurrence and timing of predialysis signs and symptoms of renal disease, including albuminuria, abnormal kidney function tests such as elevated blood urea nitrogen or creatinine, hematuria, peripheral edema as a result of kidney disease, nocturia, and when the patient first became aware of his or her kidney disease; other relevant medical history, including hypertension, episodes of kidney stones, gout, and family history of dialysis; lifetime history of analgesic use; consumption of tobacco, alcohol, and caffeinated beverages; and history of employment in occupations that are reported to increase the risk for renal insufficiency. The threshold for recording use of an analgesic drug was that it took place at least once a week for a minimum duration of 1 yr. Details obtained included time started, dose per pill, pills per day, frequency per week, duration, and reason for use. When there were multiple episodes with varying patterns of use for a particular drug, these were recorded separately.

For exploration of the reproducibility of the obtained information, 32 patients were re-interviewed 3 to 4 mo after the first interview, and the responses were compared. Information on demographic and medical details and occupational history was reported with excellent agreement (κ = 0.8 to 1.0 [18]) on both interviews. Information on drug use and beverage consumption was reported with good agreement (κ = 0.6 to 0.8). We conclude that the questionnaire yielded adequately reliable information.

Definition of Index Year and Analgesic Use

Testing the hypothesis requires evaluation of analgesic use that occurred before the onset of renal insufficiency, as the etiologically relevant time is the time of the initial injury. It is not possible to know the date of injury with precision. Although we attempted to obtain information about signs and symptoms of renal disease that would allow the determination of onset for each patient within a year or two, an extensive evaluation suggested that this approach was not reliable. The earliest sign or symptom was <3 yr before dialysis for 113 (51%) patients; this interval likely is too short a time for chronic damage to occur and therefore is not credible. For another 46 (21%) patients, the only reported sign was the date when they became aware that they had renal disease. For 27 (12%) others, only one sign was reported, and we judged that its timing easily could have been misremembered. Therefore, we elected to use a standard interval of 9 yr for determining relevant analgesic use, recognizing that there would be some misclassification of exposure. This interval was derived from the median interval of 7 yr for the 108 patients for whom that interval was at least 3 yr, plus an additional 2 yr to allow for the insidious onset of renal insufficiency. An "index year" 9 yr before the year of first dialysis was set for each patient; analgesic use before this year was considered relevant.

Analgesics were classified as follows: Aspirin- and acetaminophen-containing drugs (including single-ingredient and combination products), phenacetin-containing combination drugs, a composite group ("total analgesics") that contained any of the aforementioned ingredients (also including nonaspirin salicylates), and nonaspirin nonsteroidal anti-inflammatory drugs (NSAID). Because many of the products taken were combinations of multiple ingredients, a specific product could be in more than one group. Use before the index year was quantified in two ways: Total kilograms consumed (pills/d × dose/pill × frequency × duration), with heavy use defined as at least 1 kg; and "density" of use, with heavy use defined as at least 0.3 kg/yr for at least 1 yr. In instances in which 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 that was reported as "low dose" was assumed to be 81 mg. When a patient took multiple drugs in an exposure group, the kilograms for all drugs were summed. For the composite group, the total kilograms of all relevant ingredients (aspirin, other salicylates, acetaminophen, and phenacetin) were summed.

Statistical Analyses

The four patients who were positive according to the DeBroe criteria and the 15 patients with SICK were compared separately with the 204 remaining patients who were negative according to both the DeBroe and the SICK criteria. For the comparison with SICK patients, odds ratios (OR) and exact confidence intervals (CI) (15) were estimated for categories of use of each analgesic group, relative to no use of any drug in that group at any time (because of the possibility that use after the onset of renal disease might exacerbate the clinical course, such use was kept in a separate category and not included in the reference category). The small number of patients with SICK precluded formal control for confounding, and only crude estimates are presented. The "sensitivity" and "specificity" of the SICK criteria as a test for the detection of heavy analgesic use was estimated under the assumption that the interview data accurately reflected use. It should be noted that there is no unequivocal "gold standard" measurement of analgesic use or an agreed-on definition of what constitutes heavy use.

Results

Patients

The demographic characteristics of the patients are shown in Table 3. Approximately half the patients were from the two North Carolina sites; 49% of the patients were at least 65 yr of age, 56% were male, and 50% were black. More than one third had not completed high school.

