

Chapter XII

Analytical Methods

This chapter discusses the analytical methods used in this *1996 Annual Data Report (ADR)*, and the differences between this ADR and the *1995 ADR*. The *Researcher's Guide to the USRDS Database* provides additional detail about the database and the Standard Analysis Files (See Chapter I).

What's New?

This section discusses differences between this *1996 ADR* and the *1995 ADR*, as well as areas where methodologies finally have stopped changing. It also discusses some continuing problem areas which need to be kept in mind.

1) New Data. Data are now reported through 1993 for detailed patient data and through 1994 for provider-based data. There are also changes in data which apply to years reported in past ADRs. This includes data which are delayed in reaching the USRDS database, and new sources of data, such as increased reporting by Department of Veterans Affairs facilities. See "New Data in the Database" below.

The largest change is the addition of Medicare claims data. The USRDS currently has data on institutional claims (hospital inpatient and outpatient services, outpatient dialysis, skilled nursing facilities, hospice, and home health agencies) for 1989 through 1994 and physician/supplier data for 1991 through 1993. Section K of the Reference Tables presents data on Medicare payments for 1991 through 1993. See "Medicare Claims Data" below.

2) Slight undercount of 1993 Incidence. The incidence counts for 1993 are somewhat low due apparently to some Medical Evidence forms which did not get processed. The HCFA and USRDS ESRD data systems currently are in a transition period. We expect that the data for 1994, which will be reported in the 1997 ADR, will be more complete

and that some of the missing data for 1993 will be recovered.

3) Lost to Followup Methodology. Some changes were made to the methodology for classifying patients as lost to followup. These changes delay the start of lost to followup status, resulting in counts of prevalent patients which are about 2 percent higher and death rates which are correspondingly lower than with the old methodology.

4) Death Notification Form. A new ESRD Death Notification Form (HCFA-2746) came into use during 1990. More detailed cause of death categories are used on the new form. On the old form, withdrawal from dialysis was one of the possible causes of death. On the new form, whether ESRD treatment was stopped prior to death is asked separately from the cause of death question, and withdrawal from dialysis is no longer in the list of causes of death.

Because of these changes, the cause of death tables in Section D of the Reference Tables use only the new codes. In the *1995 ADR*, these tables used a 2 year period. Since we now have three full years of data using the new forms, this *1996 ADR* is able to use 3 years for these tables as is done for the other death rate tables.

5) Death and hospitalization rate methodology. This year the news is that there have been no changes to the annual death rate methodology. The hospitalization rate methodology has been revised to more exactly parallel the death rate methodology. Specifically, incident patients and all types of dialysis patients are included in the analysis.

6) Standard population for adjustment. The adjusted incidence, prevalence, survival, and death rates in this ADR now all use 1992, the year before the most recently reported year, as the standard

population for adjustment. This is consistent with the 1995 ADR, for which 1991 was used as the standard population for adjustment.

7) State and Network Rates. In the 1995 ADR, we began presenting incidence and prevalence rates by state and by ESRD Network for the past 10 years rather than just for the most recent year as in earlier reports. These rates show some instability from year to year which means that these rates must be interpreted with caution.

8) Race code. As discussed in the 1995 ADR, the race codes now in use by the Medicare system include a Hispanic category. This creates an inconsistency with the race categories reported on the ESRD Medical Evidence Form. This will result in a slight increase in the number of patients classified as "Other" race.

9) Minimum Aggregation Size. Starting with the 1994 ADR, in the Reference Tables and in the graphics and tables in the text, aggregations of less than 10 persons generally are not reported. This is consistent with the restriction imposed by Section F of the USRDS Agreement for Release of Data.

This restriction is imposed to insure the confidentiality of patient-specific data. In a cross-tabulation of patients by patient characteristics, if an individual cell in the table represents only one person, then it is conceivable that one could identify that person and learn something about that person. By restricting cell sizes to 10 or more persons, such identification becomes virtually impossible. The general convention in the Reference Tables is to replace the number with an asterisk (*) in cells which are based on less than 10 patients.

The above changes in data, methodologies, and methods mean that care must be used when comparing data between ADRs for different years. However, this ADR follows the USRDS policy that each ADR present data for a series of years using a consistent set of definitions so that the reader can make valid comparisons among years within that report.

New Data in the Database

With each ADR, the counts of new patients by year increase somewhat for years which were reported in previous ADRs. The counts for the year before the ADR year (1995 for this 1996) ADR are not reported because these counts show changes on the order of 8 percent to 14 percent by the time of the

next ADR. See the discussion of the "15 Month Rule" below. The incidence counts for the most recent year reported (1993 for this 1996 ADR) can be expected to increase by about 2 percent in the next ADR and can be expected to increase by 5 percent over the 8 years starting with the ADR in which the year is first reported.

The increases in incidence counts for past years from update to update may be due to at least the following factors:

Delays in receiving data. This is probably the main factor in the most recent years. The USRDS generally waits 15 months before reporting data for a given time period because of this problem.

An example of a delay in receiving data is a patient who is covered under private health insurance and for whom Medicare would be the secondary payer for the first 21 months. If the private health insurance covers all of the dialysis costs, then the Medical Evidence Form might not be filed until a year or more after the start of ESRD. Since there may be no Medicare claims, the patient will not appear in the USRDS database until the Medical Evidence Form is filed. When the form is filed, the date of start of chronic renal dialysis entered on the form should be the date dialysis started, even though it was not covered by Medicare at that point. Thus this patient appears as an incident case in the earlier year rather than the year in which the form was filed.

New data sources. The data starting with the 1993 ADR include Department of Veterans Affairs (DVA) patients being reported to HCFA for the first time. The HCFA and USRDS ESRD data systems are currently in a transition period and will incorporate data from the ESRD Network data systems and from the Medicare claims data. These new sources will contribute to more complete counts in the 1997 ADR.

