The USRDS Role in Enhancing End-Stage Renal Disease Research: Interpretation of Research and Sharing of Analysis Files

Friedrich K. Port, Randall Webb, Robert A. Wolfe, and Philip J. Held

Kidney Epidemiology and Cost Center, USRDS, University of Michigan, Ann Arbor, Michigan

The United States Renal Data System (USRDS) was initiated in 1988. During the past 10 years it has grown to a remarkably complete and validated data system that has become the model and envy for national or regional end-stage renal disease (ESRD) registries.

The National Institute of Diabetes and Digestive and Kidney Diseases (NIDDK), in cooperation with the Health Care Financing Administration (HCFA), established the six goals of the USRDS for the current five-year contract. The first five deal with research and reports, whereas the sixth one deals with supporting other investigators.

1. Design and implement a consolidated renal disease data system that will provide the biostatistical, data management, and analytical expertise necessary to characterize the total renal patient population and to describe the distribution of patients by sociodemographic variables across treatment modalities.
2. Report on the incidence, prevalence, mortality rates, and trends over time of renal disease by primary diagnosis, treatment modality, and other sociodemographic variables.
3. Develop and analyze aggregate data on the effect of various modalities of treatment by disease and patient group categories. These data will be used to analyze the prevention and progression of renal disease with special emphasis on morbidity and mortality.
4. Identify problems and opportunities for more focused special studies of renal research issues currently not addressed by the consolidated data system.
5. Conduct cost effectiveness and other economic studies pertaining to biomedical and epidemiologic aspects of ESRD.
6. Support investigator-initiated research by making data from the database widely available in convenient formats to the biomedical and economic research communities.

This brief review focuses on the sixth goal for the purpose of assisting clinicians and researchers to enhance their use of the USRDS data. In pursuit of this purpose we describe here a) the data sources available to the USRDS and through the USRDS for others, b) more detailed data collections for USRDS special studies on random samples of patients that are also available to researchers, c) examples of the utility of data for answering clinically relevant questions including issues of appropriate study design and interpretation of results, and d) access to USRDS data for clinicians and researchers.

Data Sources for the USRDS

The HCFA provides most of the existing data in the USRDS database. In addition to all the data from its ESRD Program Management and Medical Information System and the Annual Facility Survey, HCFA provides data on transplant follow-up and all Medicare-covered services derived from Medicare paid claims. These HCFA-supplied data are the core of the USRDS database.

Since July 1990, selected data on non-Medicare patients treated by the U.S. Department of Veterans Affairs (DVA) facilities have also been incorporated into the USRDS database. In July 1994, HCFA and the Health Resources Services Administration (HRSA) consolidated transplant data into a single collection, the United Network for Organ Sharing (UNOS), under its contract with HRSA. The expanded transplant data are shared among HRSA, HCFA, and the National Institutes of Health (NIH) and thus are available to the USRDS. The HRSA-collected transplant data include non-Medicare as well as Medicare patients.

In addition, HCFA helps the USRDS with special studies. Most of the new, primary data for special studies are collected through the 18 ESRD networks, which are funded by HCFA under separate contracts. Data from the special studies are fully integrated into the USRDS database. Data collection began in March 1995 for an important new USRDS special study, the Dialysis Morbidity and Mortality Study (DMMS). This special study was drawn from all dialysis units in the United States and obtained data on over 20,000 patients.

Data in the USRDS database collected by HCFA's ESRD networks, federal insurance carriers, and fiscal intermediaries are supplemented by data from the Social

Security Administration, the U.S. Bureau of the Census, local and national ESRD provider databases, and international ESRD registries.

USRDS Special Studies

The USRDS has carried out a number of special studies, most of which have been based on national random samples. Special study topics are approved by NIDDK, with recommendations from HCFA, the USRDS Scientific Advisory Committees, the ESRD networks, and the renal community. For each study, design and sampling plans were developed, samples were selected, and data collection forms and instructions were drafted, tested, finalized, and implemented with assistance from the networks. These studies have resulted in standard analysis files (SAFs) as described below. The data collection forms used for several of the special studies can be found in the appendix of the USRDS Annual Data Report (1).