More than half of the patients with SICK were from one of the North Carolina sites, WFU, compared with 26% of those without SICK. The patients with SICK also seemed to be somewhat older, more likely to be male, and have a higher level of education; they were considerably less likely to be black (20 *versus* 53%). The apparent differences were not tested statistically because of the small numbers of patients with SICK.

CT Scan Appearance

The typical CT appearance of the kidneys from patients with and without SICK are shown in Figure 1.

Analgesic Use among Patients with ESRD and with and without SICK

The prevalence of analgesic use among the four DeBroe-positive patients was 50%, although the small number of such patients precluded a detailed analysis (data not shown). Analgesic use in patients with and without SICK is shown in Table

Table 3. Distribution of 221 ESRD patients according to CT classification and demographic factors[a]

	SICK ($n = 15$)		Other Patients ($n = 204$)		Total ($n = 221$)[b]	
	n	%	*n*	%	*n*	%
Age (yr)						
35 to 49	1	7	57	28	58	26
50 to 64	3	20	51	25	55	25
65 to 74	5	33	51	25	57	26
≥75	6	40	45	22	51	23
Gender						
male	10	67	113	55	124	56
female	5	33	91	45	97	44
Race						
black	3	20	108	53	111	50
other	12	80	96	47	110	50
Educational level						
<8th grade	2	13	26	13	28	13
8th to 11th grade	2	13	49	24	52	24
completed high school	3	20	46	23	49	22
vocational college	0	—	9	4	9	4
some college	5	33	49	24	54	24
completed college	3	20	25	12	29	13
Region						
Chapel Hill, NC (UNC)	3	20	51	25	54	24
Winston-Salem, NC (WFU)	8	53	53	26	61	28
Toledo, OH/Detroit, MI (MCO)	2	13	36	18	39	18
Dallas, TX (DNA)	1	7	49	24	51	23
Portland, OR (OHSU)	1	7	15	7	16	7

[a]DNA, Dallas Nephrology Associates and the University of Texas Southwestern Medical Center; MCO, Medical College of Ohio; OHSU, Oregon Health Sciences University; UNC, University of North Carolina; WFU, Wake Forest University.
[b]Includes two patients who met the DeBroe criteria but not SICK.

4. The prevalence of use was higher among patients with SICK for all categories of aspirin- and acetaminophen-containing products. There were significant associations of SICK with any regular use of aspirin before the index year (OR 8.8; 95% CI 2.4 to 39), use of at least 1 kg before the index year (OR 7.4; 95% CI 1.2 to 43), and use of at least 0.3 kg/yr for at least 1 yr (OR 9.8; 95% CI 1.2 to 66). There also were significant associations of SICK with use of the combined group of analgesic products, including any use before the index year (OR 5.0; 95% CI 1.4 to 22) and use of at least 0.3 kg/yr (OR 8.2; 95% CI 1.5 to 45). Mantel-Haenszel adjustment (16) for age did not materially decrease any of the elevated OR for aspirin or the all-analgesics group, and most were higher than the crude estimates (data not shown).

Among the 15 patients with SICK in the study, only four took at least 1 kg of analgesics before the index year, and all of these patients were from the WFU site in Winston-Salem, NC; four of the five patients who took at least 0.3 kg/yr analgesics were from WFU, and the fifth was from the DNA site in Dallas, TX. None of the OR for acetaminophen-containing products was significant. Approximately two thirds of the patients with SICK were *not* heavy users of acetaminophen- or aspirin-containing products, according to our definitions. With regard to other analgesics, three patients with SICK and one without SICK had taken products that contained phenacetin, all of which also included aspirin, giving a higher prevalence of phenacetin use among the patients with SICK. The prevalence of use of other "mixture" analgesics—products that contain at least two analgesic/antipyretic components (aspirin, other salicylates, or acetaminophen)—was low: None of the patients with SICK and five (2%) other patients had taken at least 1 kg of such products before the index year. There also were very few users of NSAID before the index year and no users of cyclooxygenase-2 inhibitors. The latter drugs were not available before the index year (for the most recently enrolled patients, the index year was 1993).

