

ESRD Data Sources Overview

This document provides an overview of publicly-available data sources related to analysis of the U.S. dialysis industry and the U.S. dialysis patient population. We review about 25 data sources in this report.

1. Medicare ESRD and Related Public-Use Files

CMS makes available a number of data sets on a regular basis that include a wide range of information on dialysis facilities and ESRD patient experience. These files fall into two categories: 1) those that are free to download without restriction and 2) those that require prior approval and data fees. Links to most file sources can be found in the References section at the end of this report.

Free Datasets.

Dialysis Facility Compare (DFC). The Dialysis Facility Compare is the best public data source for a current (updated quarterly) snapshot of the U.S. dialysis industry, with demographic and operational characteristics and clinical and quality measures on almost all Medicare-certified facilities in the US and Territories. DFC has historically had a lag of a year or more in adding new facilities and may lag even longer in updating facility ownership information, but seems to be getting better at updating in a more timely fashion recently. Data can be downloaded in csv, MS Excel or MS Access formats and a data archive back to 2001 has recently been added to the DFC website. However, file formats have changed regularly over the years. Prima Health Analytics maintains a DFC History database back to 2001 in a common record format.

Dialysis Facility Reports (DFRs). The annual Dialysis Facility Reports contain a wealth of information about virtually all dialysis clinics and the patient populations they serve in the U.S, including data on mortality, morbidity, transplantation, vaccinations, vascular access and other clinical process and quality outcomes, patient characteristics, and state survey results. There are no economic or financial data. Data sources used include Medicare claims, Medicare Eligibility Forms, Fistula First/CROWNWeb, and a number of other secondary sources. Currently, there is no public download site for the reports in any form, although some historical reports through 2012 are available in PDF format from: <https://projects.propublica.org/dialysis/>. Prima Health Analytics has full data sets of the DFRs for 2010, 2011 and 2012, and is in process of trying to acquire the 2013 and/or 2014 data sets.

Medicare Renal Cost Reports/Hospital Cost Reports. The cost reports from HCRIS (Health Cost Report Information System) are a very rich resource for financial and utilization information for all dialysis facilities. Free-standing facilities file the Renal Cost Report and hospital-based dialysis centers' (HBDCs) statistics are filed on renal department schedules as part of the overall hospital cost report. Formats differ between hospital and free-standing cost reports and

**Prima Health Analytics
December 2014**

have changed over time for both report types. Prima Health Analytics has built and maintains a Consolidated Cost Report Database (CCRD) which combines data from pre and post PPS reporting structures and from hospital and renal cost reports into a common data record for consistent reporting across facility types and time periods. Currently, we maintain a 10-year history of dialysis cost reports which is updated quarterly.

Medicare ESRD QIP Results Files. Each year a set of QIP Performance Score Summary Reports are made available with details of the QIP Performance Score by measure and the calculation of Payment Adjustments, by facility. Facility identifiers include Medicare Cert #, Name, Address, and Facility Ownership information.

Medicare ESRD PPS Facility-level Impact Files. As part of ESRD PPS rule-making, for both proposed and final rules, CMS releases a facility-level impact file which has important payment characteristics for each facility which are not available from any other public source. These include things like low volume adjuster status, estimated outlier payment amounts/percents and information from which the prior year payment case-mix can be derived.

Medicare ESRD PPS Claims-Based Monitoring Data. Twice annually CMS releases a summary report/data file of key trends in treatment of ESRD. Trends in General Mortality & Morbidity, Anemia Management, Vascular Access, Home Dialysis, Training, & Onset, Bone & Mineral Management, and Fluid Management are reported. The data are summarized across all providers and are reported by month, but this is usually the most up-to-date set of statistics available, if close to real-time benchmarks are needed. For example, the Q2 2014 monitoring file released in early November contains statistics reported through June 2014. By contrast, the USRDS Annual Data Report released at about the same time has data only through CY 2012.