Delays in Medicare eligibility. If a patient is covered by some other insurance (such as DVA) for four years and then becomes eligible for Medicare and has a Medical Evidence Form filed, his/her first service date would be four years prior to appearing in the file.

Errors in first service date. The physician completing the Medical Evidence Form may report the first dialysis ever rather than the start of maintenance dialysis. There may be clerical or keypunch errors in the year of the first dialysis date. The software for determining first service date may

be wrong. Thus, patients added in a given update can appear as incident in a much earlier year. If a first service date is corrected, then a patient may appear as incident in a different year than in a previous database update.

The “Fifteen Month Rule”

The USRDS generally does not report data for periods later than 15 months prior to the update date. Because of delays in the submission and processing of Medicare bills, these data generally are not considered to be complete until 15 months have passed. Thus, although we have data through June 1995, the data are only considered complete through March 1994, and this ADR reports results only through December 31, 1993.

The USRDS CC is currently evaluating this rule. Some analyses in this ADR use data from 1993. These analyses are labeled as preliminary and should be treated with caution.

Slight Undercount in Incidence for 1993

As discussed in Chapter II and reported in the Reference Tables (A.1, for example), the number of incident Medicare ESRD patients reported in this ADR is only 2.0 percent higher for 1993 (57,384) than for 1992 (56,243). This very modest increase is in contrast to steady growth of nearly 10 percent per year in Medicare incidence counts between 1984 and 1992. Investigations by both HCFA and the USRDS have confirmed that this is largely the result of a sharp reduction in the expected number of Medical Evidence forms filed with HCFA during the third quarter of 1993. It does not appear to mark a permanent reduction in the growth in incidence.

A Medical Evidence form is filed with HCFA for all patients receiving dialysis or a kidney transplant under the Medicare program. The USRDS generally uses the information on this form to determine when each patient began treatment for ESRD. However, a far lower than expected number of Medical Evidence forms were filed with HCFA during the third quarter of 1993. Further, this pattern is not limited to patients of a particular age, gender or race, or living in a particular geographic region.

The USRDS has traditionally relied on other sources of information to identify ESRD patients for whom no Medical Evidence form was filed (typically 4-5 percent of new patients). For some patients the

start of dialysis is determined based on the first dialysis bill paid by Medicare or on the date the patient was first eligible for Medicare because of ESRD. These patients are included in the measurements of both incidence and prevalence for 1993, and are reflected in the ‘missing diagnosis’ category since the Medical Evidence form is our only source of information on primary diagnosis. However, even after including these patients there remains a far lower than expected number of incident patients during the third quarter of 1993.

A separate measurement of the incidence for 1993 is available from the Annual Facility Survey (AFS), which includes both Medicare and non-Medicare patients. Based on 6.0 percent growth in incidence from the AFS for 1992-93, we would expect over 59,600 new Medicare patients during 1993, or approximately 2,200 more patients than reported in this ADR. Despite considerable efforts by both HCFA and the USRDS to investigate this matter, there are currently no clear explanations for this decline.

We expect that we will ultimately be able to identify many if not most of the 1993 incident Medicare patients for whom no Medical Evidence form has been filed and are not currently captured by the USRDS database. Patients with Medicare as a secondary insurance payer for up to 21 months of ESRD may be identified on the basis of dialysis claims or a delayed Medical Evidence form. In addition, patients who may have died before Medicare started paying for dialysis may have an ESRD Death Notification form filed with HCFA.

Preliminary data for 1994 do not indicate a similar decline in Medical Evidence forms filed during the third quarter, and suggest that the incidence counts are much closer (though still somewhat less) to what we would expect based on previous trends in incidence.

The USRDS also expects slight discontinuities in 1994-95, which represents a transition period for the HCFA data system. The HCFA is completing a transition from the way data are stored in the Program Management and Medical Information System (PMMIS), replacing its Medicare ESRD Support Subsystem (MESS) with an enhanced on-line data system known as the Renal Beneficiary and Utilization System (REBUS). This new data system is expected to be more complete and include more detailed information about ESRD patients. However, data for 1994 and 1995 may not be entirely consistent with prior years because of the transition to this new

data system. This issue will be addressed in further detail in the 1997 ADR.

Medicare Claims Data

One of the biggest changes now taking place in the USRDS database is the addition of Medicare claims data to the database and its use in ADR analyses. The claims data for 1989-1994 require about 16 gigabytes (billions of bytes) of disk storage, an eight-fold increase over the prior requirements.

Portions of the USRDS database, particularly the quarterly dialysis and hospital inpatient stay records have always been derived from Medicare claims data. Now the USRDS is getting data directly from the HCFA Medicare claims files. This has a number of advantages for the USRDS:

- Data about all claims, not just dialysis and inpatient stays
- Detail to the level of individual diagnoses and procedures
- Dollar amounts billed and paid
- Ability to determine more precisely when changes in dialysis modality occurred.

Section K of the Reference Tables presents extensive data about Medicare payments and is the basis for the analyses in Chapter IX.

The Medicare Claims data are obtained from HCFA's Standard Analysis Files (SAFs). The SAFs are created on a calendar year basis six months following the end of the calendar year. HCFA estimates that the SAFs are 98 percent complete. The USRDS obtained institutional claims data (hospital inpatient and outpatient facilities, outpatient dialysis facilities, skilled nursing facilities, hospice facilities, and home health agencies) for 1989 through 1994. In addition, physician and supplier claims data for 1991 through 1993 were obtained.

Data were obtained for all patient ID numbers currently in the USRDS database plus all ID numbers which had outpatient dialysis claims or a kidney transplant during 1989-1994. The HCFA enrollment database was used to gather all of the cross-reference ID numbers for these IDs, that is all of the ID numbers under which these patients may have had claims. Patients who are not currently in the PMMIS were not included in the analyses in this ADR.