For exploring the validity of the index interval of 9 yr, a sensitivity analysis was conducted for any regular use of aspirin and acetaminophen products. With a shorter index interval of 7 yr, the prevalence of aspirin use in patients with SICK was unchanged from the main analysis and slightly higher in other patients with ESRD (data not shown). With a longer interval of 11 yr, the prevalence of aspirin use was lower than in the main analysis for both patients with SICK and other patients with

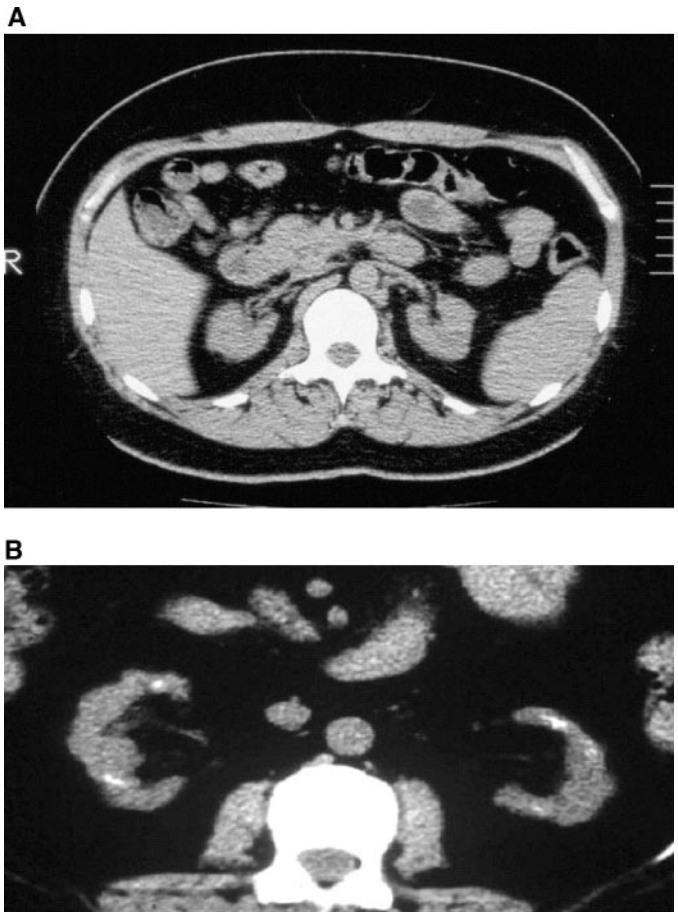

Figure 1. (A) Typical computerized tomography (CT) appearance of a patient without small indented calcified kidneys (SICK) and with ESRD; this pattern was seen in 65% of patients. Note that both kidneys are small and show no parenchymal scars or calcifications. (B) CT findings of a patient with grade 3 SICK. Both kidneys are small and demonstrate bilateral parenchymal scars and coarse papillary calcifications.

ESRD. A similar pattern was observed for acetaminophen. With both of the alternative index intervals, the OR estimates for the two drugs were lower than the estimates in Table 4 as based on the 9-yr interval.

The use of analgesics that contain aspirin or acetaminophen but not phenacetin is shown in Table 5; when users of the last compound are excluded, this leaves 12 patients with SICK and 203 other patients with ESRD for analysis. Overall, there was a significant association with regular use of aspirin-containing products before the index year (OR 6.6; 95% CI 1.6 to 31), and the OR for any regular use of the combined group of analgesics was elevated (3.7) but not statistically significant. There were only two users of acetaminophen among the patients with SICK and too few heavy users of any of the drug groups for OR estimation.

The sensitivity and the specificity of the non–contrast-enhanced CT scan as a test for detecting heavy analgesic use among patients with ESRD are shown in Table 6. The parameters were calculated with two definitions of heavy use: At least 1 kg total exposure before the index year and at least 0.3 kg/yr. For the composite variable of all analgesics, most heavy users did not meet the criteria for SICK, and sensitivity was low, ranging from 16 to 26%; it was even lower for the combined group of analgesics other than phenacetin, ranging from 5 to 13%. There were only two heavy users of phenacetin, both of whom had SICK; this was consistent with a high sensitivity of the CT criteria for this compound. Specificity was high for all categories shown, at 94 to 95%; that is, the large majority of patients who were not heavy analgesic users had negative CT criteria.