Examples of USRDS special studies include the Case Mix Severity Study, CAPD and Peritonitis Study, Pediatric Growth and Development Study, Case Mix Adequacy Study, and most recently the DMMS. The latter study was performed in four waves of four mutually exclusive random samples of dialysis units. Thus virtually each dialysis unit was included in only one data collection over a five-year period. To further reduce the burden only reasonably small random samples of patients were chosen in each facility. The DMMS included numerous research topics such as dialyzer reuse, dialysis dose, dialyzer membrane, pre-ESRD care, nutrition, rehabilitation, treatment modality choice (selection), comorbidites, peritoneal dialysis versus hemodialysis outcomes, vascular access, anemia therapy, laboratory data, and medications. Data from the first wave of the DMMS are already available to researchers through an SAF as described below. Much of the most interesting USRDS research comes from these special study data collections.

Research Examples with Emphasis on Interpretation of Observational Results

An almost endless list of examples could be provided, but we will limit this discussion to only a few examples. For additional examples readers are referred to the text of the USRDS 1997 Annual Data Report (1) and the list of USRDS publications that appears in the appendix at the end of the reference tables. Alternatively, this information can be found at the USRDS Web site at http://www.med.umich.edu/usrds.

Growth in incidence of treated ESRD. The number of patients starting ESRD therapy has been growing exponentially in the United States and in other countries (2), with a greater growth for elderly and diabetic patients. This greater growth has been attributed to increasing acceptance of older and sicker patients to ESRD therapy, that is, patients who previously had not been treated for their chronic uremia. However, even for younger adults, the USRDS has documented a similar exponential growth, albeit at a slower rate. An explanation for this observation may be that patients who used to die of

conditions such as cardiovascular complications of diabetes now live longer to reach the end stage of their renal complications. This competing risk may even explain some of the racial differences in incidence for younger age groups (3). Another explanation suggests that there may be more kidney disease in recent years through the growing prevalence of diabetes and the growing use of nonsteroidal anti-inflammatory drugs.

The incidence rates for treated ESRD are closely watched by many in the renal community and the USRDS has been reporting the ESRD incidence counts and adjusted rates for the last nine years. In the last few years there have been some changes in reporting to HCFA, which made analyses of trends more difficult. The most recent USRDS analyses suggest that the counts for 1996 were lower than expected based on the trend line. Statistical analyses indicate that the 1996 counts were barely within the 95% confidence interval (thus not statistically different from the trend line at the 0.05 level) for the projected counts according to trends over the preceding years. Figure 1 shows the actual annual counts of newly treated ESRD incidence and the prediction for 1996. Counts observed in the next few years will shed more light on these trends and will determine if this newest nonsignificant slowing in the growth of new ESRD patients will be confirmed. In the meantime, further evaluation of subgroups and statistical analyses may be used to differentiate random variation from changes in trends.

Dialyzer membrane use and mortality. Based on various observations it was hypothesized that mortality risk among hemodialysis patients varied by the type of membrane used. Dialyzer membranes can be categorized into three groups: unmodified cellulose, modified cellulose, and synthetic membranes. An evaluation of mortality risk was possible in a USRDS special study, which collected data on case mix, including comorbid conditions and delivered dialysis dose. The latter had to be restricted to patients who had been on ESRD therapy for more than one year to minimize the role of residual renal function. It was found to be an important factor since the average dose of dialysis (Kt/V) varied by the type of membrane. A Cox proportional hazards analysis of days to death or censor among hemodialysis patients showed that the risk
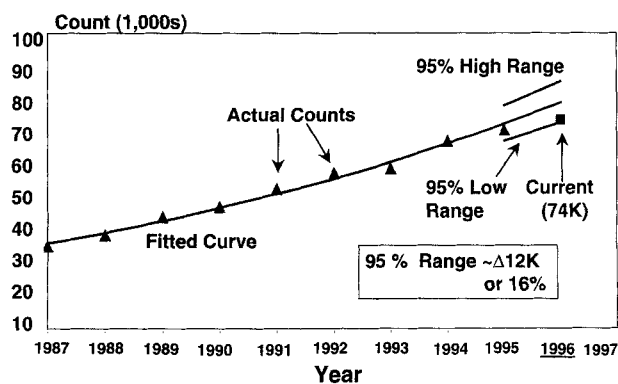

FIG. 1. Incident counts of newly treated ESRD by year with projection to 1996 and actual count for 1996. The 95% confidence band indicates the range projected for 1996 according to counts from previous years.

of death was significantly lower in patients treated with modified cellulose and synthetic membranes compared to unmodified cellulose membranes, all other measured patient factors being equal through statistical adjustment (4). Since membrane use varied by region, an additional analysis also adjusted for region covariables. Since the results remained virtually unchanged, the observed differences in outcome were independent of region, dialysis dose, or comorbid conditions.

