USRDS
United States Renal Data System

2025 Researcher's Guide to the USRDS Database

Preface

   The United States Renal Data System (USRDS) provides this Researcher's Guide for investigators within and outside of the USRDS Coordinating Center (CC) who wish to undertake research projects using data from the USRDS database. This guide places particular emphasis on the USRDS Standard Analysis File (SAF) dataset, the primary means by which we make USRDS data available for use. The Researcher's Guide includes information to help investigators select and use the appropriate SAFs for their intended project.

Acknowledgments

   Most of the original data contained in the USRDS End-Stage Renal Disease (ESRD) database are provided by the Centers for Medicare & Medicaid Services (CMS) through an inter-agency agreement between the National Institute of Diabetes and Digestive and Kidney Diseases (NIDDK)/Division of Kidney, Urologic, and Hematologic Diseases (DKUHD) of the National Institutes of Health (NIH), and the CMS Bureau of Data Management and Strategy. The USRDS offers its gratitude and appreciation to everyone who contributed to the USRDS database and to this guide, especially all the ESRD professionals and Networks who report the data included in the ESRD database.

USRDS Coordinating Center

- Director: Kirsten L. Johansen, MD

- Deputy Director: James B. Wetmore, MD, MS

- Project Manager: David T. Gilbertson, PhD

USRDS Coordinating Center Contact Information

USRDS Coordinating Center
701 Park Avenue
Suite S4.100
Minneapolis, MN 55415
usrds@niddk.nih.gov

Publications/Suggested Citation

Publications using the USRDS data should include a citation and a disclaimer.

The suggested citation for this guide:

U.S. Renal Data System. 2025 Researcher's Guide to the USRDS Database. National Institutes of Health, National Institute of Diabetes and Digestive and Kidney Diseases, Bethesda, MD, 2025.

The suggested citation for the 2025 ADR:

U.S. Renal Data System. 2025 USRDS Annual Data Report: Epidemiology of kidney disease in the United States. National Institutes of Health, National Institute of Diabetes and Digestive and Kidney Diseases, Bethesda, MD, 2025.

Publications based upon USRDS data must include the following disclaimer:

"The data reported here have been supplied by the United States Renal Data System (USRDS). The interpretation and reporting of these data are the responsibility of the author(s) and in no way should be seen as an official policy or interpretation of the U.S. government."

Publications based upon USRDS data must also be submitted for privacy assurance to the USRDS Coordinating Center, usrds@niddk.nih.gov, with the subject line: "USRDS Manuscript". Requests should include the journal name to which the manuscript will be submitted and reference the approved Data Use Agreement number.

Introduction

USRDS ADR WEBSITE: https://usrds-adr.niddk.nih.gov/2025

The USRDS ADR website provides users with access to the Annual Data Report (ADR). ADR chapters are presented in HTML and available as downloadable portable document format (.pdf) files. Reference Tables and the data underlying the figures and tables for each chapter are available as Microsoft Excel (.xlsx) files.

The United States Renal Data System provides a range of data products and services, including Standard Analysis Files (SAFs) to researchers. To assist newer users, Figure 1 and Table 1 present a broad view of the available data.
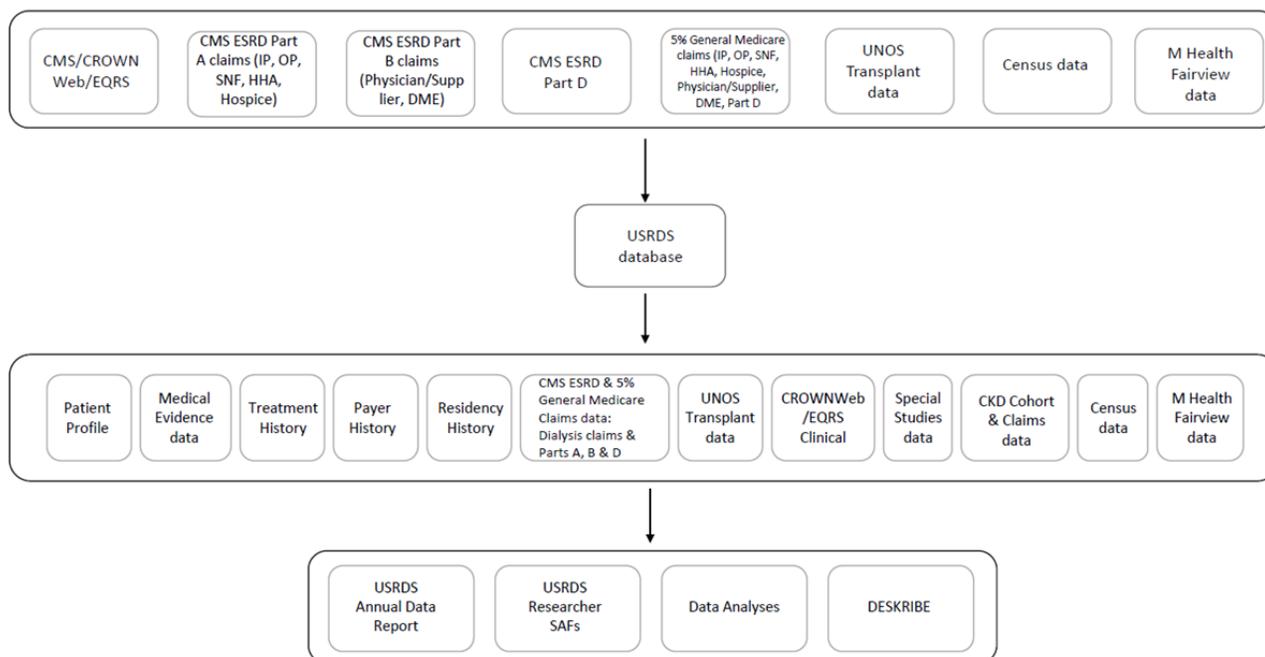
Figure 1. Structure of the USRDS database

Table 1. Year-to-year comparison of number of records in USRDS Standard Analysis Files (SAFs) 2021-2025

	2021	2022	2023	2024	2025
PATIENTS	3,385,510	3,511,615	3,649,796	3,908,455	4,040,067
RESIDENC	5,308,009	5,013,414	5,210,669	5,450,271	5,637,875
MEDEVID†	3,518,351	3,607,103	3,777,224	3,941,689	4,064,904
RXHIST	9,543,870	9,912,769	10,715,084	10,878,343	11,350,060
FACILITY	160,248	168,401	176,928	184,144	192,554
TX	605,341	631,861	644,652	681,913	707,296
TXWAIT*	935,945	981,749	993,261	1,045,086	1,088,286

*TXWAIT is the total number of observations in Waitlist_ki and Waitlist_kp.

Requesting Data

The USRDS presents the results of its research at national conferences and in peer-reviewed journals. A list and links to published journal articles can be found on the website (https://www.niddk.nih.gov/about-niddk/strategic-plans-reports/usrds/about-usrds/publications).

Making information on ESRD available to the renal community is a primary objective of the USRDS, and we are committed to the timely fulfillment of data requests. In many cases, requests can be answered through data published in the ADR or elsewhere. Requests for data not available in material published by the USRDS and that require two hours or less of staff time are often fulfilled by the Coordinating Center within one week (simple data requests, https://www.niddk.nih.gov/about-niddk/strategic-plans-reports/usrds/for-researchers/simple-data-requests). However, more complex requests require the researcher to receive the USRDS data and perform the analyses themselves.

Data Extraction System for Kidney Related Information & Basic Epidemiology (DESKRIBE)

The application system entitled "Data Extraction System for Kidney Related Information & Basic Epidemiology", or DESKRIBE (Data Query Tools, https://www.niddk.nih.gov/about-niddk/strategic-plans-reports/usrds/data-query-tools) is an online data querying application that allows users to create customized data tables and regional maps.

DESKRIBE allows users to retrieve information on the entire ESRD population or on precisely specified subgroups of interest. The system is designed to provide a high-level of interactivity and instant feedback, allowing users with varying background and knowledge to navigate the system. DESKRIBE allows users to create downloadable data tables and interactive maps. A tutorial is available on this site to help new users.

End-Stage Renal Disease Quarterly Update

The end-stage renal disease (ESRD) quarterly updates contain data on incidence and prevalence by calendar quarter over the past 2-3 years, including the most current data that the USRDS has received from CMS. It is updated quarterly. Please note that the data in the ESRD Quarterly Update is more recent than the USRDS ADR, but these data are preliminary counts and are updated quarter by quarter.

DATA REQUESTS

A primary objective of the USRDS is to make data available to the renal community. Types of data requests include:

Simple Data requests: Two-hour. Brief, detailed questions and data requests that are not answered directly by the ADR can be submitted to the Coordinating Center via email (USRDS@niddk.nih.gov); those that require less than two hours of staff time to fulfill will be completed.

Data requests: More than two hours. These types of requests will require researchers to receive the USRDS data and perform analyses to answer their questions. These must be submitted by email in either Word or PDF, and require a proposal, IRB approval and data use agreement (DUA). After processing, your request will be submitted for review and approval by the Authorizing NIDDK Official.

Standard Analysis Files: SAFs provide patient-specific and facility-specific USRDS data to support ESRD research.

Merged data files: The USRDS provides the service of linking population cohorts to the USRDS ESRD dataset to determine ESRD status and outcomes for epidemiological research. Merged files can be created by the Coordinating Center for a limited number of approved research projects per year.

The USRDS SAFs contain a limited dataset and not considered publicly available. Since the SAFs and linking files contain confidential, patient-specific data, their release requires the approval process. To request and use USRDS data files, investigators must provide a detailed description of the proposed investigation. The summary must include an in-depth description of significance of the research, study design and methodology, the requested USRDS datasets by name and why the datasets are necessary to conduct the research, and information on data security, among other information.

The project must comply with HIPAA privacy regulations, and the summary should provide enough information to enable assessment of compliance. Guidelines for adherence to HIPAA privacy regulations are found in the data use agreements found on the USRDS website.

IRB approval/documentation is required, with both a waiver of consent and a waiver of HIPAA authorization being approved. For more information on IRB documentation, please see the Outline for research located at the links in the below paragraph.

Detailed information on how to complete the forms to request data is available on the DUA form and outline form found at https://www.niddk.nih.gov/about-niddk/strategic-plans-reports/usrds/for-researchers/standard-analysis-files for standard requests and https://www.niddk.nih.gov/about-niddk/strategic-plans-reports/usrds/for-researchers/merged-data-requests for merge requests. Requests should be submitted to the USRDS Coordinating Center at USRDS@niddk.nih.gov

The investigator and the Coordinating Center will resolve any technical questions. The NIDDK will review the project for technical merit and conformity with HIPAA privacy regulations. The Authorizing NIDDK Official or USRDS Coordinating Center will notify the investigator(s) if the project is not approved, and discuss reasons for the decision. If the project is approved, the Authorizing NIDDK Official will send a copy of the approval letter to the Coordinating Center. The Coordinating Center will then prepare the files and documentation and send them to the investigator.

Any reports or articles resulting from the use of USRDS data must be submitted to the Coordinating Center for NIDDK review before submission for publication to assure adherence to HIPAA privacy regulations. The NIDDK reviewer must respond within 30 days. If a report or article is determined not to adhere to the Privacy Act, it shall not be published until compliance is achieved. Assessment of compliance will not depend on the opinions and conclusions expressed by the investigators, nor will the NIDDK reviewer's approval indicate government endorsement of the investigator's opinions and conclusions.

All publications using released data must contain the standard acknowledgment and disclaimer presented above (Page 2). Investigators are requested to send copies or links of all published manuscripts resulting from this research to the Coordinating Center.

CAVEATS

This policy establishes conditions and procedures for the release of data from the USRDS and is intended to ensure that data are made available to investigators in the pursuit of legitimate research. *Research* means a systematic investigation, including research development, testing, and evaluation, designed to develop or contribute to generalizable knowledge (45 CFR 46) directed toward a fuller scientific knowledge or understanding of the subject studied.

The USRDS will not release data that identify individual patients, however, the USRDS data is a limited dataset containing some identifiers. Since it may be possible to infer identity from SAF data, these data are considered confidential. The USRDS data use agreements (Agreement for Release of Data and Merged Dataset Agreement for

Release of Data) contain general and specific restrictions on the use of USRDS data, and investigators are expected to abide by these restrictions. If individually identifiable data are needed, the request should be submitted directly to CMS. Use of these data to identify and/or contact patients, facilities, or providers is prohibited by USRDS policy and HIPAA privacy regulations.

The USRDS cannot release data to researchers located outside of the US. For more information, see the ResDAC website https://resdac.org/articles/cms-non-us-based-researcher-policy

The USRDS Coordinating Center will provide data files only. Analytical services other than review of the proposal and preparation of the data files are not provided under the USRDS contract, although Coordinating Center personnel may participate in analyses funded by other sources.

FILE FORMATS

SAFs are provided as SAS files only. The SAS format is widely used, easily transported, and largely self- documenting. SAS is a commercially available data management and statistical analysis software system that runs on most computers and is almost universally available on university computer systems. The SAFs take full advantage of the program's ability to incorporate detailed documentation into the file. Researchers needing another format must arrange for the conversion.

RESEARCH FILES

The Coordinating Center maintains a set of SAFs that can meet the majority of research needs and provide easy access to the data used in the ADR. The SAFs were introduced in 1994, as the NIDDK began awarding new grants focusing on research using the USRDS data. The result has been an annual increase in the number of files provided by the USRDS.

Before 1994, all researcher files were created for specific projects. Since the introduction of the SAFs, however, custom files, beyond the linking crosswalk using the researcher-provided patient finder file, are not created. For more information on merged data requests, please see the USRDS website.

The Core set of SAFs contains basic patient data and is required to use all the other SAFs. Included are each patient's demographic information, payer and treatment history, limited transplant data, and provider data. Approximately half of the researchers using the USRDS data need only these files.

The Transplant dataset contains detailed transplant and transplant follow-up data collected by the Organ Procurement and Transplantation Network (OPTN).

Data on hospital inpatient stays are found in the Hospital dataset.

All Medicare billing data are available by individual year.

SAF use is governed by the USRDS policy on data release for investigator-initiated research. Research proposals must be approved by a NIDDK Official, and institutions and researchers must sign the applicable USRDS data use agreement available at the USRDS website.

Most SAFs provide patient-specific data and are considered limited datasets. Many patient identifiers are removed or encrypted, but data confidentiality remains a serious concern. The USRDS data use agreements describe restrictions on SAF use and disposition. The agreement(s) must be signed by the appropriate institutional authority responsible for Information Technology and privacy security, then co-signed by the investigators/analysts to acknowledge their responsibility for protecting

the privacy of this kind of individual patient data. SAFs include an encrypted ID number to allow patient data from multiple SAFs to be merged.

FILES AVAILABLE FOR REQUEST

A complete listing of the variables included in these files can be found in the *2025 Researcher's Guide to the USRDS Database Appendix B: USRDS Data File Descriptions*.

Files vary by size, please see Section 16 of this document for sizes.

CORE DATASET

The Core Standard Analysis Files contain the most frequently used data and **are required for the use** of the other datasets. Files included in the Core Dataset, with brief explanations, are shown below in Table 2. Please note that the Medicare Payment/Claims data lags behind the vital status information (e.g. death, graft loss) contained within the Core PATIENTS and Core TRANSPLANT files by up to a year.

The Core includes some transplant-related files
- TX includes minimum details on all transplants from all sources.
- WAITLIST_KI and WAITLIST_KP contain one record for each patient in the USRDS database per waiting-list event for kidney and kidney-pancreas transplants. These data were collected by OPTN.
- WAITSEQ_KI and WAITSEQ_KP contain condensed, center-specific, kidney and kidney- pancreas waiting list dates. These data were collected by OPTN.

TRANSPLANT DATASET

The Transplant dataset contains detailed transplant data from OPTN and includes these files:
- TXUNOS_TRR_KI, TXUNOS_TRR_KP, TXUNOS_TCR_KI, TXUNOS_TCR_KP, TXUNOS_CDR, TXUNOS_LDR, TXUNOS_DHS, TXUNOS_RHS) include transplant information collected since 1987 by OPTN, who is currently the primary source of transplant data for the USRDS.
- TXIRUNOS includes information on immunosuppressive drugs collected by OPTN at the time of transplantation events.
- TXFUUNOS_KI and TXFUUNOS_KP include transplant follow-up reports collected by OPTN since 1988 for kidney and kidney-pancreas recipients.
- TXIFUNOS includes information on immunosuppressive drugs, collected by OPTN at follow-up visits.

HOSPITALIZATION DATASET

Hospitalization data are derived from the inpatient Claims. The Hospital SAF contains diagnosis and surgical procedure codes for each stay. No payment or cost variables are included in this dataset, which is intended for researchers who need data on hospital inpatient stays and diagnoses and procedures for those stays, but who do not require payment data.

CROWNWEB/EQRS CLINICAL DATASET

EQRS (or historically CrownWeb) data, are monthly clinical data collected from the CMS End-Stage Renal Disease Quality Reporting System. This includes clinical data for hemodialysis adequacy, peritoneal dialysis adequacy, anemia management, iron, mineral metabolism, and vascular access.

CKD 5% COHORT CORE & COHORT HOSPITALIZATION DATASETS

The CKD cohort datasets are built from the 5% general Medicare Claims SAFs and contain a patient master file, a

payer sequence file, and a set of comorbidity files. The data includes all those who had at least one claim with a diagnosis code of CKD.

CLAIM DATA

Claim data are available as institutional and physician/supplier claim datasets for the ESRD population enrolled in the **Medicare fee-for-service** program. Available years can be found on the DUA form. Please note that the Medicare payment and claims data lags behind the vital status information (e.g. death, graft loss) contained in the Core PATIENTS and Core TRANSPLANT files by up to a year.

Institutional claims consist of all inpatient/outpatient claims (inpatient, outpatient, skilled nursing facility, home health agency, and hospice), including outpatient dialysis claims. Physician/Supplier claims are bills covering physician, laboratory, and radiology services, as well as medical equipment/supplies.The structure and content of the two types of claims differ, as do the files derived from them.

Institutional claims are provided in three types of files. The Institutional Claims file (CLM in the file name) indicates claim type, dollar amounts, DRG code, type of dialysis involved (if any), and dates of service. The Institutional Claims Details file (DXP in the file name) contains details such as diagnosis and procedure codes. The Revenue Center Details file (DET in the file name) contains revenue center details. Many analyses require only the Institutional Claims files.

Physician/supplier claims are provided in three types of files — claim-level file (CLM in the file name), multiple trailers for line item details of the claim (LINE in the file name), and multiple trailers for diagnoses and procedure details of the claim (DX in the file name).

The ESRD Standard Analytical Files (SAFs) contain data from final action claims submitted for Medicare beneficiaries, enrolled in the Medicare fee-for-service program, in which all adjustments have been resolved.

For Institutional Claims, the USRDS uses the following 100% SAF claims:
- Inpatient
- Outpatient
- Skilled Nursing Facility
- Home Health Agency
- Hospice

For Physician/Supplier, the USRDS uses the following 100% SAF claims:
- Physician/Supplier
- Durable Medical Equipment

The USRDS 2025 SAFs include all claims up to **December 31, 2023**.

MEDICARE CLAIMS CLINICAL DATA

Medicare Claims Clinical Data is extracted from medical claims that specifically indicate dialysis (via revenue center details, diagnoses, and procedures codes) and is available for the claim year 2011 and on. It contains dialysis-related information only, including dialysis vascular access type, dialysis session, ESA use, hemoglobin, Kt/V etc. These are extracted from Institutional SAF claims files (Outpatient, Inpatient, Home Health, Skilled Nursing, Hospice). They can be linked back to the original claims in the claim files for claim details using the claim link key (variable SEQ_KEYC) and the USRDS_ID. The Medicare Claims Clinical data does not include diagnosis or procedures codes to assess for comorbid conditions.

PRE-ESRD MEDICARE CLAIMS

The pre-ESRD claims (also known as the back-casted claims) are a collection of Medicare Institutional, and Physician/Supplier billing records and Prescription Drug events incurred before the onset of ESRD. The USRDS has made the pre-ESRD data available as SAFs. The structure of the files is identical to the ESRD claims files and organized by calendar year. In the 2011 and later pre-ESRD claim files, the variable list is identical to the ESRD claims files but is organized by incident year. The data includes claims that occurred before the first ESRD service date of the incident year and the four prior calendar years (for 2011-2017 incident years) and one prior calendar year (for 2018 and later). Also, a pre-ESRD payer sequence is provided for researchers to determine Medicare enrollment for the periods before the first ESRD service date. A listing of available years can be found on the DUA form.

PART D DATA

Section 101 of the Medicare Prescription Drug, Improvement, and Modernization Act of 2003 amended Title XVIII of the Social Security Act by establishing the Voluntary Prescription Drug Benefit Program (Part D). Effective January 1, 2006, Part D is an optional prescription drug benefit for individuals who are entitled to Medicare benefits under Part A or enrolled in Medicare benefits under Part B. The data from the first few months of 2006 when the benefit was very new may be incomplete and should be interpreted in that context.

The Part D data is obtained from CMS annually, with finder files provided by the USRDS. The Part D data are divided into two separate files: an annual enrollment file containing monthly indicators of enrollment in Part D, and a prescription drug event file (PDE) containing details of prescriptions filled by Part D beneficiaries.

As Part D is voluntary, not all Medicare beneficiaries are enrolled. The annual enrollment file contains 12 monthly indicators that detail whether the beneficiary is enrolled in Part D, and if so, the type of plan. There are also monthly indicators for dual eligibility (Medicare and Medicaid), the Retiree Drug Subsidy, and the Low-Income Subsidy (LIS).

CKD 5% GENERAL MEDICARE DATA

Claims for the 5% general Medicare CKD cohort, including Inpatient, Outpatient, Skilled Nursing Facility, Home Health, and Hospice institutional claims and physician, lab, and Durable Medical Equipment claims. The cohort includes those who had at least one claim with a CKD diagnosis code. Data are derived from the Medicare claims SAF files for a 5% random sample. Please see the DUA for years of claims available.

PHYSICIAN AND PROVIDER CROSSWALKS

The USRDS can provide two crosswalks to researchers that allow the researcher to link the USRDS data to some publicly-available datasets. The Provider crosswalk contains a crosswalk of the CMS Provider ID to the USRDS provider ID. The Physician crosswalk contains a crosswalk of the CMS Physician NPI number to the USRDS physician ID.

SPECIAL STUDIES DATASET

Some USRDS Special Studies surveys were conducted in previous years. Information on these can be found in Section 15. These are available by special request only on a separate DUA. Please contact the USRDS Coordinating Center for more details.

M HEALTH FAIRVIEW DATABASE

The M Health Fairview clinical data are integrated electronic health records (EHR) from M Health Fairview.

Information on this database can be found in the Methods & Databases section. These data are not linked to the USRDS database and therefore do not contain a USRDS_ID.