Discussion

The results of NANS demonstrate that there is an association between heavy analgesic ingestion and findings of SICK in incident US patients with ESRD that seems to be accounted for mostly by phenacetin-containing products. However, the anatomic changes only occur in a low percentage of heavy users (26% or less overall, 13% or less among those who did not take phenacetin); because most heavy analgesic users have negative (non-SICK) CT scans, the procedure is not a useful tool for identifying such users among the US ESRD population. This finding is important because it follows that, in the United States, the CT scan cannot be used to rule out analgesics as a cause of renal impairment simply because the scan is negative.

Most previous follow-up and case-control studies have shown a significant association between heavy analgesic use and kidney disease (1–9). However, until the work of Elseviers, DeBroe, and colleagues (2,12,13,17,18), this field of investigation has been hindered by the lack of a reliable diagnostic test for the entity. Their studies suggested that the CT scan has high sensitivity and specificity for certain changes, including a reduction in kidney volume, parenchymal scarring, and papillary calcifications. Because there are variations in the types and amounts of analgesics that are consumed from country to country (and within countries), the applicability of these findings to the US population was unknown. For example, Belgian patients with abnormal CT findings frequently had consumed analgesic mixtures (2,18); in this US study, mixture product consumption was uncommon. Of interest is that 10 of the 15 patients with SICK were from the North Carolina sites, with virtually all of the heavy analgesic users from the WFU site. Several of these heavy users had a remote history of phenacetin use.

With regard to specific analgesic ingredients, initial reports linked phenacetin to interstitial nephritis, papillary necrosis, and uroepithelial malignancies (19–25). Phenacetin was removed from the US market in 1983, and it might be expected that the culprit drug thereby was eliminated. However, acetaminophen frequently was substituted for phenacetin in analgesic mixture products and was already a popular single-ingredient drug. Because acetaminophen is the major active metabolite of phenacetin (26–28), concern has lingered over the possibility that its chronic use could incite renal damage.

For addressing various methodological issues, NANS was designed with several unique features. An unequivocal outcome, ESRD, was specified to allow a clear demonstration of an association between analgesic use and any anatomic changes.

Table 4. Regular use[a] of analgesic products with various ingredients among 15 patients with SICK and 204 other patients with ESRD[b]

	SICK		Other Patients		OR (Exact 95% CI)
	n	%	*n*	%	
Aspirin before index year	11	73	37	18	8.8 (2.4 to 39)
≥1 kg lifetime	4	27	16	8	7.4 (1.2 to 43)
≥0.3 kg/yr	3	20	9	4	9.8 (1.2 to 66)
Acetaminophen before index year	5	33	26	13	2.9 (0.7 to 10)
≥1 kg lifetime	2	13	5	2	—
≥0.3 kg/yr	3	20	8	4	5.7 (0.8 to 29)
Phenacetin before index year	3	20	1	0.5	—
≥1 kg lifetime	2	13	0	—	—
≥0.3 kg/yr	0	—	0	—	—
All analgesics[c] before index year	11	73	51	25	5.0 (1.4 to 22)
≥1 kg lifetime	4	27	21	10	4.4 (0.7 to 25)
≥0.3 kg/yr	5	33	14	7	8.2 (1.5 to 45)
Ibuprofen before index year	1	7	13	6	—
Other "traditional" NSAID before index year	0	—	1[d]	0.5	—
COX-2 inhibitors before index year	0	—	0	—	—

[a]Use ≥1 d/wk for ≥1 yr before the index year; reference category for OR is none or less than minimum use at any time (numbers not shown).
[b]CI, confidence interval; COX-2, cyclooxygenase-2; NSAID, nonsteroidal anti-inflammatory drugs; OR, odds ratio.
[c]Any use of products that contained aspirin, acetaminophen, or analgesic mixtures. Dose is the combined dose of all analgesic components, including aspirin, salicylamide, acetaminophen, and phenacetin.
[d]Naproxen.

Table 5. Regular use[a] of aspirin- and acetaminophen-containing products among 12 patients who had SICK and 203 other patients who had ESRD and did not use phenacetin

	SICK		Other Patients		OR (Exact 95% CI)
	n	%	*n*	%	
Aspirin before index year	8	67	36	18	6.6 (1.6 to 31)
≥1 kg lifetime	1	8	15	7	—
≥0.3 kg/yr	1	8	8	4	—
Acetaminophen before index year	2	17	25	12	—
≥1 kg lifetime	0	—	5	2	—
≥0.3 kg/yr	1	8	8	4	—
All analgesics[b] before index year	8	67	50	25	3.7 (0.9 to 17)
≥1 kg lifetime	1	8	20	10	—
≥0.3 kg/yr	2	17	14	7	—

[a]Use ≥1 d/wk for≥1 yr before the index year; reference category for OR is none or less than minimum use at any time (numbers not shown).
[b]Any use of products that contained aspirin, acetaminophen, or analgesic mixtures. Dose is the combined dose of all analgesic components, including aspirin, salicylamide, and acetaminophen.