The interpretation of these statistically significant findings, however, included several possible mechanisms: a) greater biocompatibility; b) higher flux and thus greater clearance of middle molecules; c) requirement of an ultrafiltration control device, which may allow better attainment of dry weight; d) prescription of more expensive dialyzers by medical teams that provide care that is superior in other respects ("center effect"); and e) other unexamined factors. These possible mechanisms can be explored through additional observational research at the USRDS and through prospective trials. Despite not understanding the reasons, there is little doubt ($p < 0.05$) that patients treated in dialysis units that use different membranes actually had different outcomes. This investigation, then, helps focus future research and likely leads to better outcomes for hemodialysis patients.

Comparative survival by dialytic treatment modality. Patients treated with hemodialysis (HD) can be compared to those treated with peritoneal dialysis (PD). Since PD patients in the US are on average six years younger than HD patients, it is clear that any comparison needs to take such differences into account. Additionally, differences in comorbid status beyond those explained by age do exist and reflect in part patient selection. The findings of these comparisons have shown no differences in outcomes for incident patient groups except for those with diabetic ESRD and among those only over 58 years of age (5). Among prevalent patients (treated with dialysis for various lengths of time), differences favoring HD appeared larger and emphasized differences in older and diabetic subgroups (6). The main point here is again the issue of careful interpretation of the statistically significant results. As shown in Figure 2, there are many potential factors that may contribute to these statistically significant observations (7). To mention only one: Compliance with PD may be poorer than with HD, and results among fully compliant HD and PD patients may give results that are much closer.

Nonproportional hazards. The Cox proportional hazards model is used frequently, but one needs to consider whether it is appropriate for the specific study. An example for nonproportionality is the comparison of dialysis with transplant patient mortality risk. Since transplantation adds a mortality risk due to surgery and the immediate postoperative course, standard Cox analysis restricted to the first few months after transplantation would likely show worse survival for transplant recipients than for dialysis patients, while a long term followup would come to the opposite conclusion. This has been documented by analysis with the nonproportional Cox model that allows the relative mortality risk to vary with time. The risk was more than twofold higher during the first week post transplantation and subsequently markedly lower than for a dialysis group that had been on the transplant waiting list for the same duration (8, 9). There are potential fallacies with the Cox model (10); however, we observe that most errors occur in the design and planning of studies.

Overall, to avoid erroneous or biased results, great caution has to be employed in the design of observational studies and in the use of appropriate adjustments. The second and third examples show that observational research has the limitation that it is difficult to ascribe a clear mechanism to the observations. However, there are numerous advantages to epidemiologic research based on USRDS data, including a) the description of national data and practices, b) the observation of outcomes based on the actual practice rather than the practice at research centers, c) the large volume of studies that can be performed in one year, d) the adjustment for numerous comorbid and other factors in special study data collections, and e) the relatively low cost, when compared to randomized controlled studies that can answer only few questions, usually at very great cost.

Access to USRDS Microlevel Data and Other Statistics

Goal 6 of the USRDS is to support investigator-initiated research by making data from the database widely available in convenient formats to the biomedical and economic research communities. One important means of making data available is through timely responses to data requests made by researchers, practitioners, and other members of the renal community.

As shown in Figure 3, the number of requests for information ("2-hr data requests") has steadily increased over the years. During 1996 an average of one to two requests were filled per working day. Several requests

Outcomes Using the Cox Model

Case Mix

| Age, race, sex ESRD Cause Comorbidity Severity |

+

Treatment

| Dose, prescribed Dose, delivered Compliance Medical Team Practice Patterns |

+

Modality
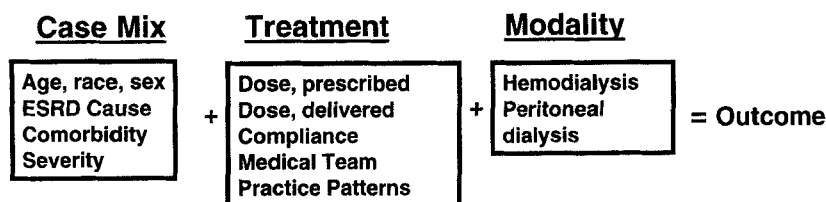
| Hemodialysis Peritoneal dialysis |