A complete listing of the variables included in the available files is available in *2025 Researcher's Guide to the USRDS Database Appendix B: USRDS Data File Descriptions*.

LINKAGE TO THE UNITED NETWORK FOR ORGAN SHARING DATA (UNOS)

An investigator can link a United Network for Organ Sharing (UNOS) center STAR file to the USRDS data. To do so, the researcher will need to (1) submit a merge request to USRDS and(2) provide their UNOS approval for the center STAR file and linkage. Upon approval the researcher will submit to USRDS the patient identifiers (SSN, DOB, etc). Please note, the National STAR File from UNOS contains encrypted patient IDs, and cannot be linked to another data source.

Table 2. USRDS Core Files

Patients
The Patients (PATIENTS) file contains one record per patient in the USRDS database and provides basic demographic and ESRD-related data. The ESRD-related data includes ESRD incident date, cause of ESRD, transplant date, and patient survival data. This file is required to use any of the other Standard Analysis Files. Most other files will need to be linked to this file using the USRDS-assigned patient ID.
Medical Evidence
The Medical Evidence file (MEDEVID) contains full data from the 1987, 1995, 2005, 2015, 2018, and 2024 versions of the CMS Medical Evidence form (CMS 2728). In April 1995, a new version of the form was introduced, collecting data on comorbidity, employment status, lab values at initiation, and Hispanic ethnicity; an expanded form was later implemented in 2005. The 2015 form was implemented to align with the transition to use of ICD-10-CM codes to report Primary Cause of Renal Failure. One record for each 2728 form filed for 1987, 1995, 2005, 2015, 2018, or 2024 versions. ESRD first service date, initial treatment modality, comorbid conditions, patient status at the start of ESRD. See the "Methods & Databases" section in this Researchers' Guide for details on Medical Evidence form and their versions.
Treatment History
The Treatment History files (RXHIST, RXHIST60) contain a record for each patient at each change in modality (in both files) or dialysis provider (RXHIST only). RXHIST60 also condensed the records less than 60 day to the previous record.
Payer History
The Payer History file (PAYHIST) contains a record for each patient at each change in insurance payer. One record for each period a patient is covered by one payer; each patient can have multiple records. The insurance payers affect the viability of clinical data from claims.
Transplant
The Transplant file (TX) contains basic data for all kidney transplants reported by CMS and OPTN, including the date of transplant, date of graft failure, demographics of kidney recipient and donor. More detailed transplant data are contained in a separate transplant dataset. One record for each transplant event; a patient can have multiple events.
Transplant Waiting List
Beginning with 2001 data (used in the 2002 ADR), the Transplant Waiting List files (WAITLIST_KI, WAITLIST_KP, WAITSEQ_KI, WAITSEQ_KP) have been updated to include basic patient demographic data and, from OPTN, all unique waiting-list periods for each dialysis patient ever listed. One or more records for each patient ever wait-listed.

Facility

Conducted annually, the CMS End-Stage Renal Disease Facility Survey is the source of data for the Facility file (FACILITY). One record for each facility each year. A facility served as both transplant center and dialysis center may have two records, one for dialysis service and one for transplant service. The survey period is January 1 through December 31. It includes information on facility type, number of patients by modality, and number of staff positions in the facility.

Death

The Death file (DEATH) contains full data from the CMS Death Notification form for ESRD patients (CMS 2746). Date, cause, and place of Death are included.

Census

Data include U.S. population from U.S. Census bureau and from the Surveillance, Epidemiology, and End Results (SEER) Program. Data include U.S. level population counts by age, sex, race/ethnicity, and calendar year (CPUS4RE, CPUS5RE, CPUS6RE, CPUS7RE), and state-level population counts by age, sex, race/ethnicity, and calendar year (CPST4RE and CPST5RE). The -4RE (4 race/ethnicity groups) and -5RE (5 race/ethnicity groups) files are created based the SEER population data (bridged races) and the -6RE and -7RE files are related from Census data (not-bridged race data).

Residence

The Residence file (RESIDENC) is a longitudinal record of each patient's residence (zip code level). For each patient, one record for each period in a different residence. Regional analyses.

Formats

All USRDS-defined SAS formats used by SAFs. Format library used to format values of categorical variables.

What's New in the 2025 SAFs?

Based on the data availability and public research demand, some variables were added to SAF files or dropped.

From October 2024, the new ESRD Medical Evidence form (CMS 2728) and ESRD Death Notification form (CMS 2746) replaced the 2018 versions. Some variables were added and the contents of some variables were changed. This *What's New* reflects most of changes, but some minor changes (such as adding a value/category for a variable) are not shown here. Comorbidities were added from a-w to a-zz and the corresponding disease of each letter was also changed for most of the letters. The variable "COMORBID" contains all comorbid conditions a patient had at onset of ESRD using the corresponding letters. To defined comorbid conditions, researchers will need to use the letters in the variable "COMORBID" and the form version. For all versions before 2024, the letters have the same meaning as that in the 2018 version. However, for the 2024 version, researchers will need to map the letters to comorbid conditions based on the 2024 Medical Evidence form. Comparing the 2018 version of the forms with the 2024 forms will be helpful before using information from CMS 2728 and CMS 2746.

1. New variables added

File Death

The following two variables were added to the ESRD Death Notification form (CMS 2746) from October 2024 and added to the 2025 Core DEATH file.

Variable Added	Label
SHORT_TERM_DLYS_PRIOR_DEATH	Did the patient experience a short-term course (acute) of dialysis prior to death?
DLYS_ENDD_AFTR_RQST_IND	Indicates whether patient/family requested end of dialysis

File Facility

The following seven variables were added to the Core FACILITY file. The content of variable "CERTCODE" was changed in EQRS data and was dropped from the Core SAF. We created a variable "CENTER_TYPE" to replace it. A facility can be a dialysis center and a transplant center; they may share the same provider number; but they usually have different ORG_IDs. So, the ORG_ID was added to the file this year. Also, some facilities are hospital based. On the facility survey file, they also provide the provider number of the hospital. We added the hospital provider number (PROVUSRD_HOSP) to the FACILITY file this year. Please pay attention to the fact that in different survey years, a facility might exchange the provider numbers between facility and hospital. To identify the same facilities, you may need to use both variables.

Variable Added	Label
CENTER_TYPE	Center/Facility is 1 dialysis center or 2-transplant center or 3-dialysis and transplant center
ORG_ID	Facility being surveyed in Form 2744
OPEN_DT	Date Provider opened
CLOSE_DT	Date Provider closed
ORG_FAC_PGM_TYPE_ID	Unique identifier of an organization facility program type
ORG_FAC_PGM_TYPE_DESC	Description of the organization facility program type
PROVUSRD_HOSP	USRDS assigned provider ID

File Medevid

Thirty-four variables were added to the Medical Evidence form (MEDEVID) file.

Variable Added	Label
HAS_LIVING_WILL	Does the patient have living will or Medical/Physician order for life sustaining treatment?
HOUSING_CONCERN	Is the patient currently concerned about where patient will live over the next 90 days?
DAILY_CARE	Does patient have caregiver support to assist with daily care?
DAILY_CARE_DLYS_TRNSPLNT	Does patient have caregiver support to assist with home dialysis/kidney transplant?
DAILY_CARE_COHABITS	Does patient have caregiver live with them to assist with daily care?
UNDERSTANDS_MEDICAL_LIT	Does the patient understand health literature in English?
NEEDS_NONWRITTEN_DOCS	Does the patient need a different way other than written documents to learn about your health?
NEEDS_TRANSLATOR	Does the patient need a translator to understand health information?
REFERRED_TRNSPLNT_CNTR	Has the patient been connected to a transplant center with a referral? (yes,no)
REFERRED_TRNSPLNT_CNTR_DT	Date Referred to Transplant Center
CATHETER_USED	Was one lumen of the Central Venous Catheter used and one needle placed in an AVF or graft?
PD_CATHETER_PRESENT	Is PD catheter present?
ACUTE_DIAGNOSIS	Was patient diagnosed with an acute kidney injury in the last 12 months?
ACUTE_DIAGNOSIS_DIALYSIS	Dialysis Required for Diagnosed with an Acute Kidney Injury
UNDERSTOOD_HOME_DIALYSIS	Does the patient understand home dialysis options at the time of admission?
NOTIFIED_HOME_DIALYSIS	Does the patient indicate they received and understood options for a home dialysis modality
NOTIFIED_TRNSPLNT	Does the patient indicate they received and understood options for a kidney transplant?
NOTIFIED_LIVING_TRNSPLNT	For Living donor transplant
NOTIFIED_NO_DLYS_20H	Does the patient indicate they received and understood the option of not starting dialysis at all, also called active medical management without dialysis?
CYSTATIN_C_DATE	Cystatin C date
CYSTATIN_C	Cystatin C value
PATNOTINFORMEDREASON_DLYS	Reason why patient NOT informed of home dialysis options
COMO_DM_DIABOINJ	Diabetes, currently use other injectable
PATTXOP_OVERWHELMING	Patient NOT informed of transplant options: Patient found info overwhelming
PATTXOP_COGNITIVEIMPAIRMENT	Patient NOT informed of transplant options: Cognitive Impairment
PATTXOP_CONTRAINDICATION	Patient NOT informed of transplant options: Patient has an absolute contraindication
PATDLYSOP_DECLINE	Patient NOT informed of home dialysis options: Patient declines information
PATDLYSOP_UNASSESSSED	Patient NOT informed of home dialysis options: Patient has not been assessed
PATDLYSOP_OTHER	Patient NOT informed of home dialysis options: Other
PATDLYSOP_OVERWHELMING	Patient NOT informed of home dialysis options: Patient found info overwhelming
PATDLYSOP_COGNITIVEIMPAIRMEN	Patient NOT informed of home dialysis options: Cognitive Impairment
PATDLYSOP_CONTRAINDICATION	Patient NOT informed of home dialysis options: Patient has an absolute contraindication
PROVUSRD_REF_TXCTR	USRDS Assigned Provider ID Number for referred transplant center
PHYSUSRD_REF_TXCTR_NPI	USRDS Assigned Physician ID referred Transplant Center

Files waitlist_ki and waitlist_kp

The race variable was not included in the OPTN files that we used to create files waitlist_ki and waitlist_kp. We added it to the files based on other files in the OPTN database this year. It contains information about race and ethnicity.

Variable Added	Label
Race	Race and ethnicity

2. Variables dropped

File Facility

Variable	Label	Drop Reason
CERTCODE	Certified Code (Unique identifier of Medicare status type)	Replaced by CENTER_TYPE

3. Format Changes

Formats Dropped

FMTNAME	Drop Reason
FSCERT	This format was applied on the variable "CERTCODE" which was replaced by the variable "CENTER_TYPE"
MESET_CUR	It only matches 1995 version of Medical Evidence form the corresponding variable has been dropped

Format Value Changed

(1) Format ACCESS

In the 2024 SAFs		In the 2025 SAFs	
Code	Label	Code	Label
1	AVF	1	AVF
2	Graft	2	Graft
3	Cathe	3	Catheter (form_vrsn_txt < 2023)
4	Other	4	Other
5	Ukn	5	PD Catheter (form_vrsn_txt = 2023
		6	Central Venous Catheter (form_vrsn_txt = 2023)

(2) EMPSTAT

In the 2024 SAFs		In the 2025 SAFs	
Code	Label	Code	Label
1	Unemployed	1	Unemployed
2	Emp full-time	2	Employed Full Time
3	Emp pt-time	3	Employed Part Time
4	Homemaker	4	Homemaker
5	Ret-age	5	Retired due to Age/Preference
6	Ret-disab	6	Retired (Disability)
7	Med LOA	7	Medical Leave of Absence

8	Student	8	Student
9	Other	9	Volunteer

(3) MESET

In the 2024 SAFs		In the 2025 SAFs	
Code	Label	Code	Label
1	Home	1	Home
2	Dialysis facility/center	2	In-center
3	Skilled nursing facility/Long-term care facility	3	SNF/Long-term Care Facility
4	Missing/Unknown	4	Transitional Care Unit

(4) DISGRPC

In the 2024 SAFs		In the 2025 SAFs	
Code	Label	Code	Label
	OTHER		MISSING CAUSE
1	DIABETES	1	DIABETES
2	HYPERTENSION	2	HYPERTENSION
3	GLOMERULONEPHRITIS	3	GLOMERULONEPHRITIS
4	CYSTIC KIDNEY	4	CYSTIC KIDNEY
5	OTHER UROLOGIC	5	OTHER UROLOGIC
6	OTHER CAUSE	6	OTHER CAUSE
7	UNKNOWN CAUSE	7	UNKNOWN CAUSE
8	MISSING CAUSE	8	MISSING CAUSE

(5) DONORTYP

In the 2024 SAFs		In the 2025 SAFs	
Code	Label	Code	Label
1	Living Related	1	Deceased Donor
2	Living Unrelated	2	Living Related
3	Deceased	3	Living Unrelated
		4	Unknown
		5	Multi-organ
		6	Paired Exchange

Getting Started

Work on a study typically consists of (1) determining the study cohort (patients) and (2) study variables, (3) selecting the variables from the datasets in which they are stored, (4) merging the selected variables into one or more datasets for analysis, and finally, (5) performing statistical analyses of the data. This section offers dataset examples of this process for first-time users of the USRDS SAFs.

Determining the study cohort is particularly critical when performing analysis of Medicare Claims data. Medicare claims data provides a complete picture only for those patients for whom Medicare is the primary insurance payer. For patients with

other healthcare insurers or for whom Medicare is a secondary payer, information is not captured for hospitalizations or other treatments paid for by an alternate insurance. For this reason, the inclusion of Medicare claims for patients with other insurers or with Medicare as a secondary payer will likely lead to inaccurate results.

To perform an analysis restricted to Medicare as primary payer patients, the Payer History File should be used to identify claims falling into periods when Medicare is confirmed as the primary payer. We suggest using the SAS syntax in *Example 8: Create a Patient Cohort of 2022 Medicare Primary Incident Patients* as a guide; the example relies on the Payer History File (see in this document *Section 4: Payer History*).

All USRDS data are stored in SAS datasets that were created in a Windows environment and can be directly used only on a Windows operating system. Using the datasets on another computer platform will require moving and converting the data to SAS datasets for that platform. Using another data analysis system will also require a conversion to a format compatible with that system. However, *R* users can directly read the data in with some R-packages. Two example packages are "SAS7bdat" and "Haven". "SAS7bdat" has a function "read.sas7bdat" and Haven" has a function "read_sas", they can read SAS data files directly into *R. R* handles memory differently than SAS, and it may be important to read in only variables that are needed from some datasets that are very large.

We provide examples using basic SAS code. New SAS users should take classes, consult colleagues, or otherwise become familiar with the SAS system. Regardless of SAS experience, all users must know the following:

- The location of the SAS dataset
- The location of the SAS catalog of FORMATS

In the examples, the datasets and the FORMAT catalog are assumed to be in the same directory, namely, C:\SAF. Two SAS statements are needed to point to this information:

LIBNAME saf "c:\saf"; * Directory location of the data;

LIBNAME LIBRARY "c:\saf"; * Directory location of format catalog;

Always assume that these two statements are required in all code.

(See *2025 Researcher's Guide to the USRDS Database Appendix C: USRDS Data Formatting* for additional information.) In actual analyses, the datasets and the format catalog may be stored in different directories.

BASIC SAS USE

The SAS CONTENTS procedure generates a list of all variables in the dataset and a label associated with each. The information in this label is usually an adequate indication of whether the variable will be used; PROC CONTENTS is always the best method for obtaining the latest variable list, however, as last-minute updates may not be reflected in the database documentation (see *2025 Researcher's Guide to the USRDS Database Appendix B: USRDS Data File Descriptions*).

Note: In the following program examples, SAS commands, keywords, and procedures are in uppercase text. Datasets, comments, and variables in which SAS will perform operations are in lowercase text.

To determine the contents of the PATIENTS dataset or any other dataset, use the following code (remember to include the two LIBNAME requirements, as seen above).

PROC CONTENTS DATA=saf.patients; TITLE1 'DATASET:

saf.patients';

RUN;

To view the data, use the SAS procedure PRINT to list all observations of the dataset:

PROC PRINT data=saf.patients;

TITLE1 'DATASET: saf.patients'; RUN;

A dataset can contain millions of observations. To print the first 500, use the following code:

PROC PRINT DATA=saf.patients (OBS=500); TITLE1 'DATASET:

saf.patients';

RUN;

   To print a group of observations other than the first N observations, use the following code; this example prints observations 1,500-2,000, inclusive:

PROC PRINT DATA=saf.patients (FIRSTOBS=1500 OBS=2000);

   TITLE1 'DATASET: saf.patients'; RUN;

A dataset can contain hundreds of variables. Print selected variables using the VAR statement:

PROC PRINT DATA=saf.patients (OBS=500);

VAR usrds_id sex race incyear;

TITLE1 'DATASET: saf.patients';

RUN;

   There are two parts to the "DATA=" expression. The first is the directory location, defined by the LIBNAME, and the second is the dataset name. The first part of the expression (e.g. saf, the directory location) implies that the dataset is permanent. Our later examples do not include this expression. Thus the datasets created will automatically be deleted when SAS is closed. To permanently save a dataset, the first part of the DATA expression must be included.

SAS FORMATS FOR USRDS DATA

   The information contained in a dataset is typically coded. Thus, the variable SEX may have the values F, M, or U, where "F" is the coded value for female, "M" is the value for male, and "U" is the value for unknown.

   Using one character instead of several saves disk storage, but because the coded values are not always easy for users to understand, a format is assigned to translate them. Many variables have been assigned formats by the USRDS. In each program, SAS must be told where the catalog of these formats are. Assuming that the format catalog is "formats64" and saved in the directory C:\SAF, then the following SAS LIBNAME makes the formats accessible to your SAS programs:

LIBNAME LIBRARY "c:\saf"; * Directory location of format catalog;

OPTIONS FMTSEARCH = (LIBRARY.formats64);

   The SAS keyword LIBRARY must be included because it tells SAS to look for formats in the directory, C:\SAF. Other methods can accomplish this, but not as easily.

To bypass the use of formats, use the following SAS statement before running any SAS procedures:

OPTIONS NOFMTERR;

Bypassing the use of formats allows you to see raw data instead of the formatted values, which may be useful when you need to write SAS IF statements to control the flow of your program in a SAS data step. (See *2025 Researcher's Guide to the USRDS D*atabase *Appendix C: USRDS Data Formatting* for a tabular list of the formats with their coded values.)

If you want to know a specific format's definition for coded values following are SAS examples statements for character format and numeric format respectively:

PROC FORMAT LIBRARY = LIBRARY.formats64 FMTLIB; SELECT $SEXFMT;

RUN;

PROC FORMTA LIBRARY = LIBRARY.formats64 FMTLIB;

 SELECT RACE RUN;

COMMENT LINES

Comment lines in the SAS code are written like this:

/* Comment line */ * Comment line;

These can include any descriptive comment. The use of comments is optional but strongly recommended.

THE SAF DIRECTORY

Throughout this section, SAF is used as the permanent SAS LIBNAME. It is assumed that all the USRDS SAF datasets and catalogs have been placed in this directory. If the datasets have been placed in the WINDOWS directory C:\SAF, then the following SAS LIBNAME would point to the SAF datasets: LIBNAME saf "c:\saf";

Note: The datasets may be loaded into any directory, with the directory in the LIBNAME changed accordingly.

SAS Examples

EXAMPLE 1: INCIDENT COHORT
-- Create a patient cohort of incident patients from 2012 to 2022 --;

data inc2012_2022 (KEEP=usrds_id incyear rxgroup);
        set saf.rxhist;
        by USRDS_ID BEGDATE;
        if first.usrds_id;
        incyear=year(begdate);
        if 2012 <= incyear <= 2022;
run;

PROC FREQ DATA=inc2012_2022;
TABLE incyear;

```
   TITLE 'Frequency Distribution of Incident Patients 2012 - 2022 by Incident Year'; RUN;


PROC FREQ DATA=inc2012_2022;
TABLE incyear * rxgroup;
   TITLE 'Frequency Distribution of Incident Patients 2012 - 2022 by Incident Year and Modality'; RUN;
```

EXAMPLE 2: INCIDENT PATIENT DISTRIBUTION BY DEMOGRAPHIC DATA

```
*-- Get year 2022 incident cohort first --*;

data inc_2022 (KEEP=usrds_id incyear rxgroup);
        set saf.rxhist;
        by USRDS_ID BEGDATE;
        if first.usrds_id;
        incyear=year(begdate);
        if incyear = 2022;
run;

*-- Get the patient demographic information from the Patients file for the incident cohort created above --*;

PROC SQL;
        create table inc_2022_demo as
        select a.*, b.incagec, b.sex, b.race, b.disgrpc, b.first_se, b.died from inc_2022  as a
        left join saf.patients as b
        on a.USRDS_ID = b.USRDS_ID
        order by USRDS_ID;
QUIT;


PROC FREQ DATA= inc_2022_demo; TABLE incagec sex race disgrpc;


TITLE 'Frequency Distribution of 2022 Incident Patients by Demographic Data'; RUN;
```

EXAMPLE 3: POINT PREVALENT DIALYSIS COHORT

```
*-- Create a patient cohort of point prevalent dialysis (not including Uncertain Dialysis) patients who were alive on
December 31st, 2022. --*;

DATA pre_2022 (KEEP=usrds_id rxgroup);
SET saf.rxhist;
BY usrds_id begdate;
IF (begdate <= MDY(12,31,2022)) AND
   (enddate > MDY(12,31,2022) or (enddate = MDY(12,31,2022) AND (death^='1'))) AND
(rxgroup NOT IN ("B", "X", "Z", "T","A")) THEN OUTPUT;
RUN;


PROC FREQ DATA=pre_2022;
```

```
TABLE rxgroup;
  TITLE 'Frequency Distribution of 12-31-2022 Point Prevalent Dialysis Patients by Modality'; RUN;
```

EXAMPLE 4: 2017 INCIDENT DIALYSIS PATIENT SURVIVAL RATES (KAPLAN-MEIER)

```
*-- Calculate survival rates of 2017 incident dialysis patients using Kaplan-Meier method, censoring at transplant.
*-- Get year 2017 incident dialysis cohort first --*;

data inc_2017(keep=usrds_id first_se rxgroup);
        set saf.rxhist;
        by USRDS_ID BEGDATE;
        if first.usrds_id;
        incyear=year(begdate);
        if incyear = 2017 and rxgroup NOT IN ("B","X","Z","T");
  run;


*-- Get the transplant date from the Patients file for the incident cohort created above --*;

PROC SQL;
        create table inc_2017_more as
        select a.*, b.tx1date, b.died, b.first_se from inc_2017  as a
        left join saf.patients as b
        on a.USRDS_ID = b.USRDS_ID
        order by USRDS_ID;
QUIT;

*-- Get the patient demographic information from the Patients file for the incident cohort created above --*;

data inc_2017_s;
        set inc_2017_more;
        t = (MIN(died, tx1date, MDY(12,31,2022)) - first_se + 1) / 30.4375; *Calculate the survival time (in months) of
each incident patient with 12/31/2022 as maximum censoring date.;
        c = (MIN(died,, tx1date, MDY(12,31,2022)) = died) AND (MIN(died, tx1date, MDY(12,31,2022)) ^= tx1date);
*Determine whether the patient is censored. 12/31/2022 is used as start date of censoring.;
run;

PROC LIFETEST DATA=inc_2017_s METHOD=KM NOTABLE PLOTS=(s) OUTSURV=surv2017
CONFTYPE=LINEAR lineprinter;
TIME t*c(0);
TITLE '2017 Incident Dialysis Patients Survival Rates'; RUN;
```