The USRDS CC is working closely with staff at HCFA on the use of these data and on the

development of analysis files which will be of maximum benefit to the research community.

Lost to Followup Methodology

There frequently are gaps in the dialysis billing data upon which the dialysis modality periods are based. Our convention is to assume that a treatment modality continues until the next modality-determining event. For a patient with a functioning transplant, the functioning transplant modality continues until a transplant failure is encountered. However, a dialysis modality is continued only for 365 days in the absence of dialysis bills or other confirmation that dialysis is continuing. After this 365 day period, the patient is counted as lost to followup until dialysis bills resume. The exception to this 365 day rule is that lost to followup does not begin until the second year of ESRD.

With the 1996 ADR, a change was made in the lost to followup methodology. Previously, the lost to followup period began at the beginning, rather than at the end, of a 365 day period with no bills. This change produces aggregate results which are more stable from one ADR to the next. The second change is that lost to followup cannot begin until the end of the second year of ESRD, where previously it could begin during the second year. The rationale for this first two years rule is that Medicare may be the secondary payer for up to the first 21 months of ESRD, delaying the appearance of Medicare claims.

The effect of these changes is to delay the start of lost to followup status. This, in turn, causes the counts of prevalent patients to increase by about two percent and death rates to decrease slightly.

Death Notification Form

In March of 1990, a new ESRD Death Notification Form (HCFA-2746) came into use. A few deaths from 1989 were reported in 1990 using the new form. Both forms were in use in 1990. After 1990 only a few old forms were filed.

The new form uses a new set of codes and categories for cause of death. Another substantial change is that withdrawal from dialysis is no longer a cause of death, but a separate question asks whether the ESRD treatment was stopped before the patient died and the reason for such stoppage.

In the 1993 and 1994 ADRs, the cause of death categories from the new form were recoded into the categories used on the old form. Patients who

stopped treatment before death were recoded into the withdrawal from dialysis category. This resulted in an increase in the withdrawal category from about 10 percent of the total in 1989 (old form) to about 16 percent in 1991 (new form). It seems likely that most of this change is due to the change in coding. Comparisons of cause of death before and after 1990 must be treated with caution.

For this reason the 1995 ADR reported cause-specific death rates only for the 1991-1992 period, during which only the new form was in use. This 1996 ADR reports cause specific death rates for 1991-1993.

Geography

Starting with the 1993 ADR, Puerto Rico and United States Territories have been reported separately from the 50 states and the District of Columbia in the incidence and prevalence Reference Tables (Tables A.1, A.3, B.1, B.4).

The population base used in computing incidence and prevalence rates is United States resident population, which includes only the 50 states and D.C. For Puerto Rico and United States Territories, only counts are reported, and rates are not computed. Puerto Rico, the Virgin Islands, and the Pacific Territories are not included in computing the incidence and prevalence rates in Sections A and B for the New Jersey and Northern California ESRD Networks, respectively. In Reference Tables sections other than A and B, all patients are counted, and tables by Network include Puerto Rico, the Virgin Islands, and the Pacific Islands.

Measurement Concepts: Incidence, Acceptance, and Prevalence

Incidence and prevalence are two concepts used in measuring the extent of a disease. The incidence of a disease is the number of persons diagnosed with that disease in a given time period, typically a year. Prevalence is the number of persons who have that disease at a given point in time (point prevalence) or who have the disease during a given time period (period prevalence). Thus, annual period prevalence consists of persons who have the disease at the end of the year plus those who had the disease at some time during the year and died before the end of the year. The prevalence measures combine the effects of those

who get the disease (incidence) and those who die (mortality).

Note that the USRDS considers successful transplantation as a therapy rather than a "recovery" from ESRD.

The terms "incidence" and "prevalence" of ESRD imply that information is available about all patients with ESRD, but, as discussed earlier, this is not the case. Actually, the data are for persons receiving ESRD therapy as reported through HCFA and do not include patients who die of ESRD before receiving treatment and those who are not reported through HCFA. For these reasons, the terms incidence and prevalence are qualified as incidence and prevalence of reported ESRD therapy.

As discussed earlier, patients are classified as lost to followup if one year has passed with no dialysis, death, or transplant data. Starting with the 1992 ADR, these patients are not included in the point prevalence counts. The Reference Tables show the number of patients not counted for this reason.

The term "acceptance into ESRD therapy" is used by some other ESRD registries, such as the European Dialysis and Transplantation Association. The USRDS chose not to use this term because "incidence of reported ESRD therapy" is more precise, and because "acceptance" implies that the remaining patients are rejected, when in fact they may be unidentified as ESRD cases or unreported through HCFA.

Incidence tends to be a more useful measure for medical and epidemiological research which examines the causes of disease and the differences in how sub-populations are affected by these causes. Point prevalence may be more useful for public health research, since it measures the overall burden of the disease on the health care delivery system. Period prevalence may be more useful for cost analysis, since it indicates the total burden of the disease over the course of the year.

Sections A and B of the Reference Tables present parallel sets of counts and rates for incidence and December 31 point prevalence. Section B also presents annual period prevalence counts (but not rates). This de-emphasis of period prevalence reflects our experience over the past year with the general lack of demand for more than overall counts of period prevalence.

Census Population Base

Incidence and prevalence counts indicate the absolute level of occurrence of ESRD as defined by its treatment. Incidence and prevalence rates per million population indicate the level of the disease relative to the population and provide a convenient method of comparing the occurrence of the disease over time, among population subgroups, and with other diseases. This report generally uses rates per million population, although for some tables rates per ten million population are used. The size of the population base is chosen largely to produce rates with three significant digits to the left of the decimal point. Rates for diseases which occur more frequently than ESRD might use a smaller population base size.