= Outcome

FIG. 2. Potential contributing factors for observed results in comparisons between peritoneal and hemodialysis patients. (From Ref. 7, used with permission.)
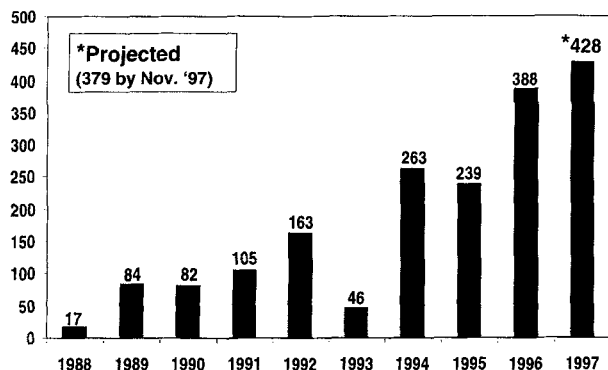
Data Requests Processed 1988-1997

FIG. 3. Number of data requests processed by the USRDS Coordinating Center by year, 1988–97.

were filled regarding specific analysis files and several have led to scientific publications. Requests for statistics not available in the Annual Data Reports but that would require 2 hr or less of computer programmer/analyst time can be provided by the Coordinating Center, usually within one week of the request. Requests that require more than 2 hr of computer programmer/analyst time will be undertaken only upon written approval by the NIDDK project officer. Many of these more complex data requests can be met at low cost by obtaining USRDS SAFs or custom data files. Both of these files contain patient-level data.

The USRDS has been sharing research data with other researchers for many years with the approval from the NIH/NIDDK project officer. Figure 4 shows the number as well as the size of requests for data files for researchers outside of the USRDS that were filled from 1992 to November 1997, including pending requests. The figure shows the number of CD-ROMs provided to researchers by calendar year. Note that over 600 high-density floppy disks fit on one CD. Thus, there has been a clear increase in both the number of researchers and in the amount of information released for research during the recent years.

The SAF make the USRDS database available to researchers in an easy-to-use and well-documented format. This approach allowed a major reduction in the production costs and thus a cost saving for researchers. These analysis files have patient-specific information, although patient identifiers and facility identifiers are encrypted. The SAFs are governed by the USRDS Policy on Data Release for Investigator-Initiated Research, and they require that the research investigator's proposal be approved by NIDDK and that the researcher sign the USRDS Agreement for Release of Data, which means the researcher agrees to observe the prescribed restrictions. These documents and further discussion of the SAFs can be found in the USRDS Annual Data Report.

There has also been a marked growth in the amount of data provided. All requests except custom requests include the core SAF CD. This CD includes basic patient data, each patient's treatment history, full transplant and transplant follow-up data, and all data from the USRDS special studies. The hospitalization CD includes data about hospital inpatient stays except for payment data items. This file is too large to be included on the core CD. Historically about 40% of researchers have requested this file in addition to the core CD. The Medicare payment SAFs are used by only about 20% of the researchers, but these files account for 70% of the CDs provided. A full set of Medicare payment files requires 36 CDs.

CD-ROM technology has been crucial to the growth in the use and usability of the USRDS SAFs. Half of the researchers obtaining patient-level microlevel data files have needed only one CD to carry out their research. The full Medicare payment data would require 120 nine-track magnetic tapes instead of 36 CDs. Providing this volume of data on tape would not be practical.

A wide variety of research topics have been addressed by researchers outside the USRDS using the USRDS SAFs. The reference tables in the annual data reports also are used extensively. In fact about half of recent American Society of Nephrology abstracts, which were based

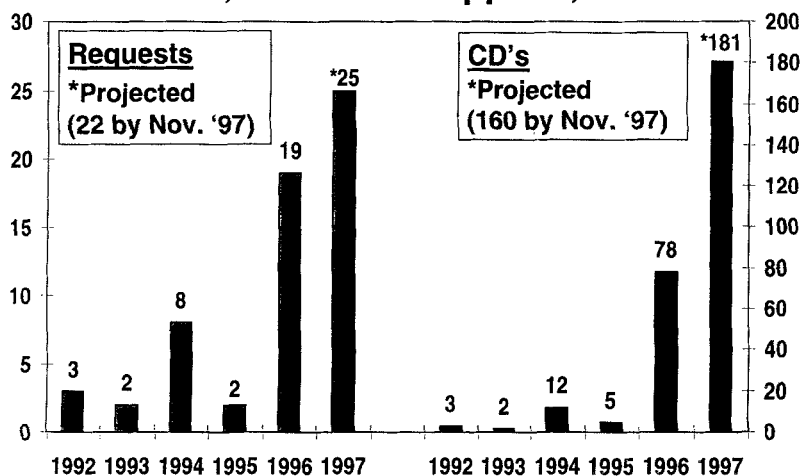## Researcher Requests for USRDS Patient Level Data Files, and CD's Supplied, 1992-1997