EXAMPLE 5: MERGE WITH MEDICAL EVIDENCE FILE (CMS 2728)

```
*-- Demonstrate a way to extract comorbid conditions from the
Medical Evidence File for a pre-defined study cohort (see Example 2 above). –*;
```

```
*-- Pick one record from the ME file since this file has multiple records for one patient, the example below is to pick
the first record --*;

data medevid_1st;
        set saf.medevid;
        by usrds_id  meseq;
        if first.usrds_id;
run;

*-- use the incident cohort created in Example 2 above --*;

DATA inc_2022_me;
MERGE inc_2022_demo  (IN=x1) medevid_1st (IN=x2 KEEP=usrds_id como_canc como_cararr como_chf como_cvatia
como_dm_ins como_diabprim como_dysrhyt como_htn como_ihd como_mi como_copd como_pvd);
BY usrds_id;
IF x1 AND x2;
IF (FIRST.usrds_id); RUN;
```

EXAMPLE 6: 2022 INCIDENT PATIENT SURVIVAL RATES (KAPLAN-MEIER)

```
    *-- Show a survival rate calculation stratified by comorbid patient condition, not censored at transplant. --*;
*-- use the incident cohort created in Example 5 above --*;

DATA inc_2022_s; SET inc_2022_me;
        IF UPCASE(como_canc) IN ("1" "Y")) THEN can = 1; ELSE can = 0;
        t = (MIN(died, MDY(12,31,2022)) - first_se + 1) / 30.4375; * Calculate the survival time (in months) of each
incident patient. For survival time in day, leave denominator 30.4775 out ;
        IF (t < 0) THEN t = 0;
        c = (MIN(died, MDY(12,31,2022)) = died); *  Determine whether the patient is censored.;
RUN;
PROC LIFETEST DATA=inc_2022_s METHOD=KM NOTABLE PLOTS=(s) OUTSURV=surv2022
CONFTYPE=LINEAR lineprinter;
TIME t*c(0);
STRATA can;
   TITLE '20022 Incident ESRD Patients Survival Rates by Co-Morbidity'; RUN;
```

EXAMPLE 7: WAITING LIST ACCESS RATE

```
    *-- Determine waiting list access rate to December 31, 2022 of 2022 incident dialysis patients who were registered
in the transplant waiting list. --*;

DATA txwait;
SET saf.waitseq_ki saf.waitseq_kp;
RUN;


PROC SORT data=txwait;
By usrds_id;
RUN;
```

```
*-- use the incident cohort created in Example 5 above --*;

DATA inc_2022_w;
          * MERGE inc_2022_demo (IN=x1) txwait (IN=x2 KEEP=usrds_id begin);
          * BY usrds_id; IF x1;
          *
          * IF (rxgroup ^= "T"); *  Extract only 2022 incident dialysis patients.;
          *
          * t = (MIN(begin, died, MDY(12,31,2022)) - first_se + 1) / 30.4375; *  Calculate the waiting list access
          time (in month) of each incident patient.;
          * IF (t < 0) THEN t = 0;
          *
          *          c = (MIN(begin, died, MDY(12,31,2022)) = begin);  * Determine whether the patient is
          censored.;
          * RUN;
PROC LIFETEST DATA=inc_2022_w METHOD=KM NOTABLE PLOTS=(s) OUTSURV=wait2022
CONFTYPE=LINEAR lineprinter;
TIME t*c(0);
  TITLE '2022 Incident Dialysis Patients Waiting List Access Rates'; RUN;
```

EXAMPLE 8: CREATE A PATIENT COHORT OF 2022 MEDICARE PRIMARY INCIDENT PATIENTS

   -- Combine payer information with original incident cohort and limit to patients with Medicare primary payers at first service date. --;
-- use the incident cohort created in Example 5 above --;

```
DATA medicare_inc_2022;
       MERGE inc_2022_demo (IN=x1)  saf.payhist(IN=x2 KEEP=usrds_id mcare payer dualelig begdate enddate);
       BY usrds_id;
       IF x1 AND (payer="MPAB" OR payer="MPO") AND begdate <= first_se and (enddate>=first_se or enddate=.);
RUN;
```

EXAMPLE 9: TOTAL ADMISSION RATE

   -- Link incident cohort file with hospitalization file; gather hospitalizations within period of interest (includes only patients with hospitalizations). --;
-- use the incident cohort created in Example 5 above --;

LIBNAME ipclms "c:\saf\claim\in"; * Directory location of the hospitalization data;

```
Data ip_clm22_23;
       Set ipclms.ip_clm_2022 (in=y KEEP=usrds_id clm_from clm_thru)
 ipclms.ip_clm_2023Q1 (in=y KEEP=usrds_id clm_from clm_thru);
run; * some claims with clam_from in year 2022 but clm_thru in year 2023 are in the ip_clm_2023Q1 claim file;

DATA hospdat;
MERGE inc_2022_demo (IN=x)  ip_clm_22_23 (in=y);
BY usrds_id;
IF x AND y;
sfu = first_se + 91;
```

```
efu = MIN(died, first_se+455, mdy(12,31.2022));
IF (died NE .) AND (died < sfu) THEN DELETE;
IF ((sfu>clm_thru) OR (efu<clm_from)) THEN DELETE;
FORMAT sfu efu MMDDYY10.;
RUN;

PROC SORT DATA=hospdat;
  BY usrds_id clm_from clm_thru; RUN;


DATA rate;
SET hospdat;
BY usrds_id;
RETAIN n_hos n_adm exptime rt_adm;
IF FIRST.usrds_id THEN
DO; /* INITIALIZE RETAIN VARIABLES FOR EACH USRDS_ID */ n_hos=0;
n_adm=0; exptime=0; rt_adm=0;
END;
* FOR HOSPITALIZATIONS OVERLAPPING STUDY START ONLY COUNT;
* HOSPITAL DAYS, NOT AS ADMISSION;
IF (clm_from<sfu<=clm_thru) THEN n_hos=n_hos+(MIN(clm_thru,efu)-sfu+1);
* IF WITHIN STUDY PERIOD COUNT DAYS (UP TO STUDY END) AND;
* ADMISSIONS;
ELSE IF (sfu<=clm_from<=efu) THEN
DO; n_hos=n_hos+(MIN(clm_thru,efu)-clm_from +1);
n_adm=n_adm+1;
END;
IF LAST.usrds_id THEN /* Output one record per USRDS_ID */
DO; exptime=(efu-sfu-n_hos+1)/365;
IF exptime>0 then rt_adm=n_adm*1000/exptime;
ELSE IF (exptime<0) THEN exptime=0;
OUTPUT;
  END; RUN;


DATA rateall;


MERGE inc_2022_demo (IN=x1) rate (IN=x2);
BY usrds_id;
sfu = first_se + 91;
efu = MIN(died, first_se+455, mdy(12,31,2022));
IF (died NE .) AND (died < sfu) THEN DELETE;
FORMAT sfu efu MMDDYY10.;
IF x2=0 THEN DO;
        n_hos=0; n_adm=0; rt_adm=0;
```

```
        exptime=(efu-sfu+1)/365;
END;
IF x1;
RUN;
```

```
/*-- Calculate mean admissions per 1,000 patient-years by sex. --*/
PROC TABULATE DATA=rateall;
VAR rt_adm;
CLASS sex;
WEIGHT exptime;
TABLE sex=""*rt_adm=""*mean*f=8.1;
KEYLABEL mean="";
  TITLE1 'Total Admission rates per 1,000 patient years by sex'; RUN;
```

EXAMPLE 10: FIRST ADMISSION RATE

```
*-- Link patients file with hospitalization file; gather hospitalizations within the period of interest
(includes only patients with hospitalizations). --*;
*-- use the hospdat dataset created in Example 9 above --*;

DATA frate;
SET hospdat;
BY usrds_id;
RETAIN n_adm exptime rt_adm flag;
IF FIRST.usrds_id THEN
DO; /* INITIALIZE RETAIN VARIABLES FOR EACH USRDS_ID */ n_adm=0;
exptime=0; rt_adm=0; flag=0;
END;
* FOR HOSPITALIZATIONS OVERLAPPING STUDY START SET EXPOSURE;
* TIME TO ZERO TO EXCLUDE FROM RATE;
  IF (clm_from<sfu<=clm_thru) THEN
  DO; exptime=0; flag=1;
  END;
* IF WITHIN STUDY PERIOD COUNT ADMISSION AND CALCULATE;
* TIME TO ZERO TO EXCLUDE FROM RATE;
  ELSE IF (sfu<=clm_from<=efu) AND flag=0 THEN
  DO; exptime=(clm_from-sfu+1)/365;
  n_adm=n_adm+1; flag=1;
  END;
  IF LAST.usrds_id THEN /* OUTPUT ONE RECORD PER USRDS_ID */ DO;
  IF exptime>0 THEN rt_fadm=n_adm*1000/exptime; OUTPUT;
  END;
RUN;
/*-- Combine result with the original incident sample to include patients without hospitalizations. --*/
```

```
DATA frateall;
MERGE inc_2022_demo (IN=x1) frate (IN=x2);
BY usrds_id; sfu = first_se + 91;
efu = MIN(died, first_se+455);
IF (died NE .) AND (died < sfu) THEN DELETE;
FORMAT sfu efu MMDDYY10.;
IF x2=0 THEN DO;
        n_adm=0;
        rt_fadm=0;
        exptime=(efu-sfu+1)/365;
END;
IF x1;
RUN;

PROC TABULATE DATA=frateall;
VAR rt_fadm;
CLASS sex;
WEIGHT exptime;
TABLE sex=""*rt_fadm=""*mean*f=8.1;
KEYLABEL mean="";
  TITLE1 'First Admission rates per 1,000 patient years'; RUN;
```

EXAMPLE 11: TOTAL COST AND COST BY SERVICE TYPE

```
/*total ESRD Medicare cost for institutional claims 2022*/
PROC SQL;
CREATE TABLE total_in_cost AS SELECT SUM(clm_amt)
AS cost
FROM ipclms.ip_clm_2022;
QUIT;

/*total ESRD Physician cost in 2022 claims line file*/

LIBNAME psclms "c:\saf\claim\ps"; * Directory location of the physician data;

PROC SQL;
CREATE TABLE total_ps_cost AS
SELECT SUM(pmtamt) AS cost
  FROM psclms.ps_line_2022; QUIT;
```

EXAMPLE 12: PAYER STATUS OF TRANSPLANT PATIENTS

```
    /*-- the payer status of 12/31/2022 prevalent transplant patients by age group (0-64, 65+) and years after
transplant (0-3 years, 3+years) --*/

PROC FORMAT; value agegrp
0-64='0-64'
```

```
65-high='65plus';
RUN;


PROC FORMAT; value pl 0-3='0-
3'
3-high='3plus';
RUN;



DATA tx_2022 (KEEP=usrds_id begdate enddate);
SET saf.rxhist;
BY usrds_id begdate;
IF (begdate <= MDY(12,31,2022)) AND
(enddate > MDY(12,31,2022) or (enddate = MDY(12,31,2022) AND (death^='1'))) AND (rxgroup='T') THEN OUTPUT;
RUN;

DATA payhist2022; set saf.payhist;
where begdate<=MDY(12,31,2022) and (enddate>=MDY(12,31,2022) or enddate=.);
RUN;

PROC SQL; create table tx2022_payhist2022 as select a.*,
b.begdate as begdate_payer,
b.enddate as enddate_payer,
b.payer,
b.mcare,
b.dualelig from tx_2022 a left join payhist2022 b on a.usrds_id=b.usrds_id; QUIT;

PROC SQL; create table tx2022_payhist2022_patinfo as select a.*, b.born,b.died, b.first_mcare_pta_start,
b.first_mcare_ptb_start,
b.first_mcare_pta_end, b.first_mcare_ptb_end,
b.first_mcare_pta_status, b.first_mcare_ptb_status,
b.first_mcare_pta_reason, b.first_mcare_ptb_reason
from tx2022_payhist2022 a left join saf.patients b on a.usrds_id=b.usrds_id; QUIT;

DATA tx2022_payhist2022_patinfo2;
set tx2022_payhist2022_patinfo; age_prev2022=floor((mdy(12,31,2022)-born)/365.25);
periodlength_year=floor((mdy(12,31,2022) -begdate)/365.25);
RUN;


PROC FREQ data=tx2022_payhist2022_patinfo2;
table periodlength_year*age_prev2022*mcare /list missing;
format age_prev2022 agegrp. periodlength_year pl.;
RUN;
```

Methods & Databases

SECTION 1: DATA SOURCES AND THE USRDS ESRD DATABASE

A primary objective of the USRDS Coordinating Center is to provide all relevant ESRD data as an integrated and consistent database system for outcomes research. The USRDS ESRD database includes ESRD patient demographic and diagnosis data, biochemical values, dialysis claims, and information on treatment history, hospitalization events, and physician/supplier services.

DATA SOURCES USED IN THE ANNUAL DATA REPORT AND FOR CREATION OF THE SAFs

The data used by the USRDS Coordinating Center originates from the End Stage Renal Disease Quality Reporting System (EQRS), the Centers for Medicare & Medicaid Services (CMS), the Organ Procurement and Transplantation Network (OPTN), the Centers for Disease Control (CDC), the ESRD Networks, the U.S. Census, and select data from past and current USRDS Special Studies.

The End Stage Renal Disease Quality Reporting System (EQRS)

The primary source of ESRD patient information for the USRDS is currently the End Stage Renal Disease Quality Reporting System (EQRS). The EQRS database contains demographic, diagnostic, treatment history and vital information for all patients with ESRD.

The original CMS ESRD database was named the Program Management and Medical Information System (PMMIS); this was replaced by the Renal Beneficiary and Utilization System (REBUS) in 1995. Having advanced its database technology, in the fall of 2003 CMS migrated the REBUS database into an Oracle relational database system named the Renal Management Information System (REMIS). In 2003, the Standard Information Management System (SIMS) database of the ESRD Networks was also established; SIMS included information to track a patient's movement in and out of ESRD facilities and their transitions from one treatment modality to another. The REMIS database originally included all patients who were alive and had ESRD as of January 1, 1995, or were incident after this date, an approach adopted from the procedure used to create the SIMS database. However, because the REMIS system as it existed did not include legacy patients for longitudinal studies, CMS expanded it in the fall of 2004 to include all ESRD patients. Together, REMIS and SIMS comprised the CROWN system. In May 2012, Internet-based access to the data system, CROWNWeb, was rolled out nationally. It replaced the functionality of SIMS, interfaced with REMIS, and also provided new data to support the calculation of clinical measures. CMS regularly updates the REMIS/CROWNWeb database, using the Medicare Enrollment Database (EDB), Medicare inpatient and outpatient claims, the Organ Procurement and Transplantation Network (OPTN) transplant database, ESRD Medical Evidence Report forms (CMS 2728), and ESRD Death Notification forms (CMS 2746). CMS has established data integrity rules to ensure accurate identification of patients in the CMS databases. Each ESRD patient (new or existing) is identified with a unique patient identification number common to both databases, guaranteeing that data for all patients are consistently managed over time. In 2020, CROWNWeb, ESRD Quality Incentive Program (QIP), and REMIS were integrated into a single, new system, the End Stage Renal Disease Quality Reporting System (EQRS).

Working with data from the EQRS, the USRDS Coordinating Center identify new patients, assign the USRDS_ID, establish the first ESRD service data, treatment history, and Medicare coverage for all ESRD patients. The Medicare claims data can add more details for Medicare-covered ESRD patients.

CMS ESRD Medical Evidence Report Form (CMS 2728)

The CMS ESRD Medical Evidence Report form (CMS 2728) is used to register patients at the onset of ESRD and must be submitted by dialysis or transplant providers within 45 days of treatment initiation. It establishes Medicare eligibility for

individuals who previously were not Medicare beneficiaries, reclassifies previously eligible Medicare beneficiaries as ESRD patients, and provides demographic and diagnostic information for all new ESRD patients regardless of Medicare entitlement. Patient demographics, primary diagnosis, comorbidities, and biochemical test results at the time of ESRD onset are collected on this form.

Before 1995, dialysis units and transplant centers were required to file the Medical Evidence Report form only for Medicare-eligible patients. Since the 1995 revision, however, providers are required to complete the form for all new ESRD patients regardless of Medicare eligibility status. The 1995 revised form included new fields for comorbid conditions, employment status, expanded race categories, ethnicity, and biochemical data at ESRD onset. The third major revision of the Medical Evidence form in May 2005 remedied several shortcomings of the 1995 form and its earlier versions and is described in more detail in the next section. Key additions target pre- ESRD care and vascular access use, and additional new fields collect information on glycosylated hemoglobin (HbA1c) and lipid testing, the frequency of hemodialysis (HD) sessions, and whether patients are informed of transplant options. The Medical Evidence form is our most reliable source of information about the primary cause of a patient's ESRD. Because the list of diseases has been revised, the USRDS stores the codes from each Medical Evidence form version so that detail is not lost through conversion of one set of codes to the other.

Only one Medical Evidence form (CMS 2728) is expected for most ESRD patients; however, multiple forms may be filed for patients who need to re-gain the ESRD Medicare benefits. For example, transplant patients with functioning grafts lose Medicare benefits after three years if ESRD was the sole qualification for Medicare eligibility. If such a patient experiences graft failure and returns to dialysis, a second Medical Evidence Report must be filed to reestablish Medicare eligibility. Dialysis patients who discontinue dialysis for more than 12 months also lose Medicare ESRD benefits. If such a patient returns to dialysis or undergoes a kidney transplant, a second Medical Evidence form must be filed to reestablish Medicare eligibility.

2005 Revision of the Medical Evidence Report Form (CMS 2728)

The revision of the Medical Evidence Report form (CMS 2728) introduced in May 2005 includes new data collection methods and new variables. It allows users to specify whether the Medicare registration is initial (new ESRD patient), re-entitlement (reinstating Medicare entitlement after a lapse due to no claims being filed for 12 or more months or a functioning graft for 36 or more months), or supplemental (updating missing or incorrect information). This clarifies the intended use of the form without recourse to the "First Regular Dialysis Start Date," and helps chronicle the historical sequence of multiple forms for the same patient.

Multiple patient races can be specified. CMS provides a single variable containing a concatenated string representing all selected race codes with binary digits (0, 1). This string must be decoded to determine patient race (or races). Similar formatting requirements apply to information on comorbid conditions, medical coverage, and reasons for not informing patients of transplant options. Because the required programming is substantial, for each of these four data categories, we include the original variable with the concatenated string and create a new variable with the decoded values (i.e. information noted on form CMS 2728). However, the decoded multiple race values must be presented as "Other" or "Multiple race" to maintain consistency with the legacy race information.

Data fields for nephrologist care, dietitian care, and access type were also added, with their respective time intervals relative to ESRD onset. Data on the laboratory values hematocrit, creatinine clearance, BUN, and urea clearance are no longer collected. Added laboratory values include HbA1c and lipid profiles (TC, LDL, HDL cholesterol, and TG). Additional fields relate to whether patients were informed of transplant options (and if not, why not), donor type, comorbid conditions, and primary diagnoses were also added.

2015 Revision of the Medical Evidence Report Form (CMS 2728). Changes in the 2015 Medical Evidence Report form include:

- ICD-10-CM codes are used to report item 15, *Primary Cause of Renal Failure*. A reference list of ICD-10-CM codes for primary cause of renal failure is included.

- An option for <6 months was added to the following three questions in item 18, *Prior to ESRD therapy*:
  - Did patient receive exogenous erythropoietin or equivalent?
  - Was patient under care of a nephrologist?
  - Was patient under care of kidney dietitian?

- The date is no longer required in the following two fields in item 19: *a.2 Serum Albumin Lower Limit* and *a.3 Lab Method Used (BCG or BCP)*

2018 Revision of the Medical Evidence Report Form (CMS 2728). Changes in the 2018 Medical Evidence Report form include:

- Removal of Medicare Claim Number and addition of Medicare Beneficiary Identifier and renumbering of fields
- Inclusion of 'Other' in Race field
- Transplant options field revised to remove 'Medically unfit', "Psychologically unfit' and 'Unsuitable due to age' and included a new selection 'Patient is not eligible medically'.

2024 Revision of the Medical Evidence Report Form (CMS 2728). The 2024 Medical Evidence Report form (started from October 2024) has some substantial changes from the 2018 version. We list some significant changes here. Many minor changes were made, such as the addition of a category "Volunteer" to the occupation status, the addition of some diagnosis codes for cause of ESRD, etc. We will not list them all.

- Separated Medicare Beneficiary Identifier and social security number as two fields.
- Added detailed subcategories for Asian and Native Hawaiian or Pacific Islander and added Middle Eastern or North Africa as a separate race category.
- Added Social Determinants of Health (SDOH) related questions, such as Housing Instability, Food Insecurity, and Transportation, etc.
- Added "Patient Education" type questions, such as has been informed/understand about kidney transplant; has been informed/understand about home dialysis, etc.
- Added questions regarding the patient's ability to understand health literature in English or if they require a translator
- For lab testing, Cystatin C was added to the form.

The 2025 version of the Medical Evidence form was on-line since July 2025. Some minor changes were made from the 2024 version and it will be included in next year's SAF and Researcher's Guide.

CMS ESRD Death Notification Form (CMS 2746)

The ESRD Death Notification form (CMS 2746) is used to report the death of ESRD patients. According to CMS policy, this form must be submitted by dialysis or transplant providers within 30 days of a patient's death, and provides the date, primary and secondary causes of death, reasons for discontinuation of renal replacement therapy, if applicable, and evidence of hospice care prior to death. It is the primary source of death information for CMS and the USRDS. The USRDS also utilizes several supplemental data sources for ascertaining date and cause of death, which include the Medicare enrollment database and OPTN data. The form also changes from version to version. But the changes are usually very minor.

CMS Medicare Enrollment Database (EDB)

The CMS Medicare EDB is the designated repository of all Medicare beneficiary enrollment and entitlement data, including current and historical information on beneficiary residence, Medicare coverage status, and Health Insurance Claim/Beneficiary Identification Code (HIC/BIC) cross-referencing.