Note that rates for a specific age/gender/race population group use that group for both the numerator and denominator of the rate. Thus the incidence rate for 40-45 year old White males uses the Census population counts for 40-45 year old White males as the base. Rates by primary disease causing ESRD for 40-45 year old White males also use the Census population counts for 40-45 year old White males as the base.

The population base for computing incidence rates is the total United States resident population on July 1 of each year by race, gender, and age. For point prevalence rates, the population on December 31 is used. Race is divided among four race groups: Black, White, Asian, and Native American. Age is divided among eighteen five-year age groups: the first group includes persons who are 0-4 years of age, while the 18th group includes persons who are 85 years of age and older. For this ADR, rates are reported for 1984 through 1993. Section J of the Reference Tables shows the July 1 population counts. The December 31 counts were computed as the midpoint between the July 1 counts for adjacent years.

Race Classifications

The race codes in the USRDS database classify patients into the categories of Native American/Alaskan Native, Asian/Pacific Islander, Black, White, other, and unknown. As mentioned earlier, the Native American and Asian categories became available only with patients starting ESRD therapy in 1982. Since 1982, the proportion of new patients in the Other and Unknown categories has remained under two percent.

Incidence and prevalence rates are computed and adjusted by race using only the following categories: Black, White, Asian, and Native American. Only counts are reported for the "other" and "unknown" categories, and these categories are dropped from all tables which report rates by race. In addition, rates by race are only reported for years starting with 1982 because of the change in the reporting of race that year.

Only four categories are used in computing incidence and prevalence rates because the Census population estimates used in computing the rates place the entire population into four categories and do not include a residual "other" or an "unknown" category. Since almost two percent of the patients in the USRDS database fall into the "other" and "unknown" categories, the rates reported in this report will be biased slightly low. If bias is uniform across races, then comparisons of rates by race are not significantly biased.

The USRDS is not able to produce tabulations for the Hispanic population because it does not have this data item for all patients. An item on Hispanic ethnicity will appear on the new HCFA Medical Evidence Form, which will go into use in 1995, so that this item will be available for most patients starting ESRD in 1995 or later.

As mentioned earlier, the race codes now in use by the Social Security Administration include codes for Asian American, Native American, and Hispanic. The treatment of Hispanic as a race category is inconsistent with the classifications used by the Census Bureau. In Census tabulations, Hispanic is a category separate from race.

The race recorded in the Social Security records is one of the sources of race in the USRDS database, along with the Medical Evidence Form. The effect of this change in the coding of race will be a small increase in the number of patients identified as Asian, Native American, and other. Patients with Hispanic as the Social Security race code are recoded as "other" if race from the Medical Evidence Form is not available.

Adjustment and Standardization of Rates

Adjustment or standardization of rates is a method for comparing rates between different sub populations or time periods to remove the effect of differences in the composition of the populations being compared.

For instance, Blacks with ESRD have a younger age distribution than do Whites. One way to remove the effect of this age difference on observed rates is to compare the rates for specific age groups between Blacks and Whites. Age adjustment provides a method of obtaining summary rates (across age groups) for Blacks and Whites with the effect of age differences removed.

Age adjusted rates indicate what the rates for the two groups, for example Blacks and Whites, would have been if the age specific rates for each age group had occurred in the same standard population. Age adjustment operates by taking the age specific rates for Blacks, applying them to a standard population distribution stratified by age, summing the resulting ESRD population, and computing a new overall rate per million population using the total of the standard population. The same procedure is followed for Whites. After age adjustment, the overall 1986 incidence rate for Blacks changes from 300 to 373, and that for Whites changes from 105 to 99. Blacks have higher rates of ESRD than Whites at almost every age group, but the fact that the Black population has a younger age distribution masks some of this difference in the across age unadjusted rates, since there are relatively fewer patients in the younger age groups.

It is now USRDS practice to use the relevant population from the year before the most recent reported year as the standard population for adjustment of rates. For adjustment of incidence and prevalence rates in Sections A and B of the Reference Tables (1984-1993), the standard population is the July 1, 1992, United States resident population by five-year age groups, four race categories, and gender. For adjustment of patient survival rates in Section E, the standard population is the 1992 incident ESRD population. For adjustment of graft survival rates in Section G, the standard population is all patients transplanted in 1991 through 1992.

The adjustment method described here is the "direct standardization" method and is the method used through most of this report. For the rates by state and ESRD Network in Sections A and B of the Reference Tables, the "indirect standardization" method is used because of problems with small population counts (Fleiss).

In addition to adjustment by age, the tables in this report frequently adjust by race and gender or appropriate combinations of the three factors. For instance, in one of the standard table formats, rates by age are adjusted by race and gender, rates by race are

adjusted by age and gender, rates by gender are adjusted by age and race, and rates by primary diagnosis and overall rates by year are adjusted by age, race, and gender. Adjustment for multiple factors requires stratifying the populations by the multiple factors rather than just by age. Such adjustments are particularly useful when comparing rates over time to control for the changing characteristics of the ESRD population. Most tables showing adjusted rates are followed by a parallel table showing standard errors of the adjusted rates.

It should be noted that the process of standardization across age, race, and gender produces rates that are different than those encountered if considering one rate separately. For example, someone familiar with the total ESRD incidence rate for 1987 would find the adjusted rate (135/million population, see Reference Table A.6) to be lower than the real or "crude" rate (139/million). This difference does not indicate that the standardized rate is incorrect; rather, the standardized rate is "adjusted" so that rates for 1987 can be compared to rates for all of the other years (1982-90), regardless of changing age, race, and gender compositions of the ESRD population during these years.

Thus, while standardized rates are not "real" in one sense, they are the most useful measures for comparing rates across strata or years. If one is interested in the actual rate for any one year, then the unadjusted rates should be used.