This was considered a first step; if an association were to be identified, then it would be necessary in subsequent studies to evaluate the relationship at an earlier, less definitive point in the disease process, when intervention could have a clinically relevant effect on outcome. Only incident patients with ESRD were enrolled to avoid inclusion of patients with renal anatomic changes that are associated with dialysis (29). Of the 2300 screened patients with ESRD, 14% were eligible for study; therefore, the patients who had ESRD and were enrolled in NANS were more likely to have primary analgesic-related disease than the general ESRD population. In-person interviews were conducted in all sites by centrally trained personnel. In addition, the analysis was confined to a period that ended 9 yr before dialysis began so that we could approximate use before the onset of chronic kidney disease (CKD) and avoid the problem of including analgesic use that was influenced by the

Table 6. Sensitivity and specificity of SICK in the detection of analgesic-associated ESRD[a]

	SICK	Other Patients	Total	Sensitivity/Specificity (%)
All patients (15 SICK, 204 other ESRD)				
all analgesics[b]				
≥1 kg before index year	**4**	21	**25**	Sensitivity = 16
other[c]	11	**183**	**194**	Specificity = 94
≥0.3 kg/yr before index yr	**5**	14	**19**	Sensitivity = 26
other[c]	10	**190**	**200**	Specificity = 95
Phenacetin				
≥1 kg before index year	**2**	0	**2**	Sensitivity = 100
other[c]	13	**204**	**217**	Specificity = 94
All patients except phenacetin users (12 SICK, 203 other ESRD)				
analgesics other than phenacetin				
≥1 kg before index year	**1**	20	**21**	Sensitivity = 5
other[c]	11	**183**	**194**	Specificity = 94
≥0.3 kg/yr before index yr	2	14	**16**	Sensitivity = 13
other[c]	10	**189**	**200**	Specificity = 95

[a]Assumption: Patients must have taken the specified dose before the index year to be considered analgesic associated. Numbers in bold type are used in calculating sensitivity and specificity.
[b]Dose is the combined dose of all analgesic components, including aspirin, salicylamide, acetaminophen, and phenacetin.
[c]Including lower total doses before the index year, use after the index year only, and no use.

disease (confounding by indication) (30). Finally, emphasis was placed on defining renal anatomic changes in size, parenchymal scarring, and papillary calcifications by contemporary imaging techniques with greater precision than previously possible. This included three determinations of renal volume and a CT study of 61 healthy volunteers of varying age to obtain the distribution of "normal" kidney size.

The study results reveal a significant association between analgesic use and findings of SICK, but this is explained mostly by previous use of phenacetin-containing products. Although phenacetin has not been available in the United States for >20 yr, the long-term drug history that was recorded in NANS extended into the period when it was still in use. When patients who took phenacetin were excluded, there were too few heavy users of the remaining aspirin- and acetaminophen-containing analgesics among the patients with SICK to determine whether an association is present. There was a significant overall association between aspirin use and SICK in non–phenacetin-exposed patients, but this finding is less compelling in the absence of information for heavy use. Regardless of whether there is a relation between currently available analgesics and SICK, however, these results demonstrate convincingly that the CT criteria do not identify most heavy users, thereby answering the main question that NANS was designed to address. Furthermore, the total number of patients who met the criteria for SICK in this study is small in both absolute and relative terms. Overall, 15 patients were identified with SICK, and only four patients met the criteria described by Elseviers, Debroe, and colleagues (2,12,13,17). The patients with SICK represented approximately 7% of the entire series, which in turn represented approximately one seventh of all patients with ESRD. Therefore, the prevalence of SICK among all incident patients with ESRD could be estimated to be 1%. On a population basis, it can be concluded that there are few such patients in the United States.