Medicare Part B Drug Average Sales Price (ASP) Files. While not strictly an ESRD data source, the Part B Drug ASP files provide a series of cost benchmarks for ESRD drugs, reflecting the national average sales price, by calendar quarter, for each drug used in the treatment of dialysis patients. These files are useful for benchmarking and modeling costs over time and relative to national averages.

Restricted Datasets.

Medicare ESRD PPS Limited Data Set. This file contains select claim level data from ESRD facility outpatient claims for all ESRD Type 72X claims processed in the most recent calendar year (currently 2013.) It includes provider numbers, service dates, revenue codes, diagnosis codes, CPT/HCPCS codes, units of service, payment amounts, etc. Available on DVD for a cost of \$600, application for a copy of this file requires completion of a Data Use Agreement and a short study protocol. This is a special type of the more general Limited Data Set category of files that are available (see next entry).

Medicare Standard Analytical Files Limited Data Sets (LDS) - 5% Sample Files. These are claims files for a 5% sample of all Medicare beneficiaries across all providers in a given year. These files are available by type of claim or collectively as a group. The 5% sample is created based on selecting records with 05, 20, 45, 70 or 95 in positions 8 and 9 of the Health Insurance Claim (HIC) number. File cost is per year and files are in CSV format. Application for these files requires completion of a Data Use Agreement and a study protocol.

2. United States Renal Data System (USRDS)

USRDS has been the go-to resource for researchers needing detailed ESRD patient-level data for decades. While some USRDS data are free to the public, generally the data sets are time-consuming and expensive to acquire.

USRDS Annual Data Reports (ADRs). The national statistics in the Annual Data Reports from USRDS are free for download in Excel Reference Tables and cover a wide range of topics, with an epidemiological focus (e.g. – incidence, prevalence, comorbidities, etc.) There is relatively little on dialysis providers or economic data, but long-term trends are available – in some cases over 30 years. The statistics lag two years; that is, the 2014 ADR contains data only through 2012.

USRDS Standard Analytical Files (SAFs). The USRDS SAFs are a set of very detailed ESRD patient-specific claims and eligibility data from the USRDS database to support ESRD research. Users must sign a data release agreement with NIH and provide documentation of approval from an Institutional Review Board (IRB). Costs are in the thousands of dollars, depending on how many years of data are needed and/or which types of claims files are desired. For general claims analysis, obtaining USRDS SAFs is more complex and expensive than getting the Medicare 5% sample SAFs, but the data sets are ESRD-specific.

3. Other ESRD Data Sources

In addition to CMS and USRDS, there are a few other organizations that make ESRD data available to the public, to ESRD providers, or both. Access is in some cases restricted to members of the organization or users of the system described.

CDC Chronic Kidney Disease Surveillance Project. The CDC CKD website provides data on incidence, prevalence and treatment of all stages of CKD, and as such may be a good resource to augment the ESRD/dialysis data sources.

DOPPS DPM/Data for Researchers. Data collected by the Dialysis Outcomes and Practice Patterns Study (DOPPS) may be available for nominal cost, subject to DOPPS approval. DOPPS collects data on dialysis practice patterns from samples of dialysis facilities in the U.S. and around the world. Some summary data are available on a range of clinical measures for free download on their DOPPS Dialysis Practice Monitor (DPM) website.

Prima Health Analytics December 2014

Elab Project. The Elab project collects lab data from all 18 ESRD Networks and nearly 100% of dialysis patients and facilities. Facility-specific reports generated from the lab data offer comparative data at the state and Network level and can be used for quality improvement purposes. National reports and Trend Reports are also available.

ESRD Networks Custom Data Reports. Elab (see above) is an example of a data collection and reporting activity of the ESRD Networks. Other data may be available from the Networks that can be turned into custom extracts upon request. For example, Prima Health Analytics received reports of the number dialysis patients by modality and zip code for all Networks for a study several years ago. For Medicare ESRD providers, contact your ESRD network data resource person for more information on current data availability.