Annual Death Rate and Hospitalization Methodologies

Section D of the Reference Tables presents death counts and death rates expressed as the number of ESRD patient deaths per 1,000 patient years at risk. Similarly, Section H presents hospitalization rates based on the number of first admissions, total admissions and days in the hospital per year at risk. Both sets of rates are computed based on calendar year periods. In order to increase the stability of the estimated rates, three years (1991-1993) are pooled for the tables. Rates are presented by age, race, gender, and diabetic status.

This rate methodology has been used by the USRDS since 1991. It was designed to be used by ESRD Networks and dialysis and transplant facilities based on data about a group of patients. These rate tables can be used to compare local ESRD mortality and hospitalization rates to national rates. (Wolfe) The time period for the local data does not need to be

a calendar year and can be more or less than one year in length as long as the period at risk is measured in units of years and all other aspects of the methodology are followed.

In response to suggestions from users of this methodology, some major changes in computing death rates were made starting with the 1994 ADR: patients starting ESRD therapy during the year were included, as well as previously transplanted dialysis patients. In the past, we have not been able to parallel these changes in the hospitalization rates due to the incompleteness of the data caused by the delay in reporting patient ESRD status in the PMMIS file. However, with usage of the HCFA Standard Analysis Files beginning this year, this problem with the reporting delay is alleviated.

However, there are still some other potential problems with hospitalization data, mostly involving patients for whom Medicare is likely to be the secondary payer. Patients in the first 18 months of Medicare eligibility may have their hospital stays covered by other insurance. During this period, hospitalization data are not expected to be complete. Similarly, some transplant patients lose Medicare eligibility after three years of a successful transplant. The procedures used in these analyses attempt to screen out these patients by requiring that Medicare dialysis bills be received before a patient is eligible for entry into the study.

Incident Patients. The original methodology was based only on patients alive on January 1 who had reached day 91 of ESRD and followed those patients for the rest of the calendar year. Patients who entered ESRD therapy during the year did not enter the analysis until January 1 of the following year. This meant that one third of the patients in a given dialysis unit were likely to be excluded from the analysis. Since the 1994 ADR, patients who reach day 91 of ESRD during the year are included in the analysis and are followed from that day until the end of the calendar year (or until the censoring date for other time periods).

Previously Transplanted Patients. In the original methodology, previously transplanted dialysis patients were excluded. The new methodology has patients reenter the analysis on the first of the year after a transplant failure. If transplant failure occurs after November 1 of the previous year, the patient is also excluded from the hospitalization analysis in the upcoming year. In the future, this may be revised to have the patient reenter the analysis 60 days after the transplant failure.

New categories. Below is a list of the categories for which mortality rates are reported in this ADR, along with an explanation of the changes in the categories:

- All dialysis patients. *No longer restricted to those who have not received a transplant.* The period at risk for death is censored at transplant date if a transplant occurs during the year. The period at risk for hospitalization is censored three days prior to transplantation so that the hospital stay during which the transplant occurs is not counted.
- Dialysis patients who are on hemodialysis (including home hemodialysis) at the start of the period and who have been on this modality for at least 60 days. *No longer restricted to those who have not received a transplant.* Censuring procedures are the same as those describe above.
- Dialysis patients not yet transplanted who are on CAPD or CCPD at the start of the period and who have been on this modality for at least 60 days. *No longer restricted to those who have not received a transplant.* Censuring procedures are the same as those describe above.
- All ESRD patients.

The “All ESRD” category will be less useful than the others but is included for completeness and as an overall reference. The following categories have been replaced by the separate methodology for patients with a functioning graft, which is described later.

- All patients with a functioning transplant.
- Patients with a functioning cadaveric first transplant.

The cohort of patients used in this method is defined as those who are alive and not lost to followup at the beginning of a year, and who have had ESRD for at least 90 days, plus those patients who reach day 91 during the year. A patient who dies on December 31 of the prior year is excluded, but a patient who dies on January 1 is included with one day at risk. In effect, we are assuming that all deaths occur at 11:59 PM on the death date. Eliminating lost to followup patients is a change introduced with the 1992 Annual Data Report.

Annual Death and Hospitalization Rates for Dialysis and All ESRD

All patients who have reached day 91 of ESRD on January 1 or who reach day 91 later in the year are included and are followed until the end of the calendar year. The starting date is either January 1 or the day the patient reaches day 91 of ESRD.

Patients who have a functioning transplant on the starting day are included only in the "All ESRD" category. Patients who are on dialysis on the starting date are also included in the "All Dialysis" category. A patient in the "All Dialysis" category may also be reported in one of two subgroups (Hemodialysis or CAPD/CCPD) based on the dialysis modality on the starting date and whether the patient has been on that modality for at least 60 days. Dialysis patients who are not on hemodialysis or CAPD/CCPD or who have been on that modality for less than 60 days are included only in the "All ESRD" and "All Dialysis" categories. Note that a given patient may be included in one ("All ESRD" only), two ("All ESRD" and "All Dialysis") or three ("All ESRD", "All Dialysis", and hemo or CAPD) categories.

For patients alive on January 1 to be classified as a hemodialysis or CAPD/CCPD patient, the patient must have been on the indicated treatment modality for 60 days before the start of the year. For incident patients, the 60 day test is applied on day 91. For patients returning from a transplant, the 60 day test is applied when they reenter the analysis on the next January 1.

In determining whether the patient has been on the same modality for at least 60 days, a short (less than 60 days) spell on a different modality is not considered a change in modality if it is preceded and followed by periods of at least 60 days on the primary modality. See Chapter XIII of the 1992 ADR or the *Researcher's Guide to the USRDS Database* for a description of the application of this "60 day rule."