FIG. 4. Researcher requests for USRDS patient-level data files and number of CDs supplied by year, 1992–97. Prior to early 1996, the numbers indicate the files supplied on nine-track magnetic tape. For 1997, the projected total requests CDs are shown.

on USRDS data, used data from the annual data report reference tables rather than the SAFs.

Conclusion

The USRDS provides a wealth of information for clinicians and researchers. Specifically, it serves clinicians through analyses of a) national trends in number of treated patients, treatment, and outcome, b) patient care issues such as vascular access, patient compliance, iron therapy, and dose of dialysis, c) patient issues such as quality of life and access to care, d) facility-specific outcomes such as standardized mortality, hospitalization, and transplantation with national and regional standards. Additionally, the USRDS is a resource for inquiries and serves as a national reference. The USRDS also serves the researcher by making data and analysis files available a) through the World Wide Web (site noted above), b) through low-cost SAFs for special studies, and c) through tailored research files. Although many observational studies using USRDS data provide useful information, there has to be vigilance and experience to avoid potential biases. Some of the USRDS studies have led to new questions and hypotheses that need to be answered by more detailed prospective trials. Members of the USRDS Coordinating Center are available for research questions and advice. Thus the USRDS serves numerous needs and has helped to improve patient care and to provide large research data sets to numerous research teams. It is our hope that this service to the clinical and scientific renal community will continue in the future.

References

1. U.S. Renal Data System (USRDS): USRDS 1997 Annual Data Report. Bethesda, MD, The National Institutes of Health, National Institute of Diabetes and Digestive and Kidney Diseases, 1997
2. Port FK: End-stage renal disease: Magnitude of the problem, prognosis of future trends and possible solutions. Kidney Int 48(Suppl 50):S3-S6, 1995
3. Lopez AAL, Port FK: Differences in the patterns of age-specific black/white comparisons between end-stage renal disease attributed and not attributed to diabetes. J Am Soc Nephrol 25:714-721, 1995
4. Hakim RM, Held PJ, Stannard D, Wolfe R, Port FK, Daugirdas JT, Agodoa L: The effect of the dialysis membrane on mortality of chronic hemodialysis patients. Kidney Int 50:566-570, 1996
5. Held PJ, Port FK, Turenne MN, Gaylin DS, Hamburger RJ, Wolfe RA: Continuous ambulatory peritoneal dialysis and hemodialysis: A comparison of patient mortality with adjustment for comorbid conditions. Kidney Int 45:1163-1169, 1994
6. Bloembergen WE, Port FK, Mauger EA, Wolfe RA: A comparison of mortality between patients treated with hemodialysis and peritoneal dialysis. J Am Soc Nephrol 6:177-183, 1995
7. Port FK, Wolfe RA, Bloembergen WE, Held PJ, Young EW: The study of outcomes for CAPD versus hemodialysis patients. Perit Dial Int 16:628-633, 1996
8. Port FK, Wolfe RA, Mauger EA, Berling DP, Jiang K: Comparison of survival probabilities for dialysis patients versus cadaveric renal transplant recipients. JAMA 270:1339-1343, 1993
9. Wolfe RA, Ashby VB, Milford EL, Ojo AO, Ettenger RE, Agodoa LYC, Held PJ, Port FK: Patient survival for wait-listed (WL) dialysis versus cadaveric renal transplant (tx) patients in the U.S. J Am Soc Nephrol 8:708A, 1997
10. Wolfe RA, Strawderman RL: Logical and statistical fallacies in the use of the Cox regression model. Am J Kidney Dis 27:124-129, 1996

Note added in proof

The estimates in Figure 1 likely represent undercounts for 1996. More recent estimates (Jan. 1998) suggest that counts for 1996 are closer to the projected line.