CROWNWeb. CROWNWeb is a resource available to ESRD providers, but cannot be accessed by the general public. Data availability is under development.

NRAA/RSE Benchmarking Tool. Organizations that are members of the National Renal Administrators Association (NRAA) now have access to the RSE Benchmarking Tool, a set of national benchmark reports on the US Dialysis Industry created for the NRAA by Prima Health Analytics. Reports are available for download in PDF format and cover a wide range of cost, clinical and quality measures for the past 5 years. While no member-accessible database is available at this time, custom reports are available from Prima Health Analytics for a modest cost per clinic.

Prima Health Analytics' ESRD Data Warehouse. Prima Health Analytics maintains an ESRD Data Warehouse that links data from many of the data sources listed above by facility, and then enhances the data with geographic, SES, and up-to-date ownership and affiliation data for each facility. The resulting data warehouse contains over 6000 cost, utilization, quality, demographic and geographic variables, covering 100% of current Medicare ESRD providers (N=6568 facilities as of June 2014). Depending on data source, the database covers between 6 and 14 calendar years (2001-14). The national scope of the database can help decision-makers a) define the magnitude of a problem, b) identify where to focus efforts and c) set policy. It can provide an efficient and speedy method to acquire intelligence before investing in larger, more time consuming initiatives. Possible applications include clinical trial design, market access planning, product positioning and life cycle management.

4. General Population and Geographic Databases

For many studies, it is important to understand the geographic and socio-economic background for the ESRD data being analyzed. For example, regional and/or SES-related disparities that exist in the general health care system or social environment may have important impacts on ESRD outcomes. The following free data resources can be sources for this kind of background “confounder” or “explanatory variable” data. Linking geographically-based healthcare data to dialysis facility data is a specialty skill at Prima Health Analytics, where we use a spatially-enabled ESRD database to match up the data at whatever level of geographic or geo-political

Prima Health Analytics December 2014

granularity is needed, from census tract or block level up to zip code, county, congressional district, state, or region levels.

County Health Rankings. County Health Rankings provides a set of annual free analytic data sets in various formats with healthcare statistics for all states and counties in the U.S. The County Health Rankings provide a snapshot of a community's health and a starting point for investigating and discussing ways to improve health. Health measures are weighted to provide overall or summary measures as well.

Dartmouth Atlas of Health Care. For more than 20 years, the Dartmouth Atlas Project has documented variations in how medical resources are distributed and used in the United States. The project uses Medicare data to provide information and analysis about national, regional, and local markets, as well as hospitals and their affiliated physicians. Data are organized around Hospital Referral Regions (mapped from zip codes) and cover topics such as hospital discharges, end of life care, Medicare spending, and physician density.

Medicare Geographic Variation Data Sets. CMS has developed data that enables researchers and policymakers to evaluate geographic variation in the utilization and quality of health care services for the Medicare fee-for-service population. They have aggregated this data into a Geographic Variation Public Use File that has demographic, spending, utilization, and quality indicators at the state level (including the District of Columbia, Puerto Rico, and the Virgin Islands), hospital referral region (HRR) level, and county level. There are statistics related to ESRD services included as well. The most recent file covers 2012 data.

U.S. Census Bureau American Fact Finder. The U.S. Census Bureau makes available a vast range of socio-economic data on the U.S. population at the census block, tract, ZCTA, county, state, and Census Region levels, downloadable in flat file formats from the American Fact Finder website. By linking these socio-economic statistics to the geography of dialysis facilities, ESRD research can be enhanced with important data on the age and racial composition, educational status, income characteristics, etc. of the community served by the facility. Data from the annual American Community Survey (ACS) are updated annually in December for the prior year, so that 3 and 5 year estimates of all census statistics are regularly updated at all levels of geographic granularity.