The denominator for both rates is patient years at risk. For the "All Dialysis" category, the period at risk is from the starting date through the death date or December 31, whichever is earliest. For patients on dialysis at the beginning of the year, the period at risk is from the starting date through the death date or December 31 (whichever is earliest). If a patient is transplanted, the time at risk for death ends on the date of the transplant; the time at risk for hospitalization ends three days prior to

transplantation. Note that a dialysis patient who is transplanted during the year will have different periods of risks when he is used in the "All ESRD" and "All Dialysis" categories.

Note that the period at risk is not censored at the start of a lost to followup period during the year. Note also that if a patient is on dialysis and receives a transplant on January 1, then the days at risk are censored at the transplant date, and the patient has 0 days at risk.

Each patient contributes his/her years at risk to the denominator. The numerator is the number of events (either deaths or hospitalizations) which occurred during the at risk periods. Days at risk are divided by 365.25 to yield years at risk. Due to the low numbers in the numerator for death rates, this ratio is multiplied by 1,000 to yield deaths per 1,000 years at risk. Hospitalization rates are in terms of one year at risk.

Years may be pooled in this method to increase the cell sizes, as is done in Sections D and H of the Reference Tables. The analysis is still done in one year increments. A patient who is alive at the beginning of more than one of the pooled years will be used more than once in this method. A patient who dies during the second of the three years would contribute a full year at risk for the first year and a partial year at risk and one death for the second year. For each year, the patient's age and treatment modality will be determined as of the beginning of that year. Thus, a patient could be classified as a 45 year old dialysis patient one year and as a 46 year old functioning transplant patient the next.

Annual Death Rates for Transplant Patients

Beginning with the 1994 this ADR, annual death rates for the first year post transplant and for the next two years post transplant are reported for patients transplanted in the previous three years, i.e. 1988-1990. The last transplant year is 1990 so that it is possible for all of the patients to experience three years post transplant, if they survive. The denominator for both rates is still the patient years at risk. For the first year rate, patient years are measured in days from the date of transplant to the death date or the end of the year for the first year at risk, i.e., 365 days exactly. For the rate for the second two years, the time at risk is measured from the anniversary date, i.e., exactly one year (365 days) following the transplant date, to the death date or the end of the period, i.e., exactly two years at risk.

Mortality rates for transplant patients with functioning grafts older than three years are not reported. In earlier ADRs, such patients rates were mixed with survival rates for more recent transplants; this practice somewhat distorted the mortality rates.

Patient Survival Analysis Methodology

All new Medicare eligible ESRD patients who passed the USRDS quality control filters and who had a first service (dialysis or transplant) between January 1, 1977, and December 21, 1993, were included in the survival analysis. They were followed until December 21, 1994, giving a maximum followup of 15 years and a minimum followup of one year.

In each table in Section E, the survival rates for the most recent year are preliminary and must be used with caution. The analysis periods for the most recent year in each table extends into 1992. The period for which data are considered complete for the current database update ends in February 1992. Thus, the death data for the most recent year in each table may be incomplete.

Patient survival probabilities and standard errors are estimated using the Kaplan-Meier method (KAPLAN) as implemented in the SAS LIFETEST (SAS) procedure and Greenwood's formula. The probabilities are expressed as percentages varying from 0 to 100 (rather than probabilities varying from 0 to 1).

Survival probabilities are computed for annual cohorts for the groups shown below. For the four groups which represent subsequent rather than initial modalities, the median times from ESRD to these modalities are also presented.

In the tables for all patients and for dialysis patients, the cohorts are defined as patients starting ESRD therapy in a calendar year and surviving at least 90 days after starting ESRD therapy. Thus, the one year survival probability is the probability of surviving from day 91 to one year plus 90 days (days 91 to 455, or 3 to 15 months). The two year survival probability is the probability of surviving from day 91 to two years plus 90 days. This delay is necessary because many patients under age 65 do not become eligible for Medicare for up to 90 days, and the database does not have data until the patient becomes eligible. Additional tables for patients 65 and over are presented, showing survival from date of ESRD to day 90 and from date of ESRD to day 365. The latter

results can be compared with the earlier tables for survival from day 91 to day 455.

Given the likely selection bias involved in categorizing patients into modalities, patient characteristics may vary substantially between modality groups. Therefore, differences in survival across modality groups may not necessarily be ascribed to differences in the efficacy of the treatment modalities, even after adjustments for age, race, gender, and primary disease.

One, two, five, and ten-year Kaplan-Meier product limit estimates of survival probabilities were calculated for each age, race, gender, and primary disease group by the year of incidence.

Some survival results in Section E of the Reference Tables are adjusted for age, race, gender, and primary disease. Starting with this ADR, the standard population for these adjustments is the incident ESRD population for the year before the most recently reported year, 1991 for this ADR.

The procedure used is the direct method of adjustment described earlier in this chapter and used for incidence and prevalence. The number of cells was limited to four age, two gender, three race, and four diagnosis groups. These limitations were imposed because more cells would have led to undesirable smaller sample sizes and the Kaplan-Meier survival estimates are biased high with small cell sizes.

All Patients

The cohort consists of all ESRD patients starting renal replacement therapy in a calendar year and surviving beyond day 90. Patients are censored only at the end of followup (December 21, 1994).

Dialysis Patients (Censored at First Transplant)

The cohort consists of all ESRD patients starting renal replacement therapy in a calendar year, surviving beyond day 90, and not receiving a transplant by day 91. Patients are censored at transplantation or at end of followup (December 21, 1994).

First Renal Transplant (Cadaveric)

Patients in this cohort satisfy two conditions: 1) they received their first transplant in the designated year, and 2) the transplant donor is cadaveric.