There are several potential explanations for the lower prevalence of analgesic-related findings on CT scans in our study *versus* the earlier work by Elseviers, DeBroe, and colleagues (2,12,13,17). The ingested drugs differ, with little phenacetin or analgesic mixture use among US patients as compared with Belgian patients. Other potentially relevant factors (*e.g.*, smoking, gout) may differ between the populations, and there also were differences in CT techniques. One-centimeter-slice thickness was used by the DeBroe group *versus* 0.5 cm in NANS. Higher spatial resolution and tissue contrast helical CT technology was not available at the time of the earlier work, thereby limiting the ability to distinguish as accurately between bumpy, contour-altering cysts and true parenchymal scarring. In addition, DeBroe's method of renal size determination depends on a renoform shape, which could under- or overestimate true renal volume.

As with other observational studies, selection and information bias, confounding, and misclassification of exposure must be considered in interpreting our results. The participation rate was 54%, leaving open the potential for selective enrollment; however, although heavy analgesic users may have been more or less likely to participate than other patients with ESRD, the reading of CT scans and determination of SICK was performed blind to exposure and should not have been biased. Given the association that was observed between heavy analgesic use and SICK, there may have been some underascertainment of SICK if analgesic "abusers" tended to refuse to participate more than other patients. Such a dynamic also could explain partly the very low sensitivity of the criteria for identifying analgesic users. This possibility should be kept in mind in interpreting

the results, but it is unlikely to account for the substantial portion of heavy analgesic users who do not meet the SICK criteria.

The interview was designed carefully to maximize the recall of analgesic histories: All interviewers received extensive centralized training in its administration, and there was a high "threshold," at least once a week for at least 1 yr, for recording analgesic use. The re-interview of 32 patients demonstrated that consistent responses were obtained. However, there was no practical way to validate the information, and it is conceivable that incomplete analgesic histories were obtained. Again, however, this is unlikely to have been related to SICK, which had not been determined at the time of the interviews.

Misclassification of exposure could not be excluded; indeed, the use of a common rather than an individually set index date for all patients undoubtedly resulted in some misclassification of analgesic users as nonusers and *vice versa*. A related question concerns the appropriateness of the interval of 9 yr that was used to set the index date. This was based on information about renal signs and symptoms that we judged to be clinically reasonable, and a sensitivity analysis that changed the index date 2 yr in either direction resulted in lower OR estimates. We also did not take body weight into account in determining heavy analgesic use. A cutoff of 1 kg was used to define heavy use for all patients, whereas a lower dose may have been sufficient to constitute heavy use in individuals of low body weight. However, none of the patients with SICK had a lifetime analgesic dose between 0.5 and 0.9 kg. Therefore, any reclassification of heavy use according to body weight would not have altered the conclusion that the non–contrast-enhanced CT scan is not a sensitive tool to detect analgesic-associated kidney injury.

Finally, the small number of heavy analgesic users among the patients with SICK emphasizes the fragility of the findings for individual compounds. It was not feasible to formally control confounding in the comparison between patients with SICK and other patients because of the small numbers in the former group, although it did not seem that the results were explained by the relatively older age distribution of the patients with SICK. Nonetheless, it does not seem plausible that any methodologic issues can explain the very large proportion of heavy analgesic users that do not meet the criteria for SICK, particularly when phenacetin is excluded from consideration.

Conclusion

Our results demonstrate that findings of SICK do not occur with sufficient frequency in US patients with ESRD and heavy analgesic exposure to render the non–contrast-enhanced CT scan a sensitive tool to identify such individuals, as proposed by Elseviers and colleagues (12,13,17,18). Although the conclusion concerning the lack of diagnostic value of the CT scan is clear despite the relatively small numbers, we stress that NANS did not address the overall relationship between analgesics and the progression of renal disease. Future studies are needed to learn whether there is an association in the US population between currently available analgesics and CKD/ESRD in general and whether the ingestion of NSAID or cyclooxygenase-2 inhibitors contribute to the development of CKD or ESRD.

Appendix: Radiology Methods

CT Protocol

Images were obtained from 1 cm above the top of the higher kidney to 1 cm below the lower kidney during a single breath hold. Reconstruction at 5-mm contiguous intervals ("slice thickness") was performed. Two sets of images were obtained. The first used a "soft tissue" window of 250 HU and a level (center) of 30 HU. The second used a "stone" window of 2 HU and a level (center) of 110 HU. This latter parameter ensured that focal kidney densities indeed were calcifications rather than quantum mottle or other artifacts.