5. Private/Proprietary Data Sources

Several large dialysis and pharmaceutical companies offer data for sale or offer collaborative opportunities for ESRD-related studies. Both Davita (Davita Clinical Research) and Fresenius (Frenova) have research divisions and databases that can support clinical studies as well as health economics research.

**Prima Health Analytics
December 2014**

References/Links:

CDC Chronic Kidney Disease Surveillance Project: <http://nccd.cdc.gov/CKD/>

County Health Rankings: <http://www.countyhealthrankings.org/>

Dartmouth Atlas of Health Care: <http://www.dartmouthatlas.org/>

Dialysis Facility Compare: Downloadable files: <https://data.medicare.gov/data/dialysis-facility-compare>

DOPPS Data for Researchers: <http://www.dopps.org/ForResearchers.aspx>

DOPPS Dialysis Practice Monitor: <http://www.dopps.org/DPM/Default.aspx>

Elab Project National and Trends Reports: <http://www.esrdnet11.org/Elab/>

Medicare ESRD QIP Results Files: Download from: http://www.cms.gov/Medicare/quality-initiatives-patient-assessment-instruments/esrdqip/08_reportandcert.html

Medicare ESRD PPS Facility-level Impact Files: 2015 Final Rule Impact File can be downloaded from this page: <http://www.cms.gov/Medicare/Medicare-Fee-for-Service-Payment/ESRDpayment/End-Stage-Renal-Disease-ESRD-Payment-Regulations-and-Notices-Items/CMS-1614-F.html?DLPage=1&DLSort=3&DLSortDir=descending>

Medicare ESRD PPS Limited Data Set: <http://www.cms.gov/Research-Statistics-Data-and-Systems/Files-for-Order/LimitedDataSets/EndStageRenalDiseaseSystemFile.html>

Medicare ESRD PPS Claims-Based Monitoring Data: The most recent file can be downloaded from: <http://www.cms.gov/Medicare/Medicare-Fee-for-Service-Payment/ESRDpayment/Downloads/ESRD-PPS-Public-Release-File-Q2-2014.zip>

Medicare Geographic Variation Data Sets: http://www.cms.gov/Research-Statistics-Data-and-Systems/Statistics-Trends-and-Reports/Medicare-Geographic-Variation/GV_PUF.html

Medicare Hospital Cost Reports: Cost Report Downloads by Year accessed from this page: <http://www.cms.gov/Research-Statistics-Data-and-Systems/Files-for-Order/CostReports/Cost-Reports-by-Fiscal-Year.html>

Medicare Standard Analytical Files Limited Data Sets (LDS) - 5% Sample Files: <http://www.cms.gov/Research-Statistics-Data-and-Systems/Files-for-Order/LimitedDataSets/StandardAnalyticalFiles.html>

Medicare Renal Cost Reports: Downloadable files: <http://www.cms.gov/Research-Statistics-Data-and-Systems/Files-for-Order/CostReports/Downloads/RNL11/RNL11-All-Years.ZIP>

**Prima Health Analytics
December 2014**

Medicare Part B Drug Average Sales Price Files: <http://www.cms.gov/Medicare/Medicare-Fee-for-Service-Part-B-Drugs/McrPartBDrugAvgSalesPrice/index.html>

NRAA/RSE Benchmarking Tool: <http://www.renalexchange.com/rse-releases-benchmarking-tool-3-0/>

Prima Health Analytics' ESRD Data Warehouse Overview:
http://primahealthanalytics.com/public_docs/PHA%20ESRD%20Provider%20Data%20Warehouse%20Overview.docx

U.S. Census Bureau American Fact Finder:
<http://factfinder2.census.gov/faces/nav/jsf/pages/searchresults.xhtml?refresh=t>

USRDS Annual Data Reports: <http://www.usrds.org/adr.aspx>

USRDS Standard Analytical Files: <http://www.usrds.org/request.aspx>