Organ Procurement and Transplantation Network Database

CMS began collecting data on all Medicare kidney transplants in the early 1980s. In 1984, the National Organ Transplant Act called for the creation of an Organ Procurement and Transplant Network (OPTN) to collect data and maintain a registry for organ matching and transplantation. The United Network for Organ Sharing (UNOS) was awarded the OPTN contract in 1988 to provide a national system to allocate donor organs and maintain a centralized data depository for organ transplants. OPTN also collects data on other organ transplants, including lung, liver, heart, pancreas, and intestine transplant. The OPTN and CMS collection efforts were consolidated in 1994, and only OPTN continues to collect data on transplant donors and recipients. However, with EQRS, transplant centers also input some patient information into the system since the end of 2020.

CMS ESRD claims Files

CMS claim data includes three parts: Medical claims for beneficiaries enrolled Medicare fee-for-service (FFS) program; Medical claims for beneficiaries enrolled Medicare Advantage (MA) programs; Medical claims for Medicaid enrollees. USRDS has used claim files for Medicare FFS enrollees from the very beginning. Recently, USRDS started receiving claims for MA enrollees and Medicaid enrollees. All three parts are used for the Annual Data Report. However, only claims for Medicare FFS enrollees are released for researcher use based on the data use agreement with CMS.

The CMS ESRD claims files contain data from final action claims submitted by Medicare FFS beneficiaries, in which all adjustments have been resolved. For Part A Institutional claims, the USRDS uses the following 100% SAF claims: Inpatient, Outpatient, Skilled Nursing Facility, Home Health Agency, Hospice. For Part B Physician/Supplier, the USRDS uses the following 100% SAF claims: Physician/Supplier, Durable Medical Equipment.

CMS files are updated each quarter through June of the following year when the CMS annual files are finalized. CMS data for the current year are created six months into the year and are updated quarterly until finalized at 18 months, after which the files are frozen and will not include late arriving claims. Annual claims files are approximately 98% complete. For more information on CMS claims, see Section 5: Medicare Claims.

CMS 5% General Medicare claims files

The CMS 5% general Medicare SAF has the same structure and data elements as the ESRD 100% SAF, but the data were extracted from the general Medicare billing database as a random sample of 5% of the entire Medicare population. Because the sample was selected using the last two digits of patient Social Security Numbers, the same group of Medicare beneficiaries is represented in the CKD 5% SAFs each year, with exceptions for death, disenrollment, and new Medicare entitlements.

The USRDS Coordinating Center uses these files to conduct studies on Healthy People 2030 objectives, comparing preventive care and other non-ESRD disease treatments in general Medicare and ESRD patients. Also, these files are used to create the CKD cohort dataset.

CMS Annual Facility Survey

All Medicare-certified dialysis units and transplant centers are required to complete the CMS Annual Facility Survey (AFS; CMS 2744) at the end of each calendar year. The AFS reports counts of patients being treated at the end of the year, new ESRD patients starting during the year, and patients who died during the year. Counts of Medicare and non-Medicare end-of-year patients are included. While AFS files do not carry patient-specific demographic and diagnosis information, they do provide independent patient counts used to complement the CMS patient-specific records.

M Health Fairview Data (2005-2021)

The M Health Fairview clinical data are integrated electronic health records (EHR) from M Health Fairview, a partnership between the University of Minnesota and the Fairview Health System, a community-based, non-profit provider ranked by Thomson Reuters as one of the top 10 health care systems in the US. The M Health Fairview Clinical Data Repository houses EHR data for ~2.5 million patients from 8 hospitals and more than 40 clinics. Although most M Health Fairview patients are in the metropolitan Minneapolis/St. Paul area, the M Health Fairview network includes clinics across Minnesota, encompassing a population that is both urban and rural. Data includes demographic characteristics, medical history, medical problem lists, immunizations, inpatient and outpatient encounters, surgeries, procedures, diagnoses, laboratory results, basic vital signs, prescribed medications, insurance type, social behavioral characteristics, and more. Data are available only for those who had an eGFR ≤60 ml/min/1.73m2 after January 1, 2005 or the earliest date when the data was entered into the system after January 1, 2005. The USRDS 2025 SAFs for M Health Fairview include data up to May 2022.

U.S. Census

The U.S. population data are from the 2000 and 2010 intercensal estimates from https://www2.census.gov/programs-surveys/popest/datasets/2000-2010/intercensal/; the 2011-2020 postcensal estimates from https://www2.census.gov/programs-surveys/popest/datasets/2010-2020/counties/asrh/ and https://www2.census.gov/programs-surveys/popest/datasets/2010-2020/national/asrh/; and 2021-2023 postcensal estimates from https://www2.census.gov/programs-surveys/popest/datasets/2020-2023/. The 2000, 2010, and 2020 data are actual census data. The 2025 ADR also used the population data received from the Surveillance, Epidemiology, and End Results (SEER) Program (https://seer.cancer.gov/popdata/download.html).

The data from census use race categories as specified in the 1997 Office of Management and Budget (OMB) standards for the collection of data on race and ethnicity: American Indian or Alaska Native, Asian, Black or African American, Native Hawaiian or Pacific Islander, White, More than one race. While data from SEER is a bridged-race population data, which split the "more than one race" group to the 5 other single-race groups using an algorithm. The bridged-race is more proper for race-specific ESRD rate calculation or race-related analyses because few ESRD patients were reported as multiracial.

USRDS ESRD Database System

The USRDS Coordinating Center has developed a centralized ESRD patient database by integrating data from the above data sources and establishing methods to identify patients with ESRD. We use this database to update and maintain data on demographics, clinical measurements, biochemical lab test results, renal replacement therapies, treatment history, and all medical service events reported in the Medicare Claims data. Through this patient-oriented database, we attempt to define each ESRD patient through multiple stages of data cleaning, conversion, validation, and consolidation. Establishment of a universal patient identification system was critical to ensure that unique patient identification numbers are assigned to each ESRD patient at the time of ESRD initiation and to accurately track counts and rates of incident and prevalent cohorts over time.

In 1994, the USRDS Coordinating Center introduced the Standard Analysis Files (SAFs) — not to be confused with the CMS Standard Analytical Files for Part A, and Part B Claims data — to help meet the ESRD data needs of a wide variety of research studies. These SAFs were subsequently enhanced to include not only ESRD clinical and claims data from CMS but also transplant and waiting list data from OPTN. All SAFs are indexed or sorted by the unique USRDS-specified patient identification number, and patient identifiers (i.e. name, address, SSN, HIC/BIC) are removed to protect patient confidentiality. The USRDS Coordinating Center also uses the USRDS ESRD database to generate datasets used to create the tables, graphs, and maps in the USRDS ADR.

The USRDS ESRD database is updated quarterly with data obtained from the various data sources. The USRDS Coordinating Center receives dialysis clinical data from EQRS on a monthly basis; Facility Survey data, Patient Lists and Patient Events, Medical Evidence Report, and Death Notification data, on a quarterly basis; OPTN transplant and waiting list data twice a year; and Medicare EDB, and CMS medical claim files on a yearly basis. These multiple updates allow the USRDS Coordinating Center to assess the growth of the ESRD population, the demographic distribution of ESRD patients, and changes in the percentage of patients diagnosed with major diseases secondary to ESRD.

SECTION 2: ESRD PATIENTS

ESRD is defined as chronic renal failure requiring renal replacement treatment — dialysis or transplant — to sustain life. It is not the same as acute renal failure, from which patients are expected to recover within weeks or months. Dialysis facilities and transplant centers must immediately complete a Medical Evidence Report for all ESRD patients, to register them in the CMS ESRD database, and to apply for Medicare eligibility if they were not previously eligible.

To establish the incident and prevalent cohorts by year, modality, primary cause of renal failure, and other factors, researchers require patient demographic, clinical, and treatment history data. This information can be obtained from the USRDS Core SAF dataset from the following files: PATIENTS, MEDEVID, RXHIST, RXHIST60, and PAYHIST.

Identifying ESRD Patients

A person is identified as having ESRD when a physician certifies the disease on the Medical Evidence Report form (CMS 2728), or when there is other evidence of chronic dialysis that meets the criteria of ESRD or eligibility for a kidney transplant. Patients with acute kidney failure who are on dialysis for days or weeks, but who then recover kidney function, are excluded from the database if their Medical Evidence forms have not been submitted. Patients who die soon after kidney failure without receiving dialysis are sometimes not included in the database.

The USRDS coordinating center receives EQRS data quarterly. Algorithms are used to identify new patients and existing patients in newly received EQRS data based on personally identifiable information (PII). Data for existing patients is linked with their extant data in the database, and new patients are assigned a USRDS ID (USRDS_ID). Identifying new and existing patients in the new data is time-consuming and takes multiple iterations.

First ESRD Service Date

The first ESRD service date (FSD) is the single most important data element in the USRDS database, and each patient must, at a minimum, have a valid FSD. This date is used to determine each new patient's incident year and the first year in which the patient is counted as prevalent. The date 90 days after the FSD is used as the starting point for many patient outcomes analyses. This 90-day rule allows each new ESRD patient to obtain Medicare services despite potential delays in completing the Medicare eligibility application process, and it provides an adequate period for patients to arrive at a stable and suitable dialytic treatment modality. In most cases, the first service date is derived by identifying the earliest date of various potential indicators:

- the start of dialysis for presumed ESRD as reported on the Medical Evidence form;
- the first EQRS event;
- a kidney transplant as reported on a CMS or OPTN transplant form, a kidney transplant as reported on EQRS transplant, or a Medical Evidence form.

Death Date Determination

After the ESRD first service date, the date of death is the next most critical piece of information in the USRDS database. Death dates are obtained from several sources: the CMS Medicare Enrollment Database (EDB), CMS forms 2746 and 2728, the OPTN transplant follow-up worksheet/form, and/or the EQRS event file. Because multiple sources report death information for the same patient, an individual may have several reported dates. For these patients, the accepted death date is based on the hierarchical order below:

- CMS 2746 Death Notification form
- EQRS patient file (from the CMS EDB)
- EQRS events
- OPTN Transplant follow-up data and transplant data
- CMS 2728 Medical Evidence form

Transplant Dates

Transplant events can be identified from the OPTN data, the Medical Evidence forms indicating kidney transplant as the initial modality, from the EQRS event file, and from the EQRS transplant events. To resolve any conflicts among the data sources, and to create a complete list of unique transplant events, the USRDS has adopted the following procedures:

- Before 1988, all transplant events found in CMS PMMIS/REBUS/REMIS Transplant files are used.
- Between 1988 and 1993, all transplant events found in OPTN files are used and additional transplant events from the CMS PMMIS/REBUS/REMIS are recorded.
- After 1994, all transplant events found in OPTN files are used.
- Additionally, transplant events reported on the Medical Evidence form are used, as are transplants found in the EQRS patient events data.

Each transplant event found in the Transplant File of the USRDS Core SAF dataset is thus a unique event derived from the OPTN database, the CMS Transplant database, Medical Evidence Report records, EQRS patient events.

Graft Failure

The USRDS Coordinating Center assumes a (kidney) graft failure date is correct as reported in the OPTN transplant follow-up file or the EQRS event file unless death or a new transplant occurs before this date. A graft failure date may not be recorded in either file, however. In this case, we use the earliest of the following events:

- date of death
- date of subsequent transplant
- date of return to regular dialysis, indicated by a continuous period of dialysis billing records covering a minimum of 60 days or
- date of return to dialysis reported on the Medical Evidence form, or the date of graft nephrectomy from the OPTN transplant follow-up record or a Medicare claim

Medicare and Non-Medicare Patients

Beneficiaries are enrolled in Medicare based on criteria defined in Title XVIII of the Social Security Act of 1965 and subsequent amendments to the Act. A person in one of these four categories is eligible to apply for Medicare entitlement:

- Aged 65 years and older
- Disabled
- Enrolled in the ESRD program
- Railroad Retirement Program beneficiary

Most ESRD patients are eligible to apply for Medicare as their primary insurance payer. Some, however, are not immediately eligible for Medicare primary payer coverage because of their employment status and pre- existing primary insurance payers. These patients are covered by payers such as Employer Group Health Plans (EGHPs), the Department of Veterans Affairs, and private insurers. These patients will have first service dates established by the Medical Evidence forms or EQRS events, but will not have any dialysis claims or hospitalization events in the CMS claims database. Patients covered by EGHPs need to wait 30-33 months before becoming eligible for Medicare as primary payer. They can apply for Medicare as the secondary payer status before changing to Medicare as the primary payer.

Treatment history events can be tracked by the EQRS Events File, which follows changes in modality, transplants, transfers between facilities, and death. All ESRD patients, regardless of their Medicare eligibility status, are in the EQRS system.

The USRDS recognizes that non-Medicare patients are "true" ESRD patients and should, therefore, be included in patient counts for incidence, prevalence, and treatment modality, as well as in mortality and transplant rate calculations. Calculations of hospitalization statistics or any outcomes derived from Medicare claims (e.g., any specific diagnostic or therapeutic code), however, should only include these patients with Medicare as the primary payer. It is important to understand that a fraction of the patients in the USRDS database do not have Medicare as their primary payer at any given time. We suggest that for any claims-based analyses cohorts should be constructed using this file (PAYHIST).

SECTION 3: TREATMENT HISTORY

The USRDS uses all available data to create a treatment history for all ESRD patients in the database, including all modality events, their duration, and the renal providers involved in each patient's care. This history can be used to identify the incident and prevalent cohorts and determine the censoring points and outcomes for observational studies. The EQRS event database is the primary source of the modality sequence file, and the dialysis claims are used as a way of confirming placements and identifying problem cases. As described in the previous section, we use all available sources to determine first service dates, deaths, transplants, and transplant failures.

Treatment Modality Categories

Table 3 lists modality categories used by the USRDS. They can be described as detailed or general. For most analyses, the general categories, which combine detailed modality categories, are sufficient. The Recovered Renal Function (RRF) modality/event is similar to the lost-to-follow-up event in that the patients with an RRF event are not included in the prevalent populations for outcomes analyses. However, as with lost-to-follow-up events, these patients remain in the modality sequence so that subsequent renal episodes can be closely and promptly tracked.

Table 3. Treatment modality categories

Type	Detailed
1	In Center Hemodialysis (HD)
2	In Center Self HD
3	Home HD
4	HD Training
5	Continuous Ambulatory Peritoneal Dialysis (CAPD)
6	CAPD TRAINING
7	Continuous cyclic peritoneal dialysis (CCPD)
8	CCPD Training
9	Other Peritoneal Dialysis
10	Nursing Home HD
A	Uncertain Dialysis
B	Discontinued Dialysis
T	Transplant
X	Lost-to-follow-up
Z	Recover Kidney Function

How Treatment Modality is Determined

The Treatment History Files in the USRDS Core SAF dataset record the sequence of modalities for each patient. The variables included in these files are presented in Table 4. Each record in the file indicates a period of therapy with a given modality and any change in provider or detailed modality results in a new record. Several data sources are used to determine the treatment history, including the Medical Evidence file, EQRS Event file, OPTN Transplant Events File, Medicare Claims Files, and Death Notification form (CMS 2746).

Table 4. Treatment history SAF variables

File name	Variable name	Variable label
RXHIST (Detailed Treatment History)	**USRDS_ID**	USRDS patient identification
	BEGDATE	Beginning date of a modality period
	ENDDATE	Ending date of a modality period
	BEGDAY	Start day of modality period (First Service Date = 1)
	ENDDAY	End day of modality period (First Service Date = 1)
	RXDETAIL	Detailed treatment modality for a period
	RXGROUP	Grouped treatment modality for a period
	PROVUSRD	USRDS assigned facility identification
	DEATH	Death: 1 if enddate=deathdate; 0 if alive
RXHIST60 (Condensed Treatment History)	**USRDS_ID**	USRDS patient identification
	BEGDATE	Beginning date of a modality period
	ENDDATE	Ending date of a modality period
	BEGDAY	Start day of modality period (First service date = 1)
	ENDDAY	End day of modality period (First service date = 1)
	RXGROUP	Grouped treatment modality for a period
	DEATH	Death: 1 if enddate=deathdate; 0 if alive

In constructing the Treatment History File (RXHIST), these conventions are followed:

- The sequence always begins with the first ESRD service date and ends with the earliest of date-of-death or end of the period for which the data are complete. Other categories always have an end date. If the patient is still alive, the end date of last period is assigned a study end date, e.g. in the 2025 SAF, the end date of last period is "12/31/2024".

- If two dialysis billing periods overlap, the earlier is terminated at the start of the later period.

- A functioning transplant is considered treatment, not recovery from ESRD.

- If a dialysis billing period overlaps a transplant date, the dialysis period is closed at the transplant date.

- If a graft failure is recorded in the database, but there are no subsequent records of dialysis, or death, or another transplant, a period of "unknown dialysis" is inserted. If no dialysis billing records or re- transplant appears within one year, the patient is designated as lost-to-follow-up at the end of a one-year period, and this status continues until dialysis or new transplant records appear.

- Once established, a modality is assumed to continue until a change in provider or detailed modality occurs. For dialysis patients, if no information on treatment modality appears for one year, the patient is reclassified as lost-to-follow-up at the end of that one-year period.

- A functioning graft is assumed to continue until an indication of graft failure or death appears, or evidence of regular or maintenance dialysis.

The 60-Day Collapsing Rule

The USRDS Coordinating Center uses the convention that a dialysis modality must continue for at least 60 days to be considered stable. A transplant is considered a stable modality regardless of duration.

Because the dialysis treatment history is derived from the EQRS Events File, the Medical Evidence Report, OPTN Transplant Events, and Medicare billing records, it includes intermixed and overlapping details that are not required or are unsuitable for most analyses. A long period of CAPD, for example, may be interrupted by a short period of inpatient hemodialysis treatment, or a patient may go on vacation and receive dialysis from a different provider. By applying the 60-day rule, we collapse modality periods of less than 60 days and periods that differ only by the provider, giving a less complex treatment history for analyses that require less detail than is available in the detailed treatment history. For maximum flexibility, we provide two Treatment History Files, one with full detail (RXHIST), and one applying the 60-day collapsing rule (RXHIST60). Tables 5 and 6 show examples of how these two files differ. The 60-day file is constructed from the detailed file as follows:

- Consecutive records with the same dialysis modality group are collapsed into a single record regardless of the length of the period, detailed dialysis modalities as well as changes in the provider. Thus the collapsed file cannot be used to track the changes of provider and detailed dialysis modalities within the same general dialysis modality group.

- For events with a duration period of less than 60 days, the dialysis modality periods are labeled 'uncertain dialysis' and consecutive dialysis records labeled 'uncertain dialysis' are combined. This rule does not apply to functioning transplant and other non-dialysis modalities.

- If a dialysis modality is interrupted by a different dialysis modality that lasts less than 60 days (e.g. a period of peritoneal dialysis between two longer periods of hemodialysis), the short dialysis modality period is ignored, and the longer dialysis modality extends over the entire period. This rule is not applied to functioning transplant and other non-dialysis modalities.

Table 5. RXHIST examples

USRDS_ID	BEGDATE	ENDDATE	PROVUSRD	DEATH	RXDETAIL	RXGROUP
5	2/10/1992	6/1/1992	922	0	Hemodialysis	Hemodialysis
5	6/2/1992	1/16/1994	922	1	CAPD	CAPD
6	2/7/1974	6/12/1974	1355	0	CAPD Training	CAPD
6	6/13/1974	3/30/2008	1355	0	Transplant	Transplant[a]
6	3/31/2008	4/21/2008	7635	0	CAPD	CAPD [b]
6	4/22/2008	9/5/2008	7187	0	Hemodialysis	Hemodialysis
6	9/6/2008	10/5/2008	5130	1	Hemodialysis	Hemodialysis
7	3/20/1981	4/12/1981	1822	0	Hemodialysis	Hemodialysis
7	4/13/1981	6/4/1981	1822	0	Hemodialysis	Hemodialysis
7	6/5/1981	7/5/1981	1822	0	Transplant	Transplant[a]
7	7/6/1981	10/1/1981	1822	0	Hemodialysis	Hemodialysis
7	10/2/1981	5/7/1982	1822	0	Hemodialysis	Hemodialysis
7	5/8/1982	9/7/1982	1822	0	Hemodialysis	Hemodialysis
7	9/8/1982	4/8/1998	1822	0	Transplant	Transplant[a]
7	4/9/1998	4/10/1998	3310	1	Hemodialysis	Hemodialysis[c]
8	9/12/1983	9/30/1983	2081	0	CAPD Training	CAPD
8	10/1/1983	12/31/1983	2081	0	CAPD	CAPD
8	1/1/1984	6/19/1984	2081	0	CAPD	CAPD
8	6/20/1984	7/17/1984	2081	0	Transplant	Transplant[a]
8	7/18/1984	9/30/1984	2081	0	CAPD	CAPD
8	10/1/1984	1/31/1985	2081	0	CAPD	CAPD
8	2/1/1985	3/10/1985	2081	1	CAPD	CAPD

Table 6. RXHIST60 examples

USRDS_ID	BEGDATE	ENDDATE	DEATH	RXGROUP
5	2/10/1992	6/1/1992	0	Hemodialysis
5	6/2/1992	1/16/1994	1	CAPD
6	2/7/1974	6/12/1974	0	CAPD
6	6/13/1974	3/30/2008	0	Transplant[a]
6	3/31/2008	10/5/2008	1	Hemodialysis
7	3/20/1981	6/4/1981	0	Hemodialysis
7	6/5/1981	7/5/1981	0	Transplant[a]
7	7/6/1981	9/7/1982	0	Hemodialysis
7	9/8/1982	4/8/1998	0	Transplant[a]
7	4/9/1998	4/10/1998	1	Uncertain Dialysis[c]
8	9/12/1983	6/19/1984	0	CAPD
8	6/20/1984	7/17/1984	0	Transplant[a]
8	7/18/1984	3/10/1985	1	CAPD

[a] The 60-day rule does not apply to transplants. [b] CAPD record less than 60 days in RHIST, collapsed down with next dialysis modality in RXHIST60.
[c] Hemodialysis record less than 60 days in RHIST. It is the patient's last record, with previous event "Transplant," so it cannot be collapsed down or up with the previous or next event. Recode to "Uncertain Dialysis" in RXHIST60.

Integration of the EQRS and CMS Claims Databases

To define treatment modality, EQRS data usually has a higher priority than claims data. For patients who either do not appear in the EQRS Events File or for whom the only event is "New ESRD Patient," and patients who have transfer-out gaps in EQRS event file, Medicare dialysis claims were used to fill the gaps.

However, claims data capture the modality "Center Self Hemodialysis" more accurately than the EQRS data; this claims-based designation overrides other dialysis modalities from EQRS. Any EQRS dialysis event that falls into a "Center Self Hemodialysis" period as determined by claims is recoded as "Center Self Hemodialysis."