Transplants where the donor type is recorded as “other” or “unknown” are excluded from both the cadaveric and living related groups. The cohort is defined based on the year of first transplant regardless of the year of first ESRD service. Patients are followed from date of transplantation, and age is computed at the time of transplantation. Patients are censored only at end of followup (December 21, 1994). These patients appear in the “All ESRD” group above, and most also appear in the dialysis modality group. The median time in days from first ESRD service to first transplantation is also presented for these patients.

First Renal Transplant (Living Related)

Patients in this cohort satisfy two conditions: 1) they received their first transplant in the designated year, and 2) the transplant donor is living related. Transplants where the donor type is recorded as “other” or “unknown” are excluded from both the cadaveric and living related groups. The cohort is defined based on the year of first transplant regardless of the year of first ESRD service. Patients are followed from date of transplantation, and age is computed at the time of transplantation. Patients are censored only at end of followup (December 21, 1994). These patients appear in the “All ESRD” group above, and most also appear in the dialysis modality group. The median time in days from first ESRD service to first transplantation is presented for these patients.

Expected Remaining Lifetime Methodology

The expected remaining lifetime for a patient group is the *average* life expectancy for that group. Some individual patients within the cohort will live longer than, and some less than, the average. Although the average will not be known until all of the patients in the cohort have died, the expected remaining lifetime can be *projected* by assuming that the cohort will experience the same death rates that have been observed among groups of recently prevalent ESRD patients.

For a subgroup of ESRD patients of a particular age, the expected remaining lifetime is the result of a calculation based on a survival function, which is in turn the result of a calculation based on observed death rates. The calculations start with USRDS ADR tables, which include observed death rates among ESRD patients for successive age ranges. Let $r(A)$

denote the death rate for a 5 year age group from those tables, where A identifies one of the listed age ranges. These death rates for successive age intervals, $r(A)$, are plotted versus age, A , and the area under the curve up through age A is denoted by $R(A)$. The survival function, $S(A)$, at age A is the fraction of patients that would survive to age A , for a hypothetical patient cohort that is subjected to those death rates throughout their lifetimes. The survival function at age A is related to the death rates by the equation $S(A)=\exp(-R(A))$, where “exp” denotes the exponential function. Among patients alive at age denoted by A , the fraction who survive for X more years is then $S_A(X) = S(A+X)/S(A)$. For a given starting age, A , the expected remaining lifetime is then equal to the area under the curve of $S_A(X)$ plotted versus X . This area is truncated at the upper age limit $A+X=100$, since few patients live beyond age 100.

Graft Survival Analysis Methodology

The methodology for graft survival analysis is the same Kaplan-Meier methodology used for patient survival. For persons who have received multiple transplants, only the first transplant is analyzed. For the patient survival analysis, once a patient receives a transplant, that patient is thereafter classified as a transplant patient even if the graft fails. Graft survival analyses are presented separately for cadaveric and living related first transplants. Transplants where the donor type is recorded as “other” or “unknown” are excluded.

The starting date for the graft survival analysis is the date of the first transplant. The ending date is the earlier of death or graft failure. The derivation of the graft failure date is discussed in the *Researcher's Guide to the USRDS Database*.

Standard errors (after 90 days, 1, 2, 3, 5, and 10 years) are presented for both patient survival and graft survival analyses. The calculation of the standard error of the difference between two graft survival curves is the square root of the sum of the squared individual standard errors, since the covariance is zero. The p values reported in Chapter VII for the differences between survival curves are based on a t -test of the difference in survival.

Some adjusted survival probabilities are computed using the same methods as for patient survival. For this ADR, the standard population for

these adjustments is patients who received a transplant in 1991-1992.

Treatment Modality

The treatment modality categories and the procedures for determining treatment modality are described in the *Researcher's Guide to the USRDS Database*. In Section C of the Reference Tables, the "60 day rule" is applied. This rule requires that the patient be on a new modality for at least 60 days before it is considered a change in modality. Patients who are lost to followup are not included in the year end modality counts.

Section C of the Reference Tables also includes tabulations of patients by modality at two years. Lost to followup is included as a category in these tables.

Cost Effectiveness Methodology

Section K of the Reference Tables and the Cost Effectiveness Chapter are based on the analysis of Medicare Claims data. The claims data were merged first with patient demographics and second with treatment modality information obtained from the USRDS data base.

Medicare Claims Data

There are two types of Medicare claims: institutional and physician/supplier. All of the physician/supplier claims are Medicare Part B. The institutional claims consist of all Part A claims (Inpatient, Outpatient, Skilled Nursing Facility, Home Health Agency, and Hospice) and some Part B claims, notably outpatient dialysis. Physician/supplier claims account for about 80 percent of the claims but only 20 percent of the dollars.

The structure and content of the two types of claims are different, and so are the files derived from them. For institutional claims, there are two types of file: the Institutional Claims (Claims) file and the Institutional Claims Detail file. The Claims file indicates the type of claim, the dollar amounts, and the type of dialysis involved (if any), and the dates of service. The Claims Detail file contains details like DRG, diagnoses, and procedures. For many analyses, the Claims Details file would not be needed.

For the physician/supplier claims, there is one type of file with one record for each claim line item. The file includes dollar amounts, dates of service, diagnosis and procedure codes, and type and place of service.

The Medicare Claims data are obtained from HCFA's Standard Analysis Files (HCFA SAFs). The HCFA SAFs are created on a calendar year basis six months following the end of the calendar year. HCFA estimates that the HCFA SAFs are 98 percent complete. The USRDS obtained data for about 24 million institutional claims (hospital inpatient and outpatient facilities, outpatient dialysis facilities, skilled nursing facilities, hospice facilities, and home health agencies) for 1989 through 1994. In addition, data for 143 million physician/supplier claims line items for 1991 through 1994 were obtained.