Determining Renal Size

The ellipse method uses the following formula: Renal length (cm) × width at hilum × anterior–posterior diameter at the hilum × 0.49 = volume in cm^3 (31). The voxel-count method (modified from Bakker *et al.* [15]) was used later in the study and required that the digital data reflect the true renal volume regardless of unusual renal shapes as a result of masses (mostly cysts) or scars (indentations). A standard volumetric program was applied to the data, which summed voxels in all of the manually outlined 0.5-cm axial slices of each kidney. This method now is considered the "gold standard" for renal volume determination because it avoids the inaccuracies that are introduced by the ellipse method that considers all kidneys to be renoform (ellipsoid). A pilot comparison study demonstrated that the ellipse and voxel volume methods were comparable (within 10% agreement) in renoform kidneys.

Parenchymal thickness was measured by documenting the anterior, posterior, and lateral parenchymal thickness from three adjacent central CT slices (32). Renal pelvis and peripelvic fat were excluded. The nine values (in cm) from each kidney were averaged, resulting in a single parenchymal thickness value for each kidney.

Acknowledgments

This study was sponsored by a contract (RFP 98-03) from the National Institute of Diabetes and Digestive and Kidney Diseases of the National Institutes of Health and by National Institutes of Health grants N01-RR-0-7122 and N01-DK-0-2400.

We thank Kathy Wilber, Dr. Jonathan Loewen, and Lauren Brubaker for assistance with radiology interpretation and Drs. David Warshauer and Wendolyn Hayes for providing radiologic review. Nirma Perricone, Vicki Robins, Heather Main, Pat Lemons, Linda Owens, Mellary Chrisstie, and Emilia Baram performed patient enrollment and interviewed patients and control subjects, and Marilyn Wasti and Moira McDade (Slone Epidemiology Center) performed study coordination. John Roberge and Michael Bairos provided expert computer assistance, and Phyllis Farrell provided expert clerical assistance. We also thank Dr. Josephine Briggs for assistance with the design and performance of the project.

References

1. Dubach UC, Rosner B, Pfister E: Epidemiologic study of abuse of analgesics containing phenacetin. Renal morbidity and mortality (1968–1979). *N Engl J Med* 308: 357–362, 1983
2. Elseviers MM, DeBroe ME: A long-term prospective controlled study of analgesic abuse in Belgium. *Kidney Int* 48: 1912–1919, 1995
3. McCredie M, Steward JH, Mahony JF: Is phenacetin responsible for analgesic nephropathy in New Wales? *Clin Nephrol* 17: 134–140, 1982