Events that are implausible are removed. These include events that occur before a patient's first service date, those falling between "Transplant" and "Transplant Failure," "Transfer Out for A Transplant" events that occur 60 days or less after the corresponding "Transplant," and events occurring after "Death."

Determining Lost-to-Follow-up Periods

Gaps frequently exist in the EQRS and billing data upon which modality periods are based. The USRDS assumes that a modality continues until death or the next modality-determining event. A patient with a functioning transplant is assumed to maintain it unless a new EQRS event, claim event, or death date is encountered in the data. A dialysis modality, in contrast, is assumed to continue for only 365 days from the date of discharge from a facility without next dialysis facility admission, in the absence of a transplant or death date or dialysis claims. After this period, the patient is declared lost- to-follow-up, until the occurrence of a new EQRS event, dialysis claim, or transplant event. Discharged with reason of "recovered renal function" was defined as "recovered renal function", not "lost-to-follow-up".

Patients are considered lost-to-follow-up beginning 365 days after a "Transplant Failure" event or "Discontinued Dialysis". Patients for whom the only event is a first service date, and who do not exist in any other files were also treated as lost-to-follow-up, beginning one year after the first service date. Some events can result in a lack of dialysis data and eventual reclassification of a patient as lost-to-follow-up:

- The patient has recovered renal function (RRF),but without a discharge reason of "recovered renal function".

- The patient no longer resides in the United States.

- The patient died, but has not been reported to the Social Security Administration or CMS.

SECTION 4: PAYER HISTORY

ESRD Payer Sequence File

The ESRD Payer History File is similar to the Treatment History File; it contains a sequence of payer status for each ESRD patient. CMS payer information is used to create a continuous sequential history of payers for each patient, beginning with the first ESRD service date (FSD). Each patient's FSD in the Payer History File is the same date reported in the Treatment History File. Data from the Medicare Enrollment Database (EDB) and dialysis claims information are used to categorize payer status (Table 7).

The EDB file provides information on payer with categories listed in Table 7 from their enrollment date to Medicare, except for the "OTH" category. All patients who were not found in the EDB were classified as Other (OTH). The Other can include EGHP not enrolled in Medicare, VA administration, private insurance, and uninsured. Medicare dialysis claims were just used as a supplement to define "Medicare as the primary payer (MPP)" or "Medicare as the second payer (MSP)" status when there is no information from the EDB.

The Payer History File in the USRDS Core SAF records the sequence of payers for each patient. Table 8 shows the variables included in this file. Each record in the file represents a period covered by a particular payer. Any change in payer, Medicare status, or dual eligibility status results in a new record. If a patient is with a payer and still with this payer at the end of the current data period, we assign the end date of this data period to ENDDATE of the corresponding record. For the 2025 SAF, the end of data period is December 31, 2024.

Medicare claims are complete for MPP enrollees only. Therefore, claim-based analysis, such as hospitalization and medical cost, should include MPP patients only. The Pay history file can be used to identify those patients in the study period.

Table 7. Payer categories

Code	Description
HMO	Health Maintenance Organization (HMO) plan, or Medicare Advantage Plan
MPAB	Medicare FFS Primary pay, for both Part A and Part B
MPO	Medicare FFS Primary pay, for Other
MSP-EGHP	Medicare FFS Secondary Payer with Employer group health plan (EGHP)
MSP-nonEGHP	Medicare FFS Secondary Payer, without Employer group health plan (EGHP)
OTH	Other/Unknown, including VA, EGHP, but not enrolled in Medicare
WAIT	90 Day Waiting Period

Pre-ESRD Payer Sequence File

The Pre-ESRD Payer Sequence File can be used to determine payer status for the periods before onset of ESRD. This Payer Sequence File is similar to the standard ESRD Payer Service File except that the Pre- ESRD Payer Sequence File begins at the first evidence of Medicare enrollment from the Enrollment Database, rather than first ESRD service date, as is the case with the ESRD payer sequence. The pre-ESRD payer sequence ends the day before the first ESRD service date.

Table 8. Payer History SAF variables

Variable label	Variable description
USRDS_ID	USRDS Patient Identification
BEGDATE	Beginning date of a payer period
ENDDATE	Ending date of a payer period
PAYER	Payer for the period
MCARE	Payer includes Medicare (Y/N)
DUALELIG	Payer includes Medicare and Medicaid (Y/N)

Example of Using the Payer History File

One example of using payer history file is SAS Example 12. Transplant patients with functioning grafts would lose Medicare benefits after three years if ESRD was the sole qualification for Medicare eligibility. We want to see how much the data reflect this rule in this example. The example shows how payer history file and treatment history files are used to investigate the payer status change for 2015 prevalent transplant patients after transplant. Patients aged 65 and over are automatically enrolled in Medicare so this factor is included in the analysis.

SECTION 5: MEDICARE CLAIMS

Introduction

The USRDS Coordinating Center has created files from CMS billing data and incorporated them into the USRDS database. These files include all Medicare claims after onset of ESRD and claims before the start of ESRD. Note, the claim data is about one year behind the Core SAF data as previously discussed. For the 2025 SAF, the claim data is complete until December 31, 2023, except for the 2024 quarter one inpatient file. While the CORE SAF contains data until December 31, 2024..

CMS Data Sources

Medicare claims are of two types: Physician/Supplier claims and Institutional claims . The structure and content of the two types of claims are different, as are the files derived from them. The Institutional claims files include:

- Inpatient
- Outpatient
- Skilled Nursing Facility
- Home Health Agency
- Hospice

For Institutional and Physician/Supplier claims files, data for a given year is frozen at the end of June of the subsequent year, so claims submitted after June of the year following the year of service are not included. All data are resolved to final bills, with duplicates and correction transactions resolved into a single final bill for the service in question.

CMS Institutional claims and Physician/Supplier and DME claims are available on a quarterly basis and are finalized and closed for a particular service year 18 months following the start of that service year. For example, the CMS 2014 claims are finalized and closed in June 2015. This means that for a claim incurred on January 1, 2014, this system allows for an 18-month window for processing, while claims incurred in December 2014 have a 6-month window for processing to be completed. The inclusion of a claim to a SAF year is determined based on the end date of the service in the claim, not the begin date. For example, a hospitalization with admission on December 26, 2022, and discharged on January 3, 2023, is included in the 2023 claims, not the 2022 claims.

Patients and Time Periods Included

The claims data contains data only for Medicare as Primary Payer (MPP) and Medicare as Secondary Payer (MSP) patients, so claims-based analyses, such as analyses of cost and hospitalization, will not include non-Medicare patients. Non-Medicare patients, therefore, must be identified and excluded when determining numbers of patients or patient years at risk for analyses of cost per patient or hospitalization rates. Also, as it is impossible to determine the complete hospitalization history or complete cost of care for ESRD patients with MSP coverage; such analyses should also exclude patients during the periods when they have this coverage. The Payer History File can be used to make these exclusions (see syntax on creating a Medicare primary cohort in *Example 8: Create a Patient Cohort of 2008 Medicare Primary Incident Patients* as a guide.*)*

The USRDS database also includes pre-ESRD claims for patients who were entitled to Medicare due to age or disability before they developed ESRD. Because these data are not available for all patients, and because it is likely that patients entitled to Medicare before ESRD are systematically different from those not entitled, analyses of these data must be designed with care.

It is important to note that some overlap may exist between the pre-ESRD claims and the ESRD claims. Care must be taken to avoid double counting of claims. Our recommendation is to search the pre-ESRD claims up to, but not including, first ESRD service date, and the ESRD claims starting at the first ESRD service date.

To obtain claims from the ESRD period only, merge the claims file with PATIENTS to identify the first service date (FIRST_SE), and select only those claims occurring on or after this date. It is up to researchers to determine how or whether to include claims that straddle the first service date.

FILE STRUCTURE

ESRD Claims File Structure

The ESRD claims are organized by claim service year. The records in a particular claim service year file have a claim through service date within that claim service year. For example, in PS2004 (Physician/Supplier records incurred in claim service year 2004) all of the records have a claim through date within the calendar year 2004.

ESRD Institutional Claims File Structure

Institutional claims are for hospital inpatient stays, hospital outpatient services, and care provided at dialysis facilities, skilled nursing facilities, home health agencies, and hospices. Institutional claims are submitted via Part A claims forms, which have a large header portion followed by variable length trailers. Possible trailer fields include diagnoses, procedures, and revenue centers. The header portion is in the "Institutional claims" file; the trailer fields are in the "Claims Details" file. Starting with the calendar year 2002, "Claims Details" are separated into a "detailed diagnosis/procedure (_DXP)"file and a "Revenue Center Details (_det)" file; the "_DXP" file contains diagnostic and procedural codes, and the "_det" file contains revenue center details.

Revenue Center Details Added Since the 2008 claims

Effective January 1, 2008, CMS required an important change in billing requirements for ESRD facilities. This change, phased in over the entire 2007 calendar year, had a substantial effect on the content of the Institutional claims files that the USRDS receives from CMS for ESRD patients. The "_det" files accommodated these adjustments for claims in year 2008 and later.

In the past, ESRD facilities were able to bill an entire month of each separately billable services as a summary single line item (e.g. costs for dialysis, erythropoietin (EPO), and IV Iron were all included as one billable charge). The CMS mandated billing changes essentially now require ESRD facilities to report each separately billable service as a separate revenue center line item with individual dates of service. A typical month of dialysis now requires 13 separate line items for each session of dialysis, as opposed to one line item in previous years, 13 separate line items for EPO administrations, and so forth. The practical result of this change was an enormous increase in revenue center detail records, from 76 million in 2006 to over 175 million in 2008.

The increased detail in ESRD billing may be useful for studying patterns of utilization, so the USRDS made these additional revenue line items available to researchers. CMS also altered the method for ESRD facilities reporting of EPO administration and dosage.

Claims details can be linked back to the original claims. Complete listings of the variables are available in *Appendix B: Data File Descriptions.* Some variables are also described in *Section 5: Contents and Variables*. The claims are uniquely identified by a compound key consisting of three variables (USRDS_ID, CLM_FROM, and SEQ_KEYC), which are used to uniquely link claims records among the Institutional claims, Claims Details, and Revenue Details files. The records in all Institutional claims files are sorted and indexed by this compound key.

Data Structure of the claims files before 2002

In 2001 and prior years' files, each of "Claims" (INC in file name) and "Claims Details" (DET in file name) consisted of six

types of claims, and these six types of claims can be distinguished by the HCFASAF variable, as indicated in Table 10. There are some observations in claims with HCFASAF = "M" or HCFASAF = "Q" from PMMIS/REBUS/REMIS. These two categories are not included in Table 9 due to the small number of observations.

Table 9. Claims types in the HCFASAF variable

Type	HCFASAF
Home Health Agency	H
Hospice	S
Inpatient	I
Outpatient	O
Dialysis	D
Skilled Nursing Facility	N

Data Structure of the claims files of 2002 and later

In 2002 Institutional Claims data, three types of files were created for five types of institutional claims services. The three types of files are: "Claims" file (CLM in file name), "Claims Details," or multiple trailers for diagnoses and procedure details (DXP in file name) files, and "Revenue Details," or multiple trailers for revenue center details (DET in file name). The five types of institutional claims services are Outpatient (OP in the file name), Inpatient (IP), Home Health Agency (HH), Hospice (HS), and Skilled Nursing Facility (SN in the file name). Therefore, instead of only two ESRD Claims files, 2002 and subsequent years ESRD Claims SAFs include 15 different files for each year.

Table 10 shows the relationship of institutional claims datasets from 2002 with files 2001 and prior years. In 2001 and prior years' files dialysis claims were identified with HCFASAF = "D". In 2002 and later claims, dialysis claims can be find from outpatient claims files using TYPE_OF_BILL="72" and corresponding revenue center codes. Medicare Claim Clinical data are available for the claim year 2011 and on. Each year contains approximately four million records at the monthly level. It includes monthly hemoglobin level, ESA use, iron use, KT/v, etc. A complete list of variables included in the Medicare Claim Clinical data are in *2025 Researcher's Guide to the USRDS Database Appendix B: Data File Descriptions*.

Table 10. Relationship of institutional claims datasets from 2002 with 2001 and prior years' claims

	Institutional Claims dataset structure (2001 and prior files, yyyy is year)	Institutional Claims dataset structure (2002 and on files, yyyy is year)
Claim level file	in one file incyyyy level file: hcfasaf="H" (home health), hcfasaf="S" (hospice), hcfasaf="I" (inpatient), hcfasaf="O" (outpatient), hcfasaf="N" (skilled nursing facility), hcfasaf="D" (dialysis), hcfasaf="M"(REBUS Inpatient Stay), hcfasaf="P"(Quarterly Dialysis Records)	in separate files below: hh_clm_yyyy (home health), hs_clm_yyyy (hospice), ip_clm_yyyy (inpatient), op_clm_yyyy (outpatient), sn_clm_yyyy (skilled nursing facility), claim_clinical_yyyy

Diagnoses and Procedure Details	in one file det*yyyy:* hcfasaf="H" (home health), hcfasaf="S" (hospice), hcfasaf="I" (inpatient), hcfasaf="O" (outpatient), hcfasaf="N" (skilled nursing facility), hcfasaf="D" (dialysis), hcfasaf="M"(REBUS Inpatient Stay), hcfasaf="P"(Quarterly Dialysis Records)	in separate files below: hh_dxp_yyyy (home health), hs_dxp_yyyy (hospice), ip_dxp_yyyy (inpatient), op_dxp_yyyy (outpatient), sn_dxp_yyyy (skilled nursing facility)
Revenue Center Details	in one file detyyyy: hcfasaf="H" (home health), hcfasaf="S" (hospice), hcfasaf="I" (inpatient), hcfasaf="O" (outpatient), hcfasaf="N" (skilled nursing facility), hcfasaf="D" (dialysis), hcfasaf="M"(REBUS Inpatient Stay), hcfasaf="P"(Quarterly Dialysis Records)	in separate files below: hh_det_yyyy (home health), hs_det_yyyy (hospice), ip_det_yyyy (inpatient), op_det_yyyy (outpatient), sn_det_yyyy (skilled nursing facility)

ESRD Physician/Supplier Claims File Structure

Physician/Supplier claims are bills covering physician, laboratory, and radiology services, as well as medical supplies. They account for approximately 80% of the claims but only 20% of the dollars. One diagnostic and one procedural code can occur on each physician/supplier record, which is a line item of the entire claim. One visit to a physician can generate multiple line-item records.

While there are only minor differences in the structure of the data included in the five institutional claims types (Outpatient, Inpatient, Home Health Agency, Skilled Nursing Facility, Hospice), the structure of the Physician/Supplier claims is substantially different from that of the Institutional Claims.

Compared to Institutional Claims, Physician/Supplier claims have a simpler header portion and fewer trailer fields, including the revenue center with a CMS Healthcare Common Procedure Code Standard (HCPCS) procedural code. Unlike the International Classification of Diseases, Ninth Edition, Clinical Modification (ICD-9- CM) procedural codes on the Institutional Claims, which primarily record invasive surgical procedures, HCPCS codes record all procedures performed by physicians (e.g. patient histories) and all supplies, ranging from Band- Aids® to dialysis machines.

Data Structure Prior to the 2002 claims

Before the 2002 claims, the related claim-level and line-item data files for each type were merged together and then stacked. Multiple trailer records for diagnosis details were also included (Figure 2).

Figure 2. Physician/Supplier claims file structure prior to 2002 claim files

Stacked ↑↑↑←	Physician/Supplier claims	←←←*Merged*	Physician/Supplier line records
	Physician/Supplier multiple trailers for diagnosis details		
	Durable Medical Equipment claims	←←←*Merged*	Durable Medical Equipment line records
	Durable Medical Equipment multiple trailers for diagnosis details		

Data Structure of 2002 and later claims

For 2002 and later Physician/Supplier claims, three types of files were created for the Physician/Supplier claims (PS in file name) and DME claims (DME): claim-level file (CLM in file name), multiple trailers for line-item details of the claim (LINE), and multiple trailers for diagnoses and procedure details of the claim (DX). Therefore, instead of only one ESRD PS claims file, beginning with 2002, ESRD PS claims include six different files. The relationship of 2002 and on files, with PS claims files prior to 2001 PS claims are listed in Table 11.

Table 11. Relationship of 2002 and later datasets with 2001 and prior years' Physician/Supplier Claims datasets

	prior to 2002 PS Claims	2002 and on PS Claims
File structure	one file included claim, line and dx information File name: psyyyy	6 files with claim type and claim, line, dx separated File names: ps_clm_yyyy, ps_line_yyyy, ps_dx_yyyy, dme_clm_yyyy, dme_line_yyyy, dme_dx_yyyy
CLMTYP	No variable can be used to distinguish Physician/Supplier and Durable Medical Equipment claims	Claims were separated to different files based on claim type (file name with "ps" is for physician/supplier, file name with "dme" for durable medical equipment). No claim type indicator is necessary.
CDTYPE	CDTYPE="B" (claim level and line level records) CDTYPE="I" (diagnosis code from dx files)	Claim, line and dx claims are in separate files; no indicator is necessary. However, in the PS file, CDTYPE variable is used to distinguish ICD-9 and ICD-10 diagnosis codes.

Pre-ESRD Claims File Structure

Pre-ESRD Medicare Institutional claims and Physician/Supplier claims are available starting from the 1993 claim service year.

### Obtaining Pre-ESRD Claims for Patients Incident Prior to 2010

Before 2011, the structure of the Pre-ESRD claims files is identical to the ESRD claims files. Each pre- ESRD claim service year file contains claims for multiple incident years. For example, the claims labeled 2004 contain the pre-ESRD claims for incidents patient in 2004, 2005, and 2006. The claims are NOT constrained to an exact two-year period.

Patients who are treated for ESRD for the first time are known as incident ESRD patients. Patients who are incident for ESRD at the end of a calendar year may have claims extending back for nearly three years before their incident date.

To obtain a patient's complete pre-ESRD records for the those incident before 2011, three years of Pre-ESRD claims files should be used. These are the claim service year file that corresponds to the calendar year of incidence for that patient, the claim service year file that corresponds to the incident calendar year minus 1, and the claim service year file that corresponds

to the incident calendar year minus 2. For example, to obtain pre-ESRD Physician/Supplier records for a patient incident in 2006, PS2006, PS2005 and PS2004 should all be used (Figure 3).

Figure 3. Example of Pre-ESRD Physician/Supplier Record for an incident ESRD patient in 2006 (before 2011)
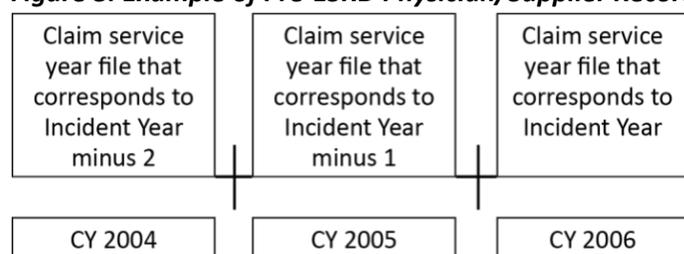
It is important to note that Pre-ESRD claims before 2011 include complete claims information for the calendar year of incidence, which means ESRD claims for the calendar year of incidence are also included in the Pre-ESRD claims. To obtain claims for the pre-ESRD period only, merge the Pre-ESRD claims with the Core PATIENTS file to identify the first service date, and select only those claims occurring before this date. Claim through date (CLM_THRU) is recommended to separate ESRD and pre-ESRD records, to be consistent with Pre-ESRD claims files 2011 and later (see next section, *Obtaining pre-ESRD Claims for Patients Incident in 2011 or Later*).

Obtaining Pre-ESRD Claims for Patients Incident in 2011 or Later

In the 2011 to 2017 Pre-ESRD claims files, the patient cohort is incident patients for an incident year in each file , and each file contains pre-ESRD records from four years of claims files before the incident year and pre-ESRD records in incident year claim files, which is a total of up to five years of pre-ESRD records. In the 2015 SAFs, files for patients incident in 2011, 2012, and 2013 were created. In years following the 2015 SAFs, files for patients incident in SAF year minus two were created. For example, in the 2016 SAFs, files for patients incident in 2014 were created. From 2020 SAFs (2018 incident cohort) on, the pre-ESRD data includes up to two years of data before ESRD. File names ending with "incyyyy" include pre-ESRD records for patients incident in year yyyy; file names starting with "preesrd5y" include up to five years of pre-ESRD records; file names starting with "preesrd2y" include up to two years of pre-ESRD records.

CONTENTS AND VARIABLES

Institutional Claims File

The Institutional claims files have one record per claim, with a claim representing a single instance of service, such as a hospital inpatient stay, an outpatient surgery, or a month of dialysis. Dollar values for total charges and payment amounts are stored in the claims file, which also shows the type and number of dialysis sessions included in the claims. Data in the Institutional claims files allow researchers to determine dialysis treatment modality over time, compute hospitalization rates, and determine aggregate costs by period and type of cost.

These data are sufficient for many research studies. Analyses of particular diagnoses, procedures, or revenue centers require the claims detail files, which are described in the *Institutional Claims Details and Revenue Center Details* section. Complete listings of the variables appearing in the Institutional claims files are available in *2025 Researcher's Guide to the USRDS Database Appendix B: USRDS Data File Descriptions*. Detailed information for some Institutional claims file variables is provided in the next section.

Medicare Payment Variables

CLM_TOT is the total amount billed for the claim, while CLM_AMT is the amount paid by Medicare. For inpatient claims, the cost also includes an amount for the CMS pass-through payments for items such as indirect medical education, capital, and kidney acquisition for transplants. To obtain this pass-through payment amount, multiply the per diem amount (PER_DIEM) by the count of covered days (CVR_DCNT). From the 2015 SAFs and on, third-party payer payment PRPAYAMT, and beneficiary payment (co-pays/deductibles) TDEDAMT are added to the Institutional claims file.

Institutional Claims Details and Revenue Center Details

The Institutional claims Details File includes a variety of types of codes to describe the details about the service claimed for, including:

- ICD CM diagnosis codes
- ICD CM procedure codes
-
- CMS claim related condition codes
- CMS claim related value codes
- CMS claim related occurrence codes

There can be number of Institutional Details records for each Institutional claims . The Claims Detail files are indexed by the same four-part compound key as the claims files so that this key can be used to link the files.

The Revenue Center details correspond to the Revenue Center "trailers" on the CMS SAF records. The Revenue Center details (or record trailers) provide data about the breakdown of the total charges into charges for each (type) of service(s) identified by the HCPCS code and number of services described by variable "units". Each Revenue Center detail contains a variable for the amount charged (REV_CH in the Institutional Claims Revenue Center Details File). Be caution for that, Revenue Center code "0001", is the sum of all of the REV_CH for all other Revenue Center details in that claim. Don't double-count the charges.

To test the consistency of the Revenue Center Details and the CLM_TOT variable, the Coordinating Center examined the original CMS SAF records for 10,000 inpatient and 10,000 outpatient claims. In all cases, the '0001' Revenue Center amount was the sum of the other Revenue Center amounts. In about 3% of the inpatient records, however, CLM_TOT was greater than this sum.