Data were obtained for all patient ID numbers currently in the USRDS database plus all ID numbers which had outpatient dialysis claims or a kidney transplant during 1989-1994. The HCFA enrollment database was used to gather all of the cross-reference ID numbers for these IDs, that is all of the ID numbers under which these patients may have had claims. Patients who are not currently in the PMMIS were not included in the analyses in this ADR.

Intent to Treat Model

Table K.1 in the Reference Tables shows aggregate Medicare payments by year and type of payment. The remaining tables are based on an "intent to treat" model. This model computes payments per patient year at risk over the three year period 1991 through 1993. It is referred to as an "intent to treat" model because patients are classified based on their treatment modality at the start of the analysis period and retain that classification even if their dialysis modality changes. Aggregation of Medicare payments was done on an "intent to treat" basis, attributing all subsequent payments to the patient's starting modality. The only exception to this rule was dialysis patients who later received a transplant. These patients were censored at date of transplant and a new record was created with transplant as the intention-to-treat modality. Only patients switching from dialysis to transplant could enter the model twice. All others remained assigned to their initial modality.

Patients are classified into four intent to treat modality categories: hemodialysis, CAPD/CCPD, uncertain dialysis, and transplant. The uncertain dialysis category includes cases where the dialysis modality is unknown or is not hemodialysis or CAPD/CCPD or where the patient has not been on the modality for 60 days at the start of the period.

The transplant category includes patients who have a functioning transplant at the start of the period

and patients who receive a transplant during the period. For patients who receive a transplant during the period, the period is split into two periods. The patient is classified into one of the dialysis periods until the date of the transplant and then begins a new period in the transplant category.

Table K.3 shows the distribution of patients into these groups and the aggregate payments of each type for these groups. The remaining tables in Section K drop "uncertain dialysis" as a separate category and add two aggregate categories: All ESRD and All Dialysis. The "uncertain dialysis" patients are included in the All ESRD and All Dialysis categories.

The intent to treat model spans the three years 1991 through 1993. ESRD patients prevalent on January 1, 1991 or incident at any time during 1991, 1992, or 1993 were potentially eligible for inclusion in the study sample. The study start date for a given patient was defined as thirty days after the latest of the following:

- January 1, 1991.
- The first ESRD service date in the USRDS database for the patient.
- The Medicare Part B entitlement date from the USRDS database.
- The first month in which dialysis payments exceed \$675 (the tenth percentile of monthly dialysis reimbursement for all patients with dialysis in 1991).

Patients for whom Medicare is the secondary payer (MSP status) for all or part of the study period are excluded from the analysis. For patients who have employer paid group health insurance coverage, Medicare is the secondary payer for the first 18 months of ESRD. Data on the MSP status of patients is obtained from the Medicare enrollment database. MSP patients are excluded from the analysis because it is impossible to characterize the total costs of their care. Table K.2 shows that about 6 percent of the patients with Medicare payments during 1991 through 1993 were excluded because they were MSP for all or part of the period.

Medicare payments were aggregated from the study start date until the patient's date of death, date of transplantation, date lost to followup, or December 31, 1993, whichever came first. Patients were defined as lost to followup if there were a period of three consecutive months in which dialysis payments (institutional plus physician/supplier) did not sum to \$675 per month.

Total costs during the followup period were divided by the length of the followup period in order to express all costs as dollars per year at risk (YAR). Costs per year at risk were calculated for all ESRD patients, for all patients by modality, and stratified by modality, age, gender, race, and diabetic vs. nondiabetic cause of ESRD.

Diabetic ESRD status is based on the primary disease causing ESRD. A patient with a non-diabetes cause in some cases may also have diabetes, but the diabetes is not judged to be the cause of ESRD. Persons with cause of ESRD missing are included only in the All category for diabetic ESRD status.

References

- Fleiss JL. *Statistical Methods for Rates and Proportions*, second edition. New York: John Wiley & Sons, 1981.
- Gross AJ, Clark VA. *Survival Distributions: Reliability Applications in the Biomedical Sciences*. New York: John Wiley & Sons. 1975.
- Kaplan EL, Meier P. Nonparametric Estimation from Incomplete Observations. *J Am Statist Assoc* 1958; 53:457-481.
- SAS Institute Inc. *SAS User's Guide: Statistics*, version 5 edition. Cary, NC: SAS Institute Inc. 1985.
- U.S. Bureau of the Census, *Estimates of the Population of States by Age, Sex, Race, and Hispanic Origin: 1981 to 1989* (machine readable data file). U.S. Bureau of the Census, Washington, DC.
- U.S. Bureau of the Census, *Revised Estimates of the Population of Counties by Age, Sex, and Race, 1980 to 1989* (machine readable data file). U.S. Bureau of the Census, Washington, DC.
- U.S. Bureau of the Census, *U.S. Population Estimates by Age, Sex, Race, and Hispanic Origin: 1990 to 1993* (machine readable data file). U.S. Bureau of the Census, Washington, DC. 1994.
- U.S. Bureau of the Census, *Population Projections for States, by Age, Sex, Race, and Hispanic Origin: 1993 to 2020 -- Detailed Results for All States: Series A* (machine readable data file) U.S. Bureau of the Census, Washington, DC, 1994.
- U.S. Bureau of the Census, *Estimates of the Population of Counties by Age, Sex, and Race:*

1991 (machine readable data file). U.S. Bureau of the Census, Washington, DC, 1992.

U.S. Bureau of the Census, 1990 Census of Population and Housing: Modified Age/Race, Sex, and Hispanic Origin (MARS) State and County File (machine readable data file), Data Users Services Division, The Bureau of the Census, Washington, DC, 1992.

Wolfe RA, Gaylin DS, Port FK, Held PJ, Wood CL. Using USRDS generated mortality tables to compare local ESRD mortality rates to national rates. *Kidney Int* 1992; 42:991-996.