4. Murray TG, Stolley PD, Anthony JC, Schinnar R, Hepler-Smith E, Jeffreys JL: Epidemiologic study of regular analgesic use and end-stage renal disease. *Arch Intern Med* 143: 1687–1693, 1983
5. Sandler DP, Weinberg CR: Analgesic use and chronic renal disease. *N Engl J Med* 321: 1126–1127, 1989
6. Pommer W, Glaeske G, Molzahn M: The analgesic problem in the Federal Republic of Germany: Analgesic consumption, frequency of analgesic nephropathy and regional differences. *Clin Nephrol* 26: 273–278, 1986
7. Morlans M, LaPorte J, Vidal X, Cabeza D, Stolley PD: End-stage renal disease and non-narcotic analgesics: A case-control study. *Br J Clin Pharmacol* 30: 713–717, 1990
8. Perneger TV, Whelton PK, Klag MJ: Risk of kidney failure associated with the use of acetaminophen, aspirin, and nonsteroidal anti-inflammatory drugs. *N Engl J Med* 331: 1675–1679, 1994
9. Steenland NK, Thun MJ, Ferguson CW, Port FK: Occupational and other exposures associated with male end-stage renal disease: A case-control study. *Am J Public Health* 80: 153–159, 1990
10. Delzell E, Shapiro S: A review of epidemiologic studies of non-narcotic analgesics and chronic renal disease. *Medicine* 77: 102–121, 1998
11. Rexrode KM, Buring JE, Glynn RJ, Stampfer MJ, Youngman LD, Gaziano JM: Analgesic use and renal function in men. *JAMA* 286: 315–321, 2001
12. Elseviers MM, Bosteels V, Cambier P, De Paepe M, Godon JP, Lins R, Lornoy W, Matthys E, Moeremans C, Roose R, *et al.*: Diagnostic criteria of analgesic nephropathy in patients with end-stage renal failure: Results of the Belgian study. *Nephrol Dial Transplant* 7: 479–486, 1992
13. Elseviers MM, De Schepper A, Corthouts R, Bosmans JL, Cosyn L, Lins RL, Lornoy W, Matthys E, Roose R, Van Caesbroeck D, *et al.*: High diagnostic performance of CT scan for analgesic nephropathy in patients with incipient to severe renal failure. *Kidney Int* 48: 1316–1323, 1995
14. Buckalew VM, Schey HM: Analgesic nephropathy: A significant cause of morbidity in the United States. *Am J Kidney Dis* 7: 164–166, 1986
15. Bakker J, Olree M, Kaatee R, de Lange EE, Moons KG, Beutler JJ, Beek FJ: Renal volume measurements: Accuracy and repeatability of US compared with that of MRI imaging. *Radiology* 211: 623–628, 1999
16. Mantel N, Haenszel W: Statistical aspects of the analysis of data from retrospective studies of disease. *J Natl Cancer Inst* 22: 719–748, 1959
17. Elseviers MM, Waller I, Nenoy D, Levora J, Matousovic K, Tanquerel T, Pommer W, Schwarz A, Keller E, Thieler H, *et al.*: Evaluation of diagnostic criteria for analgesic nephropathy in patients with end-stage renal failure: Results of the ANNE study. Analgesic Nephropathy Network of Europe. *Nephrol Dial Transplant* 10: 808–814, 1995
18. Landis JR, Koch GG: The measurement of observer agreement for categorical data. *Biometrics* 33: 159–174, 1977
19. Spuhler O, Zollinger HU: [Chronic interstitial nephritis.] *Helv Med Acta* 17: 564–567, 1950
20. Mihatsch MJ, Knusli C: Phenacetin abuse and malignant tumors. An autopsy study covering 25 years (1953–1977). *Klin Wochenschr* 60: 1339–1349, 1982
21. Bengtsson U, Angervail L, Ekman H, Lehmann L: Transitional cell tumors of the renal pelvis in analgesic abusers. *Scand J Urol Nephrol* 2: 145–150, 1968
22. Johansson S, Wahlqvist L: Tumors of urinary bladder and ureter associated with abuse of phenacetin-containing analgesics. *Acta Pathol Microbiol Scand [A]* 85: 768–774, 1977
23. Taylor JS. Carcinoma of the urinary tract and analgesic abuse. *Med J Aust* 1: 407–409, 1972
24. Mahony JF, Storey BG, Ibafiez RC, Stewart JH: Analgesic abuse, renal parenchymal disease and carcinoma of the kidney or ureter. *Aust N Z J Med* 7: 463–469, 1977
25. Steffens J, Nagel R: Tumors of the renal pelvis and ureter. Observations in 170 patients. *Br J Urol* 61: 277–283, 1988
26. Smith GE, Griffiths LA: Comparative metabolic studies of phenacetin and structurally related compounds in the rat. *Xenobiotica* 6: 217–236, 1976
27. Brodie BB, Axelrod J: The fate of acetaphenetidin (phenacetin) in man and its metabolites in biological materials. *J Pharmacol Exp Ther* 97: 58–67, 1949
28. Prescott LF, Steel RF, Ferrier WR: The effects of particle size on the absorption of phenacetin in man. A correlation between plasma concentration of phenacetin and effects on the central nervous system. *Clin Pharmacol Ther* 11: 496–500, 1970
29. Lafayette RA, Lai G: Acquired cystic kidney disease in dialysis patients. In *Principles and Practice of Dialysis*, 3rd ed., edited by Henrich WL, Philadelphia, Lippincott Williams & Wilkins, 2003, pp 513–524
30. McLaughlin JK, Lipworth L, Chow W-H, Blot WJ: Analgesic use and chronic renal failure: A critical review of the epidemiologic literature. *Kidney Int* 54: 679–686, 1998
31. Hricak H, Lieto RP: Sonographic determination of renal volume. *Radiology* 148: 311–812, 1983
32. Gourtsoyiannis N, Prassopoulos P, Cavouras D, Pantelidis N: The thickness of the renal parenchyma decreases with age: A CT study of 360 patients. *AJR Am J Roentgenol* 155: 541–544, 1990