Beginning in late 2000, CMS began providing a field in the revenue trailers called revenue center payment amount, which corresponds to the payment amount for each revenue center trailer for all outpatient claims. This field is included as an additional variable (REVPMT) in the Institutional Revenue Center Details SAF, starting with the calendar year 2001. This value allows researchers to more accurately determine the payment amount for individual types of Revenue Center services (such as laboratory service, EPO, and dialysis) billed on outpatient claims. One caveat for using this variable is that the REVPMT summed overall Revenue Center Details for a given claim will not always agree with the CLM_AMT variable contained in the Institutional Claims SAF for that claim. Our analysis showed that the sum of REVPMT for all outpatient claims exceeded the CLM_AMT for all outpatient claims by approximately 3%. The SAS format $REVCEN gives labels for the Revenue Center codes (the CODE variable on records with CDTYPE = "R").

When a claim has only one dialysis Revenue Center code, as is usually the case for years before 2001, a Revenue Center Details record is not produced because the relevant data items are recorded on the Institutional Claims record. If a claim has multiple Revenue Center Details indicating dialysis, the dialysis variables are derived from the first Revenue Center code

encountered, giving precedence to the more specific codes. In this case, a detailed record is created for each Revenue Center detail on the claim, so users have the opportunity to interpret the multiple details. Other Institutional Revenue Center details are of lesser interest unless an HCPCS code is included indicating a more specific service. A code showing that a claim is for laboratory services, for example, frequently includes an HCPCS code indicating the specific test performed.

Revenue Center Details records are included regardless of the presence or absence of an HCPCS code. Before the calendar year 2008, Revenue Center Details records were included with other Claim Details records, and after 2008, were placed in a separate file, Revenue Center Details File. For many analyses, the Revenue Center Details records may not be required. Complete listings of the variables appearing in the Revenue Center Details Files are available in the 2025 Researcher's Guide to the USRDS Database Appendix B: Data File Descriptions. The detailed information for some Revenue Center Details File variables is described in the next section.

CDTYPE and CODE Variables

The variable CDTYPE indicates the type of code contained in the CODE variable (e.g. a diagnosis code, a procedure code, a revenue center code). For instance, a code of "I" in the CDTYPE variable indicates that the value of the CODE variable is ICD-9 diagnosis code; a code of "P" in the CDTYPE variable indicates that the value of the CODE variable is an ICD-9 procedure code. Both variables are present on every record, but the variables UNITS, REV_CH, HCPCS, and URR_CD are not present for some CDTYPEs. The SAS format $CDTYPEI indicates the meaning of each CDTYPE.

UNITS

Use of the UNITS variable varies with CDTYPE. When CDTYPE = "P" (ICD-9-CM Surgical Procedures) or CDTYPE = "Q" (ICD-10-CM Surgical Procedures), UNITS is a value created by the USRDS to indicate when the surgical procedure was performed, and time is expressed as the number of days from the date given by CLM_FROM, with CLM_FROM counted as 1. A value of 1 for UNITS indicates that the procedure was performed on the date given by CLM_FROM, and 2 indicates the day after CLM_FROM. When CDTYPE = "R", UNITS is described in the CMS file documentation as a quantitative measure (unit) of services provided to a beneficiary associated with accommodation and ancillary revenue centers". From 2008, CDTYPE = "R" records are all in Revenue Center Details File. Depending on the type of service, units are measured by the number of covered days in a particular accommodation, emergency room visits, clinic visits, dialysis treatments (sessions or days), outpatient therapy visits, and outpatient clinical diagnostic laboratory tests. The revenue center code or the HCPCS code indicates the type of service. Because the meaning of UNITS varies greatly, the variable must be used with caution. When using this variable, tabulate the distribution of values over the records being analyzed to ensure that the values look correct. When CDTYPE = "I" (ICD-9-CM Diagnosis) or CDTYPE = "D" (ICD-10-CM Diagnosis), UNITS has a value of 1 to 25 which is the order of diagnosis codes on the form. Value 1 indicates that this was the primary diagnosis for this claim. The claim details are not necessarily sorted with the primary diagnosis first.

REV_CH

REV_CH occurs only on Revenue Center details records (CDTYPE = "R") and indicates "the total charges (covered and non-covered) for all accommodations and services (related to the revenue code) for a billing period before reduction for the deductible and co-insurance amounts and before an adjustment for the cost of services provided." REV_CH corresponds in concept to the CLM_TOT variable on the Institutional claims file, as discussed above under Revenue Center Details.

HCPCS and MOD1-MOD5

The CMS Common Procedure Coding Standard (HCPCS) "is a collection of codes that represent procedures, supplies, products, and services that may be provided to Medicare beneficiaries and to individuals enrolled in private health insurance programs." In the Institutional claims files, the HCPCS code occurs only on Revenue Center (CDTYPE = "R") details,

but may not be present on all such records. HCPCS are an extension of the American Medical Association CPT-4 codes, and CPT-4 codes can only be found in HCPCS in Revenue Center Details Files. Codes for certain pharmaceuticals, laboratory procedures, durable medical equipment, and radiology procedures are added to the CPT-4 codes to form HCPCS. Beginning in the calendar year 2010, the HCPCS code may be accompanied by up to five HCPCS modifier codes (MOD1-MOD5). These modifiers are used to report information related to the services described by the HCPCS code. For example, beginning July 1, 2010, to report absence or presence of access infection on the last dialysis line item of each dialysis bill, CMS requires the use of the modifiers V5 (catheter), V6 (graft), and V7 (fistula) to report the type of vascular access in use, and V8 (infection), and V9 (no infection). Another example would be modifiers for IV (JA) and subcutaneous (JB) administration of ESA.

URR_CD/URR_CAT

Starting in 1998, CMS began requiring the reporting of Urea Reduction Ratios (URRs) on outpatient hemodialysis claims. The URR is reported as a range that reflects the results of the month being billed. This information appeared as a formatted value in the Revenue Center Details for hemodialysis claims in 2012 and prior years' files. In 2013 to 2017 files, URR_CD is in the "Claims" file. From 2018 and later files, this information only in the "CLAIM_CLINICAL_yyyy" file, variable name is "URR_CAT".

Physician/Supplier Claims

In the Physician/Supplier claims files, a claim does not necessarily correspond to a logical instance of service but is more likely to represent all services provided to a patient during the provider's billing period. In 2012 and prior years' files, the file is constructed as a line item file, with one record per claim line item. For physician claims, the from/through dates can be used to identify a series of line items associated with a single visit. For supplier claims, however, the instance of service is more difficult to define. Bills for home dialysis dialysate, for example, specify the quantity and delivery date of the dialysate but not the period over which it is to be dispensed. Claims and line records are in separate files in the 2002 and subsequent year files.

CDTYPE and DIAG

Prior to 2002 files, claims and line records were merged and then detailed diagnosis codes were stacked to claims/ line merged records. Claims and line records have a one-to-many relationship, meaning there is one claims record to one or more line records. CDTYPE = "B" indicates a physician/supplier claims/line merged item and the DIAG variable indicates the diagnosis code associated with this line item. CDTYPE = "I" indicates that this record contains only one diagnosis code; in this case, the diagnosis code is associated with all line items on the claim. Payment variables for these records should be missing.

In 2002 and subsequent year files, claims, line, and detailed diagnosis code records are in separate files, and the DIAG variable is kept in all files.

HCPCS and MOD1-MOD2

HCPCS in Physician/Supplier Claims contain procedure codes, as described in the *HCPCS and MOD1-MOD5* section *in Institutional Claims Details and Revenue Center Details.* Variables MOD1-MOD2 are included to identify further the type of service billed on the line item. They are used in conjunction with the HCPCS/CPT code on the line item, and their meaning can be found in the Current Procedural Terminology codebook and the HCPCS Level II codebook.

SPCLTY

The Physician/Supplier specialty code (SPCLTY) can be useful for untangling the bills for a specific surgical procedure. The principal surgeon, physician surgical assistants, and anesthesiologist use the HCPCS referring to the major invasive surgery to bill for that surgery. The code for nephrologists is "39".

PLCSRV

The place of service variable (PLCSRV) indicates where the service was rendered. It can be used to distinguish between inpatient and outpatient services and between home dialysis and in-unit dialysis supplies. The value "6" refers to an ESRD treatment center.

HCSRVC

The CMS service code variable (HCSRVC) can be used to distinguish between the principal surgeon and assistants. The value for immunosuppressive drugs is G, for the renal supplier in the home L, for monthly capitation payment (dialysis) M, and for kidney donor N.

Payment Variables

The following three payment fields appear in each Physician/Supplier line item in all years' Physician/ Supplier claims: submitted charges (SBMTCH), allowed charges (ALOWCH, the lower of prevailing, customary, or actual as determined by CMS); and the payment amount (PMTAMT, the amount paid to the provider and/or beneficiary after deductible and co-insurance amounts have been paid for the services included as a line item on a physician/supplier claim). A complete listing of the variables in the Physician/Supplier Claims Files is available in the 2025 Researcher's Guide to the USRDS Database Appendix B: Data File Descriptions.

After the implementation of the bundled payment system in 2011, approximately 15% fewer physician/ supplier claims were submitted to CMS. We believe this is because the implementation of bundled care changed the way in which health care providers applied for Medicare reimbursement for ESRD patient care. Specifically, a given patient's dialysis facility now submits claims for any bundled care, even if it is performed by an outside physician. In turn, the facility reimburses the provider for any bundled care provided to the ESRD patient.

Therefore, when looking at temporal trends in payment data, for example, researchers will see a decline in payments after January 2011. Also, when looking at trends that cross the bundle implementation date, it is necessary to restrict payment analyses to care that falls outside of the bundle. Researchers may use the list of bundle-covered medication and laboratory procedures at the Medicare website: *http://www.cms.gov/Medicare/Medicare-Fee-for-Service-Payment/ESRDpayment/Consolidated_Billing.html*.

Occurrence of HCPCS

HCPCS codes exist in the following claim files:

- In institutional claims, the HCPCS code occurs only on Revenue Center (CDTYPE = "R") details, which are in Institutional DET claim files (REV claims prior to 2002 ESRD and 2010 pre-ESRD), but may not be present on all such records. See *HCPCS and MOD1-MOD5* in *Institutional Claims Details and Revenue Center Details* for more details.
- In Physician/Supplier and DME claims, HCPCS code occurs in DME Line files, and PS Line files See *HCPCS and MOD1-MOD2* in *Physician/Supplier Claims* for more information.
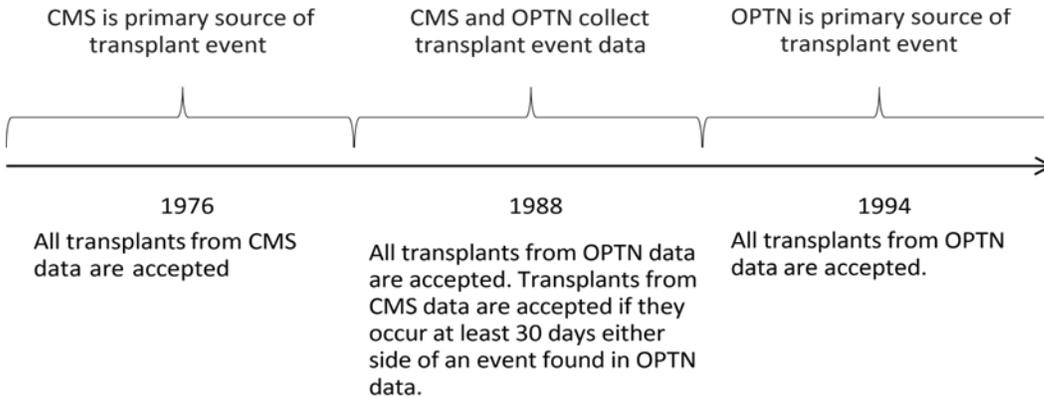
SECTION 6: TRANSPLANT PROCESS AND OUTCOMES

Transplant patients constitute a unique subset of the ESRD population and are often studied separately from dialysis patients. Researchers may wish to simply count the number of transplant events that meet certain criteria or calculate transplant event rates and survival probabilities. Using the USRDS transplant data, researchers can obtain information on transplant events, such as donor and recipient characteristics, and on patient- and graft- related outcomes.

Transplant Data Sources

Basic transplant variables are contained in the TRANSPLANT file in the Core SAF. Transplant event data are combined from CMS and OPTN databases. Before 1988, CMS was the primary source for all transplant event data. Between 1988 and 1993, both CMS and OPTN collected information regarding transplant events. Since 1994, OPTN has been the primary source for all transplant event data (Figure 4).

Figure 4. Source of transplant events in the USRDS population

CMS is primary source of transplant event | CMS and OPTN collect transplant event data | OPTN is primary source of transplant event

1976
All transplants from CMS data are accepted

1988
All transplants from OPTN data are accepted. Transplants from CMS data are accepted if they occur at least 30 days either side of an event found in OPTN data.

1994
All transplants from OPTN data are accepted.

Reconciliation of CMS and OPTN Events

The USRDS has implemented a decision algorithm, described in detail in *Section 2: ESRD Patients*, which reconciles identical transplant events when multiple sources contain conflicting information. Each transplant event found in the Transplant file (TX) of the Core SAF dataset is thus a unique event that the USRDS believes occurred after reviewing both OPTN and CMS REMIS Transplant and Medical Evidence Report files. Table 12 details the various USRDS SAFs related to transplant. The Transplant file (TX) in the Core SAF contains transplant dates, basic descriptive data, and causes of graft failure when known.

Table 12. USRDS Transplant files

File	Data Sources	Comments	Period	Dataset
TX	Constructed by USRDS from OPTN and EQRS files.	Identifying all transplant	Simce 2/26/1960	CORE
WAITLIST_KI	OPTN transplant waiting list file for kidney transplant	Data on kidney waiting list, per listing event	Listing date 3/1/1970 or later	CORE
WAITLIST_KP	OPTN transplant waiting list file for kidney and pancreas transplant	Data on kidney-pancrease waiting list, per listing event	Listing date 9/2/1986 or later	CORE
WAITSEQ_KI	OPTN transplant waiting list sequence file for kidney transplant	Condensed, center-specific, kidney transplant waiting list	Listing date 3/1/1970 or later	CORE
WAITSEQ_KP	OPTN transplant waiting list sequence file for kidney and pancreas transplant	Condensed, center-specific, kidney-pancreas transplant waiting list	Listing date 9/2/1986 or later	CORE
TXUNOS_TRR_KI	OPTN Kidney Transplant Recipient Registration file	Kidney transplant details, event-level	Transplant 02/29/1960 or later	TX
TXUNOS_TRR_KP	OPTN Kidney and Pancreas Transplant Recipient Registration file	Kidney-pancreas transplant details, event-level	Transplant 09/29/1969 or later	TX
TXUNOS_TCR_KI	OPTN Kidney Transplant Candidate Registration file	Kidney recipient data, as condidate	listing date 03/01/1973 or later	TX
TXUNOS_TCR_KP	OPTN Kidney and Pancreas Transplant Candidate Registration file	Kidney-pancreas recipient data, as condidate	listing date 07/11/1986 or later	TX
TXUNOS_CDR	OPTN Deceased Donor Registration file	Data of deceased donor for kidney and kidney-pancreas transplant with recipient	Transplant date 03/30/1964 or later	TX
TXUNOS_LDR	OPTN living Donor Registration file	Data of living donor for kidney and kidney-pancreas transplant with recipient	Transplant date 02/29/1960 or later	TX
TXUNOS_DHS	OPTN Donor Histocompatibility file	Donors' histocompatibility data for kidney and kidney-pancreas transplant	Transplant date 02/29/1960 or later	TX
TXUNOS_RHS	OPTN Recipient Histocompatibility file	Recipients' histocompatibility data for kidney and kidney-pancreas transplant	Transplant date 02/29/1960 or later	TX
TXFUUNOS_KI	OPTN Kidney Transplant Follow-up file	Follow-up details for kidney transplants	Transplant date 02/29/1960 or later	TX
TXFUUNOS_KP	OPTN Kidney and pancreas Transplant Follow-up file	Follow-up details for kidney-pancreas transplants	Transplant date 09/29/1969 or later	TX
TXIRUNOS	OPTN Immunosuppression at Transplant file	Immunosuppression details for kidney and kidney-pancreas recipients at transplants		TX
TXIFUNOS	OPTN Immunosuppression at Follow-up file	Immunosuppression details for kidney and kidney-pancreas recipients at follow-up		TX

Kidney and Pancreas Waiting List

The Core SAF also contains four files with information on the kidney and simultaneous kidney-pancreas waiting list files from OPTN. Most investigators are simply interested in the dates of listing at specific transplant centers. The WAITSEQ_KI and WAITSEQ_KP Files contain entry and removal date sequences per patient per transplant center for the kidney alone and simultaneous kidney-pancreas waiting lists, respectively. These records are collapsed from the raw OPTN Waiting List Files WAITLIST_KI and WAITLIST_KP. Both the raw OPTN Waiting List Files and the Sequence Files contain two patient identifiers: PID and USRDS_ID. ESRD patients included in the USRDS patient profile have a USRDS_ID.

Because the Sequence Files are collapsed versions of the Waiting List Files, investigators should use caution when linking back to the raw waiting list files to obtain information such as Panel Reactive Antibody (PRA) values. Investigators will need to obtain all raw waiting list records and then make decisions as to which record contains the information most relevant to their needs. Some specific assumptions the USRDS Coordinating Center used to collapse waiting list records are as follows:

- Candidates listed in error were removed (REMCODE = 10).
- Patients with missing listing dates (EDATE = .) were removed.
- Patients with listing and removal dates on the same date were removed.
- Transplant dates were taken from the USRDS Transplant File and not from the OPTN Waiting List File.
- The transplant dates on the USRDS Transplant File are cleaned and reconciled with other sources of data, so these dates were used to ensure consistency.
- Known transplant dates are used to truncate waiting periods at all centers at which a patient is currently listed. For example, if a patient underwent transplant but was not removed from a center's waiting list, a removal date would be imputed on the transplant date. This is true for all centers at which a patient is listed.
- Patients with inactive waiting periods were removed; the waiting list sequence lists only active periods.
- Overlapping waiting periods at the same center were collapsed.

Transplant SAF

To obtain additional data regarding transplant events, users need the Transplant SAF, which contains detailed information, in separate files, from CMS and OPTN (UNOS). The TRANSPLANT file in the Core SAF contains reconciled transplant event data from all available sources. As some researchers may find it beneficial to see the transplant data obtained from each source, data from each source are included.

Two Transplant Files contain data collected by CMS and OPTN (UNOS) on transplant follow-up visits. The follow-up records in TXFUHCFA and TXFUUNOS overlap in time, specifically during 1988 to 1993, and contain information collected during patient follow-up visits, which typically occur at six months, one year, and yearly thereafter. The TXIRUNOS and TXIFUNOS Files contain information on immunosuppression treatment. Data in these files are from the Kidney Transplant Recipient Registration Worksheet and Kidney Transplant Recipient Follow-Up Worksheet. TXIRUNOS contains data on treatments at the time of transplant, and TXIFUNOS contains data on treatments updated at each follow-up visit, if available.

Transplant candidate, recipient, donor, and histocompatibility information, previously combined in TXUNOS_KI_PRE_JUL04, TXUNOS_KI_POST_JUL04, and TXUNOS_KP, are separated from each other, and now saved in 8 different datasets. TXUNOS_TRR_KI include kidney recipient transplant details, and the variable list is consistent with most recent OPTN Adult/Pediatric Kidney Transplant Registration Worksheet; TXUNOS_TRR_KP include kidney-pancreas recipient transplant details, and the variable list is consistent with most recent OPTN Adult/Pediatric Kidney Pancreas Worksheet.

Transplant Registration Worksheet. TXUNOS_TCR_KI includes kidney recipient data during registration as a candidate; the variable list is consistent with the most recent OPTN Adult/Pediatric Kidney Transplant Candidate Registration Worksheet. TXUNOS_TCR_KP includes kidney-pancreas recipient data while they were registered as a candidate, and the variable list is consistent with the most recent OPTN Adult/Pediatric Kidney- pancreas Transplant Candidate Registration Worksheet. Candidate files only include patients who received a kidney/kidney-pancreas transplant, those registered as a candidate but have not received a transplant yet are not included in TXUNOS_TCR_KI or TXUNOS_TCR_KP.

TXUNOS_CDR includes deceased donor data, and the variable list is consistent with the most recent OPTN Deceased Donor Registration Worksheet; TXUNOS_LDR includes living donor data, and the variable list is consistent with the most recent OPTN Living Donor Registration Worksheet. Only donors who donated one or more organs to kidney or kidney-pancreas transplant recipients are included in TXUNOS_CDR or TXUNOS_LDR. TXUNOS_DHS includes donor histocompatibility data for either deceased or living donors who donated to kidney or kidney-pancreas transplant, and the variable list is consistent with OPTN Donor Histocompatibility Worksheet. TXUNOS_RHS includes recipient histocompatibility data for both kidney and kidney-pancreas transplant, and the variable list is consistent with OPTN Recipient Histocompatibility Worksheet. Historical variables that are no longer collected by the most recent OPTN forms are dropped from the Transplant SAF.

Starting in 2003, OPTN began to use the bitmask technique to maintain multiple values (i.e. selections/choices) within one variable. This method applies to all multiple selection questions. Previously, OPTN used as many variables as needed to collect all possible answers to such questions. For example, for the question "Treatment: Other therapies," OPTN used three variables, PHOTOPH, PLASMA, and LYMPHOID with answers "Yes/No." After 2003, OPTN combined these variables into a single variable, THERAPIES, for the same question, using the bitmask technique. The basic code values are as follows: "1" = Photopheresis; "2" = Plasmapheresis; "4" = Total Lymphoid Irradiation (TLI), while "3" represents the sum of values 1 and 2; and "5" indicates that both 1 and 4 are selected. To assist researchers in decrypting these multiple-selection value variables, USRDS has identified all possible additive combinations of these bitmask-value variables and made them available as SAS formats.

Most descriptive data on transplant events can be found in the Core SAF TRANSPLANT file, which can be used to construct counts of various transplant-related events. Transplant rates can be constructed by combining transplant data with a patient treatment history file, RXHIST or RXHIST60; rates in the ADR are typically calculated using RXHIST60, the treatment history file with the 60-day rule built in. Transplant survival, both of graft and patient, can be calculated using the transplant failure date found in the Transplant File and the date of death found in the PATIENTS File.

The Transplant SAF file structure changed starting with the 2017 ADR. Transplant candidate, recipient, donor, and histocompatibility information are now separately stored in eight different datasets. This clarified file structure makes it easier for users to understand the data sources of each component. The new structure also reduces file size for ease of delivery and use. Transplant follow-up and immunosuppressive information structure remain the same as in previous years. The changes are described in more detail below.

Table 13. Transplant SAF structure change

Transplant SAF	Figure 5 illustrates the structure of USRDS 2016 TX SAF. Transplant candidate, recipient, donor, and histocompatibility information were contained in TXUNOS_KI_PRE_JUL04/TXUNOS_KI_POST_JUL04 for kidney-alone transplants, and in TXUNOS_KP for kidney-pancreas transplants.
	Transplant follow-up information was found in TXFUUNOS_KI for kidney-alone transplants and TXFUUNOS_KP for kidney- pancreas transplants.
	Immunosuppressive information at the time of transplant was found in TXIRUNOS and immunosuppressive information at follow-up in TXIFUNOS.
	Figure 6 shows the **revised structure** of the USRDS 2017 TX SAF. Transplant candidate, recipient, donor, and histocompatibility information are now separately stored in eight different datasets.
	Transplant follow-up and immunosuppressive information structure remain the same as in previous years.
	Figure 7 shows the relationship between the previous 2016 TXUNOS_KI_PRE_JUL04/TXUNOS_KI_POST_JUL04 and the new six files of the 2017 USRDS TX SAFs.
	Figure 7 also shows the 2016 TXUNOS_KP mapped to the two files in the 2017 USRDS TX SAFs

Table 14. Relationships and linkages between the TX and TXUNOS SAFs, Kidney-Alone, 2016-2017

2016 USRDS TX SAF (previous)	2017 USRDS TX SAF (current)	OPTN form	Variable label prefix	Key ID
TXUNOS_KI_PRE_JUL04 TXUNOS_KI_POST_JUL04	TXUNOS_TRR_KI	Kidney Transplant Registration Worksheet	KIR	USRDS_ID TRR_ID DONID
	TXUNOS_TCR_KI	Kidney Transplant Candidate Registration Worksheet	TCR_KI	TRR_ID
	TXUNOS_CDR	Deceased Donor Registration Worksheet	CDR	DONID
	TXUNOS_LDR	Living Donor Registration Worksheet	LDR	DONID
	TXUNOS_DHS	Donor Histocompatibility Worksheet	DHS	DONID
	TXUNOS_RHS	Recipient Histocompatibility Worksheet	RHS	USRDS_ID TRR_ID DONID
	TXUNOS_TRR_KI	Donor Histocompatibility Worksheet vs Recipient Histocompatibility Worksheet	USRDS	USRDS_ID TRR_ID DONID

Table 15. Relationships and linkages between the TX and TXUNOS SAFs, Kidney-Pancreas, 2016-2017

2016 USRDS dataset (previous)	2017 USRDS dataset (current)	OPTN form	Variable label prefix	Key ID
TXUNOS_KP	TXUNOS_TRR_KP	Kidney-Pancreas Transplant Registration Worksheet	KPR	USRDS_ID TRR_ID DONID
	TXUNOS_TCR_KP	Kidney-Pancreas Transplant Candidate Registration Worksheet	TCR_KP	TRR_ID
	TXUNOS_CDR	Deceased Donor Registration Worksheet	CDR	DONID
	TXUNOS_LDR	Living Donor Registration Worksheet	LDR	DONID
	TXUNOS_DHS	Donor Histocompatibility Worksheet	DHS	DONID
	TXUNOS_RHS	Recipient Histocompatibility Worksheet	RHS	USRDS_ID TRR_ID DONID
	TXUNOS_TRR_KP	Donor Histocompatibility Worksheet vs Recipient Histocompatibility Worksheet	USRDS	USRDS_ID TRR_ID DONID

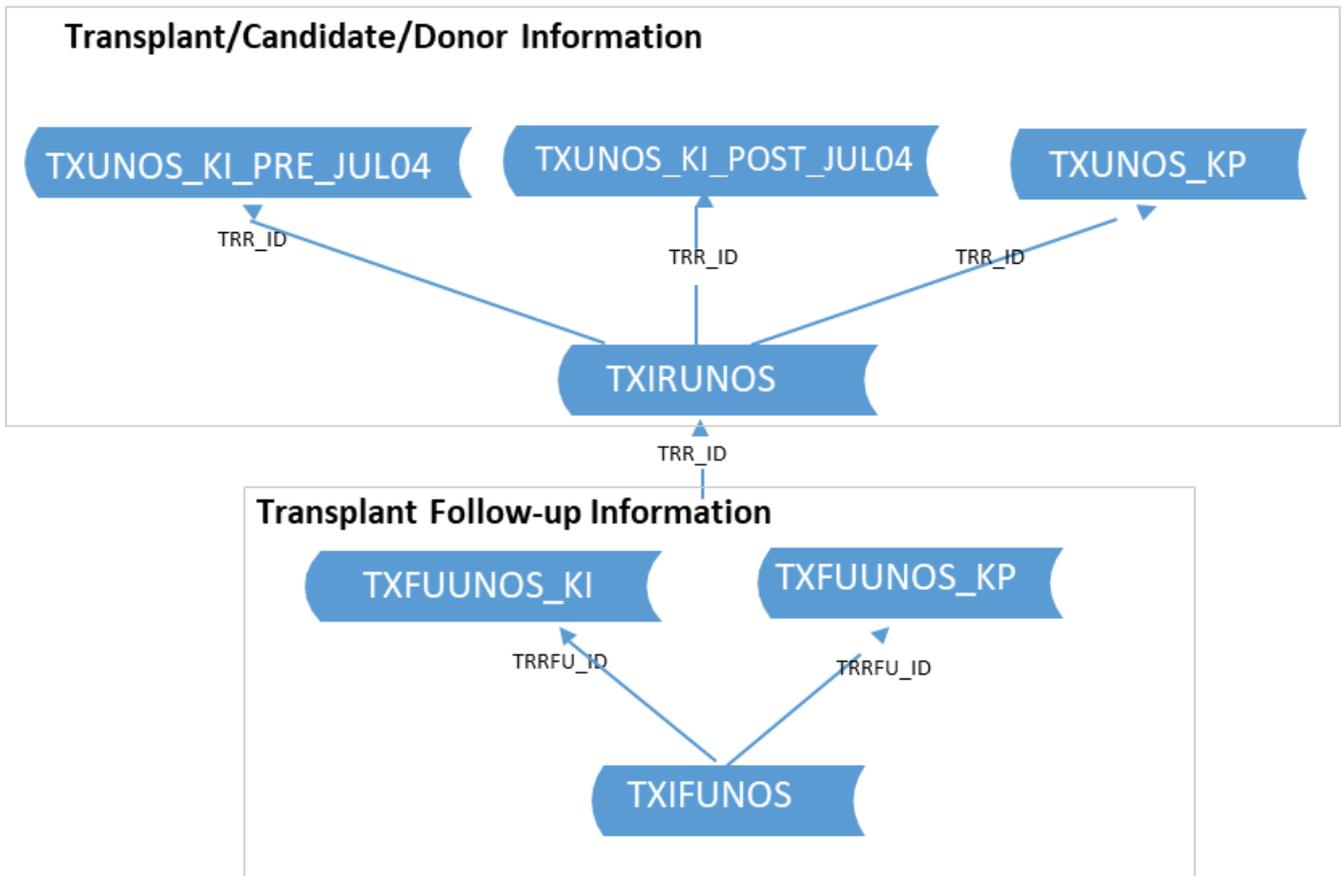**Figure 5. USRDS Transplant Standard Analysis File, 2016**

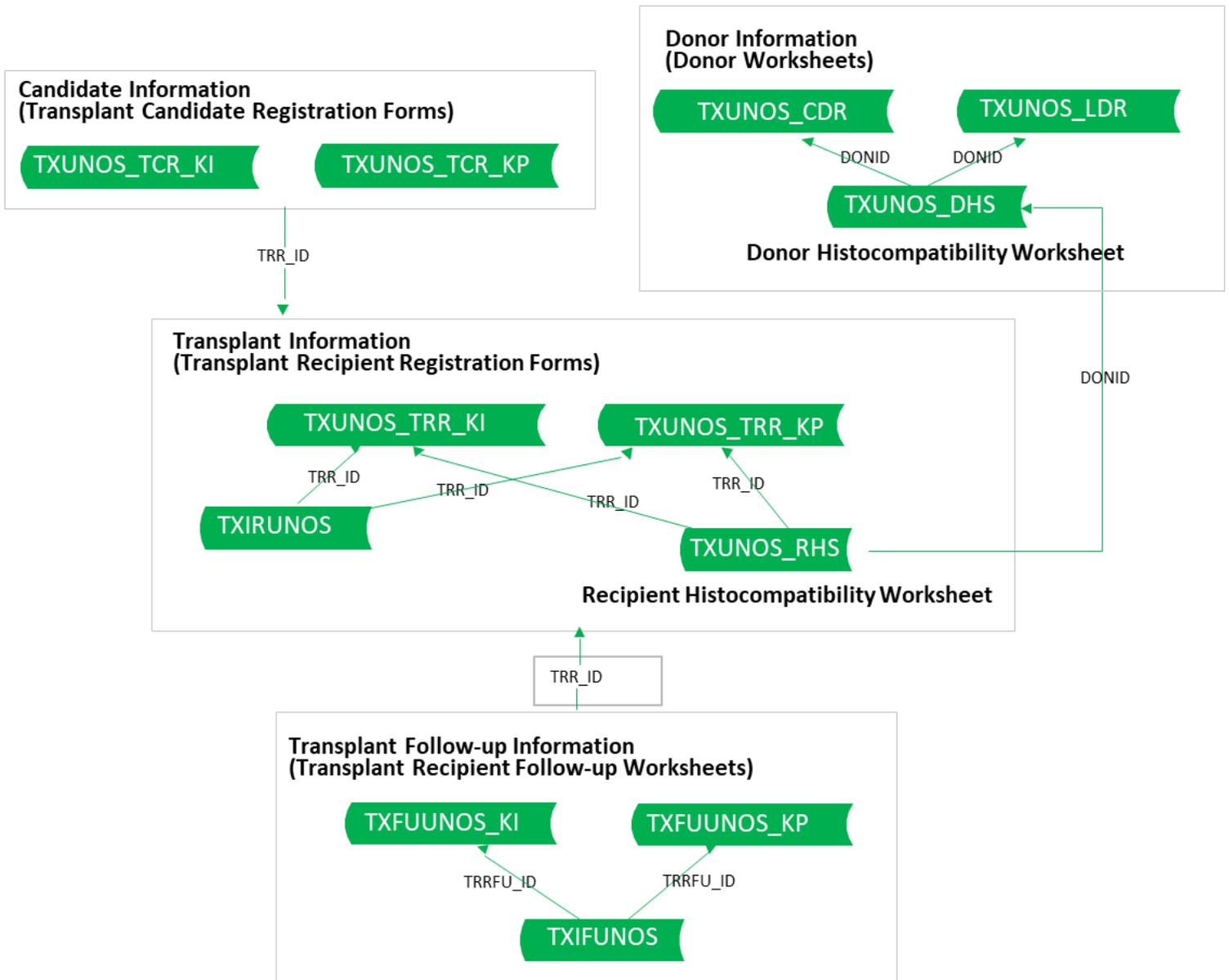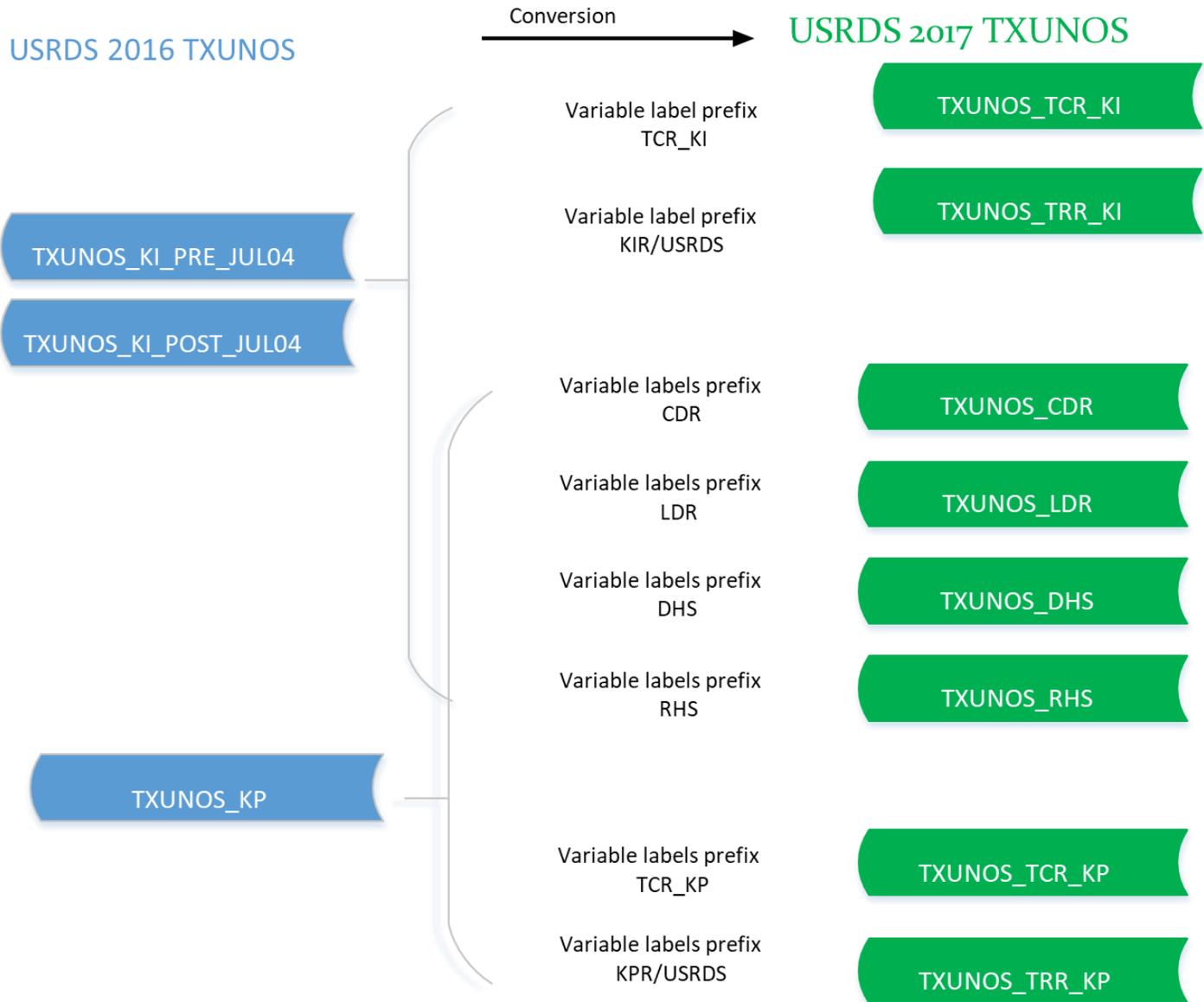**Figure 6. USRDS Transplant Standard Analysis File, Revised 2017**

Figure 7. Comparison of USRDS Transplant Standard Analysis File Structure, Pre and Post 2017 Revision

USRDS 2016 TXUNOS → Conversion → **USRDS 2017 TXUNOS**

- TXUNOS_KI_PRE_JUL04
- TXUNOS_KI_POST_JUL04
- TXUNOS_KP

2016 Variable Label	2017 File
Variable label prefix TCR_KI	TXUNOS_TCR_KI
Variable label prefix KIR/USRDS	TXUNOS_TRR_KI
Variable labels prefix CDR	TXUNOS_CDR
Variable labels prefix LDR	TXUNOS_LDR
Variable labels prefix DHS	TXUNOS_DHS
Variable labels prefix RHS	TXUNOS_RHS
Variable labels prefix TCR_KP	TXUNOS_TCR_KP
Variable labels prefix KPR/USRDS	TXUNOS_TRR_KP

SECTION 7: MORBIDITY AND HOSPITALIZATION

Morbidity associated with ESRD can be determined from information on hospitalizations and acute events, documented in Medicare claims files through ICD-9/10-CM (International Classification of Diseases, Ninth/Tenth Revision) and CPT (Current Procedural Terminology) codes. The USRDS provides data on inpatient admissions, discharges, and diagnosis and procedure codes, which may be used to produce different types of hospitalization rates, such as total admission rates and hospital day rates.

Hospitalization Data Sources

Inpatient hospitalization data, a subset of the data in the Institutional claims files, are included in the Hospitalization SAFs. These data originate from the Part A institutional inpatient claims data.

Contents of the Hospital Dataset

A complete list of variables in the Hospital SAF can be found in *2025 Researcher's Guide to the USRDS Database Appendix B: USRDS Data File Descriptions*. The following list outlines some of the key variables:

- CLM_FROM provides the from date of service, indicating the admission date.

- CLM_THRU provides the through the date of service, indicating the date of discharge.

- HSDIAG1-HSDIAG26 provide up to 10 ICD-9-CM diagnosis codes in HOSP_to2009; and up to26 ICD-9/10-CM diagnosis codes in HOSP_2010on.

- HSSURG1-HSSURG25 provide up to 10 ICD-9-CM procedure codes in HOSP_to2009; and up to 25 ICD-9/10- CM procedure codes in HOSP_2010on.

- DRG_CD provides diagnostic-related group (DRG) codes, which categorize inpatient stays by ICD-9-CM diagnosis and procedure codes.

- PRIMDIAG: equals "yes" when HSDIAG1 provides the primary diagnosis code in HOSP_to2009 and equals primary diagnosis code in HOSP_2010on

Preparing Inpatient Claims Data for Use

Before analyses, data requires cleaning to combine overlapping hospitalizations for the same patient. For the USRDS ADR, the principal diagnosis and procedure codes from the first of two overlapping hospitalizations are retained, with the combined hospitalization extending from the first admission date to the last discharge date.

To create a dataset to calculate rates for period prevalent patient cohorts, attach data for hospitalizations occurring during the selected year by patient USRDS_ID to a period prevalent cohort file for the year. In this way, all patients in the file, including those with multiple hospitalizations and those with none, are included in the rate calculation.

Because hospitalization data are incomplete for non-Medicare patients and patients classified as MSP, these patients should be excluded, thereby limiting their contribution to patient years at risk. The Payer History File in the Core SAF can be used to do this, see SAS *Example 8: Create a Patient Cohort of 2008 Medicare Primary Incident Patients.* Another method of MSP patient exclusion involves excluding dialysis patients who do not reach a certain level of Medicare paid dialysis bills. For instance, dialysis patient start dates can be required to fall between start and end dates based on Medicare paid dialysis claims as follows:

- Claims start date: the first day of the first month in which there are at least $675 (before year 2011), or at least $1,200 (in or after year 2011), of Medicare paid dialysis claims.
- Claims end date: the end of a three-month period in which there are less than $675 (before year 2011), or at least $1,200 (in or after year 2011) of paid claims in each month.

If a patient's start date does not fall between the claims start and end dates, that patient is excluded from the analysis for that year. To calculate rates by the patient characteristics, demographic data may be attached to the dataset by patient USRDS_ID from the PATIENTS File.

SECTION 8: SURVIVAL AND MORTALITY

Survival and mortality analyses are often used to compare outcomes among treatment modalities, age groups, or races, or to evaluate outcome trends over time. Primary analyses include unadjusted survival probabilities using the

Kaplan-Meier method, adjusted survival probabilities using the Cox regression model, unadjusted death rates, and adjusted death rates.

Mortality Data Sources

Survival and mortality analyses often require patient demographic information such as age, sex, race, the primary cause of ESRD, death date, the cause of death, and treatment modality history data, which can be obtained from the USRDS Core SAF Files PATIENTS and RXHIST or RXHIST60.

SECTION 9: PROVIDERS

The Facility SAF File is constructed from data supplied by the CMS Annual Facility Survey. Data in the Facility SAF are at the facility level, and there are no patient-level data. Over time, facilities may be purchased and sold. Some of these facilities might need to change their provider numbers, such as those who changed their facility type. The physical facility and location may be the same, and most staff may remain, but the provider number has changed.

The Facility SAF File contains over 100 provider characteristics, including:

- Transplant counts
- Self-dialysis information
- HD, PD, CAPD, and CCPD patient counts
- Number of in-center new dialysis patients at a facility
- Profit status, and hospital-based or freestanding status
- Chain affiliation
- Reuse practices

Provider Data Sources

The CMS Annual Facility Survey is obtained as part of the EQRS data that the USRDS receives.

Provider Numbers Assigned by the USRDS

Each provider has been assigned an identification number by CMS (CCN). However, in an effort to ensure confidentiality and conceal provider identity, the USRDS assigned an anonymous number to each provider; these are not the original provider numbers assigned by CMS. Each provider maintains its USRDS-assigned facility ID throughout the SAFs to allow linkage.

Researchers who wish to conduct patient-level studies incorporating the Core FACILITY filedata can link patient information from the Detailed Treatment History file (RXHIST), the Condensed Treatment History file (RXHIST60), or other files, to the Core FACILITY file via the USRDS provider number and year. However, investigators are cautioned that analysis of small samples does pose a risk of inadvertent patient identification, thus reporting is limited to patient groups (cells) of 11 or greater, per CMS policy.

Center Type

The center_type variable tells a center/facility be a dialysis center, transplant center, or a center for both dialysis and kidney transplant. A center for both dialysis and kidney transplant may have a unique CCN. But they might submit yearly survey separately, one for dialysis and one for kidney transplant with different organization ID (ORG_ID).

Profit Status

The profit variable, NU_P_NP, has the following values, "For-profit", "Non-profit", and "Unknown" (case sensitive). The For-profit group consists of facilities categorized on the CMS survey as individual-profit, partnership-profit, corporation-profit, or other-profit. Non-profit facilities include individual-non-profit, partnership-nonprofit, corporate non-profit, other-nonprofit, state-government non-federal, county-government non-federal, city government nonfederal, city/county-government non-federal, hospital district/authority government non-federal, other-government non-federal, Veterans Administration Government Federal, Public Health Service Government Federal, military government federal, or other-government federal. Those 18 categories are designated by the variable TYPOWNER. However, the 2002 CMS survey dropped this variable, and it is missing in subsequent survey periods. Profit status is now determined solely by the CMS Dialysis Facility Compare data.

Chain Affiliation

The chain variable, CHAIN_ID, contains a text string identifying the chain ownership of the facility. A blank value for CHAIN_ID indicates that the facility has no chain affiliation. The variable can be used to classify facilities based on the number of facilities operated by the dialysis organizations represented by CHAIN_ID.

Freestanding Vs. Hospital-Based Facility

A hospital-based dialysis facility is associated with a hospital; it can be located in the hospital or at a remote location. Freestanding dialysis facilities operate independently of a hospital. A study analyzing freestanding vs. hospital-based units will generally include three groups: freestanding, hospital-based, and unknown. The variable NU_HBFS designates this status. A code of "1" represents a hospital-based unit, and "2" represents freestanding. The unknown category will consist of any provider in the researcher's study not found in the Facility database.

SECTION 10: CKD COHORT (5% CKD COHORT DATASET)

Traditionally, the USRDS has focused on the final stages of kidney disease, due to the comprehensive data available from Medicare on all patients with ESRD. However, there is no similar national reporting system for patients with kidney disease prior to ESRD. Since 2008, the USRDS has been reporting on chronic kidney disease in Volume 1 of the USRDS ADR and making standard analytic files available using a cohort of patients diagnosed with CKD from the Medicare 5% database.

CKD Data Sources

In the 2025 ADR, the traditional algorithm was used to define CKD based on diagnosis codes for CKD in claims. Diagnosis codes for CKD were grouped into "Y" codes and "A" codes. Y codes are canonical claims whose presence can confidently be used to infer the presence of CKD. In contrast, A codes are those codes which support the presence of CKD but which, in and of themselves, are not adequate to infer the presence of CKD. Following this, the "one inpatient or two outpatient claims" algorithm was applied. To infer the presence of CKD, presence one (inpatient) or more (outpatient) Y codes were sufficient. The presence of an A code required a confirmatory Y code in scenarios where two outpatient claims were required to infer the presence of CKD. The complete code list for defining CKD are in file "Codes for Comorbidity and CKD"

under the "Supplemental Methods" section in the CKD volume of the 2025 ADR (https://usrds-adr.niddk.nih.gov/2025/chronic-kidney-disease/ckd-supplemental-methods).

To create the CKD cohort dataset by year, the USRDS Coordinating Center uses 5% General Medicare Claims data and patient demographic information. The CKD cohort dataset contains three basic data components: Patient Master File, Payer Sequence File, and a series of Co-Morbid Files. Below is a description of how the CKD cohort dataset was created using the 5% General Medicare Claims data and the underlying data structure and file organization.

Patient Master File

For patients in this file, at least one CKD ICD-9/10 diagnosis code, an "A"-code or an "Y"-code, was identified in the 5% Inpatient (IP), Outpatient (OP), Home Health (HH), Hospice (HS), Skilled Nursing Facility (SNF), and Part B (PB, including DME) SAFs, with one record per patient. The two key variables CKD_xx (1-year entry period CKD indicator) and CKD_xxyy (2-year entry period CKD indicator) are defined as follows:

CKD_xx = Y if a patient in year xx (e.g. CKD_95 or CKD_02)

- Was Part A and Part B entitled and not enrolled in an HMO for the entire year (i.e. Payer_Seq_File) and

is defined as a CKD patient. CKD_xx = N otherwise

CKD_xxyy = Y if a patient in a two-year entry period (e.g. CKD_9596, CKD_9900, or CKD_0102)

- Was Part A and Part B entitled and not enrolled in an HMO for the entire two-year entry period (i.e. Payer_Seq_File)

is defined as a CKD patients CKD_xxyy = N otherwise

Patients with CKD_95 = Y are a subset of patients in the 1995 Co-Morbid File (Co_Morbid_1995) in the Co-Morbid Files dataset; patients with CKD_9900 = Y are a subset of the patients from the 1999 and 2000 Co-Morbid Files (Co_Morbid_1999 and Co_Morbid_2000) in the Co-Morbid Files dataset, respectively.

Some patients in the Patient Master File may have all CKD_xx = N and all CKD_xxyy = N. This implies that these patients had at least one CKD ICD-9 diagnosis code identified in the 1992 to 2010 5% IP, OP, HH, HS, SNF, and PB SAFs but did not meet the one-year entry period and two-year entry period CKD eligibility criteria described above.

Each patient is identified by two patient identification numbers. One is FIVEP_ID, a unique patient ID in the 5% Medicare sample population and the primary linking key within the USRDS CKD Cohort Finder dataset. Each year, new FIVEP_ID numbers are generated only for new patients added to the 5% sample. All existing patients (ongoing patients) are referenced with their previously assigned numbers. A second patient identification number, USRDS_ID, is the unique ID in the USRDS database system for identifying ESRD patients. Only patients with ESRD are assigned a valid USRDS_ID number. FIVEP_ID and USRDS_ID are not related.

Payer Sequence File

This file was created from the 5% Medicare denominator files and the IP, OP, HH, HS, SNF and PB SAFs. It contains Medicare coverage information for 5% patients with at least one CKD ICD-9/10 diagnosis code identified through the 5% IP, OP, HH, HS, SNF, and PB SAFs. A patient might have one or many sequence records in which each record indicates different insurance coverage within a well-defined period. There are five unique insurance types:

- A: Entitled with Part A only

- B: Entitled with Part B only
- AB: Entitled with Part A and Part B
- H: Enrolled in a Medicare HMO health plan
- NM: Non-Medicare

The first record contains the earliest available date for coverage information in the 5% Medicare denominator files. The last record contains the last known available date (i.e. the most recent) for coverage information in the 5% Medicare denominator files. FIVEP_ID is the primary patient identification number.

Co-Morbid Files

The Co-Morbid files are constructed from the 5% Medicare IP, OP, HH, HS, SNF, and PB SAFs and named individually by the respective calendar years. Each file contains all patients whose CKD diagnosis was identified from the 5% Medicare IP, OP, HH, HS, SNF, and PB SAFs within that year (i.e. Co_Morbid_1995 contains all patients whose CKD disease events were identified from the 1995 5% Medicare Claims SAFs). However, patients were not necessarily required to be entitled to Parts A and B and not enrolled in an HMO for that entire year. A patient might have one or many records within a calendar year in which each record depicts one unique CKD disease event. Records are considered duplicates if they are extracted from the same type of claim file with the same ICD-9/10 diagnosis code on the same date. These files do not include duplicates.

For example, each record in Co-Morbid_1995 consists of one CKD code per claim file per service date per unique patient in 1995. The combination of claim type, CKD code, and service date for a patient establishes uniqueness in each Co-Morbid File. FIVEP_ID is the primary patient identification number.

SECTION 12: MEDICARE PART D PRESCRIPTION DRUG COVERAGE

Part D Data Sources

The ESRD and pre-ESRD Part D data are obtained from ESRD and pre-ESRD CMS annually with USRDS provided Finder Files. The Part D data are divided into two separate files: an annual enrollment file containing monthly indicators of enrollment in Part D, and a prescription drug event file (PDE) containing details of prescriptions filled by Part D beneficiaries.

Section 101 of the Medicare Prescription Drug, Improvement, and Modernization Act of 2003 amended Title XVIII of the Social Security Act by establishing the Voluntary Prescription Drug Benefit Program (Part D). Effective January 1, 2006, Part D is an optional prescription drug benefit for individuals who are entitled to Medicare benefits under Part A or are enrolled in Medicare benefits under Part B. The data from the first few months of 2006, while the benefit was very new, may not be complete, and should be interpreted with caution.

Part D Enrollment SAF

Since the Part D benefit is voluntary, not all Medicare beneficiaries are enrolled in Part D. The annual enrollment file contains 12 monthly indicators that detail whether the beneficiary is enrolled in Part D, and if so, what type of plan the beneficiary is enrolled in. There are also monthly indicators for dual eligibility (Medicare and Medicaid), monthly indicators for Retiree Drug Subsidy, and Low-Income Subsidy (LIS). A complete listing of the variables included in the annual enrollment file is available in *2025 Researcher's Guide to the USRDS Database Appendix B: USRDS Data File Descriptions*.

The variables PTD_CNTRCTID_01- PTD_ CNTRCTID_12 indicate Part D enrollment as follows:

- H = Managed care organizations other than Regional Preferred Provider Organization
- R = Regional Preferred Provider Organization
- S = Stand-alone Prescription Drug Plan (PDP)
- E = Employer-sponsored
- O = Not Medicare enrolled
- X = Medicare enrolled, but no Part D enrollment record
- N = Not Part D enrolled

Values of H, R, S, and E for these variables are generally considered to indicate Part D enrollment.

Low-Income Subsidy status can be determined from the variables CST_SHR_GRP_CD_01- CST_SHR_GRP_ CD_12. If the beneficiary is Medicare enrolled and Part D enrolled:

- 01 = 100% premium subsidy and no copayment
- 02 = 100% premium subsidy and low copayment
- 03 = 100% premium subsidy and high copayment
- 04 = LIS, 100% premium-subsidy and high copayment
- 05 = LIS, 100% premium-subsidy and 15% copayment
- 06 = LIS, 75% premium-subsidy and 15% copayment
- 07 = LIS, 50% premium-subsidy and 15% copayment
- 08 = LIS, 25% premium-subsidy and 15% copayment
- 09 = No premium subsidy or cost sharing

If the beneficiary is Medicare enrolled and not Part D enrolled:

- 10 = Not enrolled in Part D, but the employer is entitled to RDS subsidy
- 11 = Creditable coverage but no RDS
- 12 = Not Part D enrolled, no creditable coverage, no RDS
- 13 = None of the above conditions have been met

Prescription Drug Event SAF

This SAF contains details on prescription drug utilization, including brand name, generic name, dosage form, drug strength, quantity dispensed, date of service, and total prescription cost. The variables in this SAF are listed in *2025 Researcher's Guide to the USRDS Database Appendix B: USRDS Data File Descriptions*. More detailed information on the Part D benefit can be found at *http://www.cms.gov/PrescriptionDrugCovGenIn/*.

SECTION 13: CROWNWEB CLINICAL DATA/EQRS CLINICAL DATA

In 2015, CROWNWeb became an integral part of the USRDS SAFs. CROWNWeb collects administrative and clinical data from all Medicare-certified dialysis facilities in the United States. It includes patient admission, tracking, and discharge information, CMS forms, and clinical data elements. Monthly clinical data submission began in May 2012. Previously, monthly clinical data were available to the USRDS for Medicare patients only through administrative claims. The USRDS CROWNWeb Clinical Data SAF includes data collected from May 2012 and on. Some clinical measures, including vitamin D analog measures and serum parathyroid hormone (PTH) labs, have not been collected since April 2014. However, PTH collection was resumed since July 2022. Facilities are asked to submit patient clinical information each month. The last values for a month are required to be entered by the end of the following month. For example, labs drawn in August should be entered by the end of September. There are out-of-range values in some variables. Therefore, users should be cautious with out-of-range values when conducting analyses.

In 2020 the CROWNWeb Clinical data became part of the End Stage Renal Disease Quality Reporting System (EQRS). See Section 1 Data Sources in this document.

Detailed information for the clinical measures included in the CROWNWeb Clinical Data SAF can be found in *2025 Researcher's Guide to the USRDS Database Appendix B: USRDS Data File Descriptions*.

SECTION 14: USE of APPENDIX B AND C of the RESEARCHER'S GUIDE TOGETHER

Appendix B provides a brief description and variable list for each data file in the USRDS SAFs, while Appendix C provides format contents. An analyst can use SAS programs to get the contents and formatting by incorporating the formats library.

Researchers who do not have access to SAS can use Appendix B & Appendix C for such information. For example: Tab 1 (PATIENTS) in Appendix B lists names, formats, and description of the variables in the Core PATIENTS file. The variable "RACE" is a character variable with length of 1. Its description is "Race of patient" and format is "RACEFMT.". In order to get the description of the values, Appendix C, tab R would be searched for RACEFMT. Below is what you can get from Appendix B and C.

Appendix B, PATIENTS tab:

RACE	Char	1	$RACEFMT.	Race of patient

Appendix C, R tab:

RACEFMT		1	WHITE	Char
		2	BLACK/AFRICAN AMERICAN	
		3	AMERICAN INDIAN OR ALASKA NATIVE	
		4	ASIAN	
		5	NATIVE HAWAIIAN OR PACIFIC ISLANDER	
		6	OTHER OR MULTIRACIAL	
		9	UNKNOWN	

SECTION 15: SAF sizes

Table 16 below lists the sizes of available SAFs. Please note, these are approximate files sizes; sizes vary slightly year by year. You must consider how many years you are requesting to calculate the space needed; you will also need to include some space for working files.

Table 16: File sizes

Data Category	Data Group	Size (Gb)	Data Type
ESRD			
	Core	5.5	Cumulative
	Hospital	4.2	Cumulative
	Transplant	2.6	Cumulative
	CrownWeb Clinical	36.1	Cumulative
ESRD			
	Claims		
	Medicare Claims Clinical	0.4	Yearly
	Institutional	27.7	Yearly
	Physician/Supplier	19.6	Yearly
	Part D	2.8	Yearly
Pre-ESRD			
	Claims		
	Institutional	5.4	Yearly
	Physician/Supplier	2.9	Yearly
	Part D	0.9	Yearly
CKD			
	Core	6.2	Cumulative
	Hospital	1.4	Cumulative
CKD			
	Claims		
	Institutional	5.7	Yearly
	Physician/Supplier	4.9	Yearly
	Part D	3.7	Yearly

SECTION 16: SPECIAL STUDIES

Clinical Performance Measures (CPM) Project (2000-2008)

CMS developed the ESRD CPM (formerly the ESRD Core Indicators Project) to collect information on the quality of care provided to dialysis patients. The data originate from surveys of approximately 10,000 dialysis patients as completed by primary care facilities and focus on dialysis adequacy measures, anemia management, and vascular access use. Additional clinical parameters such as albumin are also available. These data were collected annually beginning in 1994, using a random sample of patients aged 18 years and older, who were alive and on dialysis at the end of each calendar year; on average, about 8,500 in-center hemodialysis patients and 1,500 peritoneal dialysis patients were surveyed each year.

Data collection for all pediatric patients aged 12 to 17 years began in 2000, and in 2002 was expanded to all in-center hemodialysis patients aged younger than 18 years. Beginning in 2005, all peritoneal dialysis (PD) patients in the United

States were sampled for the pediatric PD data collection. In anticipation of the national release of the CROWNWeb system and its supporting performance measures reports, CMS concluded its CPM project in 2009, making 2008 its final survey year.

To further expand the value and use of the CPM data, we have linked patient data from the USRDS SAFs, enabling complete claims extraction from the SAFs for all identified patients. The resulting claims history has been combined with the CPM data to form a complete mini-set of the USRDS data products with supporting files. This enables researchers to add patient-level laboratory and dialysis prescription detail to a broad range of healthcare service event data over many years.

The USRDS Coordinating Center has made the CPM data available as SAFs. The dataset contains CPM data collected in surveys from 1994-2008. This dataset is not updated with current data.

Dialysis Morbidity and Mortality Study (1993-1997)

The Dialysis Morbidity and Mortality Study (DMMS) was an observational study that collected data on demographics, comorbidity, laboratory values, treatment, socioeconomic factors, and insurance for a random sample of U.S. dialysis patients, using dialysis records. Waves I, III, and IV were each prospective historical studies of a total of 16,812 participants in which data were collected for patients on in-center hemodialysis on December 31, 1993. Data were abstracted from patient medical records, and each patient was followed from December 31, 1993, through the earliest of data abstraction, death, transplant, change in modality, or transfer to another facility. Wave II is a true prospective study of incident hemodialysis and peritoneal dialysis patients for 1996 and some incident patients entering the ESRD program in the first part of the 1997 calendar year, and included 4,024 participants. These datasets also include Medicare claims for participants in DMMS. Patients were followed through the reported claims year.

Case Mix Adequacy Study (1990)

The objectives of the USRDS Case Mix Adequacy Study (CMAS) of Dialysis were to:

- Establish the relationship between the dose of delivered dialysis therapy and mortality
- Determine the strength of this relationship when data are adjusted for comorbidity
- Assess how this relationship changes at different dialysis doses
- Assess how this relationship is affected by dialyzer reuse
- Assess the impact of different dialysis membranes on patient morbidity and mortality

The study consisted of two groups of patients: an incident sample of ESRD patients who began hemodialysis during 1990, and a prevalent sample of hemodialysis patients with ESRD onset before 1990. A total of 7,096 patients from 523 dialysis units were included, the pre- and post-BUN values needed to calculate delivered dialysis dose were present for approximately 3,300 patients. Ninety-four percent of these cases were matched to the USRDS database. The ESRD Networks collected these data in conjunction with their Medical Case Review data abstraction. These datasets also include Medicare claims for participants in CMAS. Patients were followed through the reported claims year.

Case Mix Severity Study

Data were collected for 5,255 incident patients in 1986 and 1987 at 328 dialysis units nationwide. The objectives of this study were to:

- Estimate the correlation of comorbidity and other factors present at ESRD onset with subsequent mortality and hospitalization rates, adjusting for age, sex, race, and primary diagnosis

- Evaluate possible associations of these factors with reported causes of death
- Assess the distribution of comorbidity and other factors among patients using different treatment modalities
- Compare relative mortality rates by treatment modality, adjusting for selected comorbid conditions and other factors

Pediatric Growth and Development

The objectives of the USRDS Pediatric ESRD Growth and Development Study were to:

- Establish a baseline for assessing the relation of pediatric ESRD patient growth and sexual maturation to modality
- Establish a prototype for the ongoing collection of pediatric data

All patients prevalent in 1990 and born after December 31, 1970, were included in the study, a total of 3,067 patients at 548 dialysis units.

Continuous Ambulatory Peritoneal Dialysis and Peritonitis Study

The USRDS Continuous Ambulatory Peritoneal Dialysis (CAPD) and Peritonitis Rates Study examined the relation of peritonitis episodes in CAPD patients to connection device technology and other factors. The study population included all patients newly starting CAPD in the first six months of 1989, up to a maximum of 14 patients per dialysis unit. All units providing CAPD training participated in the study. The sample includes 3,385 patients from 706 units.

Active/Adipose (2009-2013)

Active/Adipose is a cohort study to investigate the value of exercise in ESRD and was designed to examine the paradox of obesity and survival in ESRD. The Active Adipose study is a prospective, multicenter observational study designed to characterize the prevalence and progression of frailty among patients undergoing hemodialysis for end-stage renal disease (ESRD).

Patients at least 18 years of age who were on hemodialysis for at least three months were enrolled in the study. The study duration is from 2009 to 2013. A total of 771 patients in 14 dialysis centers were enrolled in the study. Participants were assessed using the validated Fried frailty index, which defines frailty as the presence of 3 or more criteria, including recent unintentional weight loss, reported exhaustion, low grip strength, slow walk speed, and low physical activity. Information was collected on participants' demographics, medical history, activity status, and mood, and physical performance was assessed via tests of grip strength, walk speed, and body composition. Additionally, blood draws were conducted semi-annually for evaluation of nutrition, inflammation, and cardiac markers. All participants were followed for two years. This dataset also contains Medicare claims for participants in AA. Patients were followed for the reported claims year.

The Transition of Care in Chronic Kidney Disease --KPSC

The Transition of Care in Chronic Kidney Disease (TCCKD) Special Study Center examines the transition of care to renal replacement therapy (RRT; i.e. dialysis or transplantation) in patients with very-late-stage (advanced) nondialysis dependent (NDD) CKD. These are often people with an estimated glomerular filtration rate (eGFR) <25 ml/min/1.73 m^2.

From 01/01/2008 through 12/31/2012, 5,778 CKD patients who transitioned to RRT in Kaiser Permanente of Southern California were identified and linked to the USRDS ESRD database. The TCCKD_MASTER_FILE contains *USRDS_ID*,

demographics (i.e. age, race and ethnicity, marital status), and comorbid conditions (i.e. coronary artery disease (CAD), weighted Charlson comorbidity index score at the time of RRT transition). TCCKD_LAB contains *USRDS_ID*, vital signs, i.e. systolic blood pressure (SBP), diastolic blood pressure (DBP), body mass index (BMI); and laboratory data, i.e. sodium, potassium, bicarbonate, uric acid, hemoglobin, white blood cells (WBC), platelets, thyroid stimulating hormone (TSH), and estimated glomerular filtration rate (eGFR). Time periods are labeled as Q12 to Q1 to stand for from 12th quarter to the quarter before the RRT transition. Quarterly averaged values of all Vital signs data, and Laboratory data are provided from 12th quarter to the quarter before the RRT transition.

The Transition of Care in Chronic Kidney Disease – Veterans Health Administration

Between 10/1/2007 and 3/30/2015, 100,318 U.S. veteran patients who transitioned to renal replacement therapy across the nation were identified. This database was created from the linkage between the national U.S. Renal Data System (USRDS) and national Veterans Health Administration databases. This dataset is not updated with current data.

Vital signs and laboratory data are presented as quarterly (3 months) averaged values collected from 1 to 12 quarters prior to renal replacement therapy transition.

In the data, variable CCI (Calculation of weighted Charlson Comorbidity Index) is calculated by the following formula: CCI= 1* Myocardial Infarction + 1* Congestive Heart Failure + 1* Peripheral Vascular Disease + 1* Cerebrovascular Disease + 1* Dementia + 1* Chronic Pulmonary Disease + 1* Rheumatic Disease + 1* Peptic Ulcer Disease + 1* Mild Liver Disease + 1* Diabetes without chronic complications + 2* Diabetes with chronic complications + 2* Paraplegia or Hemiplegia + 2* Any Cancer + 3* Moderate or Severe Liver Disease + 6* Metastatic Carcinoma + 6*AIDS/HIV

Comprehensive Dialysis Study (2006)

The Comprehensive Dialysis Study (CDS) was a USRDS special data collection study to assess rehabilitation/quality of life and nutrition issues in incident dialysis patients. Data collected included patient demographics, contact information, treatment, laboratory values, quality of life (QOL) survey interviews, and nutrition survey interviews for U.S. dialysis patients who started treatment during 2005-2007 at 335 randomly selected dialysis facilities. Patients were selected via a monthly REMIS/SIMS database custom extract process using the following criteria:

- Incident dialysis patients (age ≥ 18) who had survived for at least two months
- No prior transplant
- Dialyzed at one of the pre-determined 335 facilities

Over a period of two years, 11,292 patients were selected. Of these, 1,678 consented and participated in the surveys and/or lab assays. The CDS SAF contains the data for these patients. Of the 1,678 patients, 1,279 participated only in the QOL survey — 364 participated in the QOL and nutrition surveys and/or lab assays (227 with QOL, lab, and nutrition; 4 with QOL and lab; and 133 with QOL and nutrition; and 35 did not participate in either survey but provided lab data.

The CDS SAF dataset consists of:

- CDS Patient File. This file contains one record for each of the 1,678 patients. A patient record includes a unique identifier USRDS_ID (unique identifier used to cross-reference to other USRDS SAF files), demographic data, and a patient category indicator (PAT_CAT_BL), which indicates baseline survey and lab participation status (QOL-only, QOL w/lab data, QOL and nutrition w/o lab data, Lab data w/o survey).

- CDS QOL Baseline File. This file contains the Baseline QOL survey answer data and some derivative score data from QOL only and QOL/Nutrition participants.
- CDS Food Baseline File. This file contains the Baseline Nutrition survey answer data and data generated by "Block Dietary Data System" on QOL/Nutrition participants.
- CDS Lab File. This file contains up to five sets of lab data from a subset of QOL/Nutrition participants.
- CDS Dictionary. This file contains contents and variable formats for each CDS SAF file.