End Stage Renal Disease Population Public Use File

Technical Fact Sheet

Overview

The End Stage Renal Disease (ESRD) Population Public Use File (PUF) contains the current and historical ESRD patient population and associated clinical data for both Medicare and non-Medicare patients. The purpose of this new PUF is to provide the renal research community, providers, and other interested stakeholders access to a robust dataset free of Protected Health Information (PHI) and Personally Identifiable Information (PII) to support research on ESRD, which ultimately benefits patient care and positive outcomes in dialysis facilities. This de-identified dataset is a first for the CMS ESRD Program.

Scope

The ESRD Population PUF includes ESRD patient data from 1973 to present and ESRD patient clinical data from 2012 to present. Patient and clinical data associated with transient admissions and acute discharges are excluded from this PUF. The ESRD Population PUF consists of 16 tables in Comma-Separated Value (CSV) file format as follows:

- Patient Data Table
- Form 2728 Table
- Admit Treatment Table
- 13 Clinical Measure Tables

The ESRD patient population is selected and extracted from the Centers for Medicare & Medicaid Services (CMS) Renal Management Information System (REMIS), which contains all ESRD dialysis and transplant patients. The CMS Consolidated Renal Operations in Web-Enabled Network (CROWNWeb) system is the source of clinical data/measures extracts and supporting patient data. Only *new to ESRD* or *returning to ESRD* patients added to CROWNWeb after 06/14/2012 will have associated clinical data (the system previous to CROWNWeb did not collect clinical data.) The data is integrated such that patient clinical data, if it exists, is matched to the REMIS patient record and both patient and corresponding clinical data are included in the PUF. Together, REMIS and CROWNWeb comprise the congressionally-mandated System of Record (SOR) for all ESRD Patients in the U.S. both Medicare-entitled and non-Medicare [SOR #09-70-0520]. A complete list of the PUF data elements is included in the supporting document "ESRD Population PUF Data Descriptions". The PUF and associated documentation will be updated quarterly and posted on the cms.gov website, here:

https://www.cms.gov/Medicare/End-Stage-Renal-Disease/ESRDGeneralInformation/Data.html.

De-identification Methodology

Each patient in the PUF has a unique unintelligent identifier (PAT_ID) linked to the Clinical Measure tables. Users will be able to join tables by using the PAT_ID. Patient data elements with PHI/PII are de-identified in adherence with the Health Insurance Portability and Accountability Act (HIPAA) of 1996 Privacy Rule "Safe Harbor" guidelines (45 C.F.R. § 164.514(b)(2)) and meet the security and privacy criteria set forth by the CMS Office of Information Technology (OIT) Privacy Office. The HIPAA Safe Harbor Guidelines are posted on the hhs.gov website, here:

https://www.hhs.gov/hipaa/for-professionals/privacy/special-topics/de-identification/index.html

In order to balance the protection of patient privacy with the utility of the PUF, some patient data elements will be suppressed, such as full date of birth/death, full treatment dates, address, Medicare Beneficiary Identifier (MBI), Provider Number, and all others that could be used to reveal patient identity. In lieu of providing exact dates for patient data elements, such as Date of Birth, only the calendar year is included.

The Provider Number (PROV_NBR) will be replaced in the PUF with a non-intelligent system generated number to be maintained in a crosswalk by CMS.

In order to provide users with the ability to calculate time elapsed between events without increasing the risk of re-identification, the ***Study Day*** method is used to de-identify data in the clinical measure tables. Where available and calculable, the date the patient entered into the ESRD Program is deemed the ***Reference Date*** and all clinical collection and reporting dates are removed and reported as ***Study Days***. The ***Study Day*** is calculated as the number of days elapsed from the patient's ***Reference Date*** so the passage of time between clinical events and the sequenced order of clinical events is preserved. An example is provided in Table 1 below.

How Dates and Passage of Time are represented in the Data Files

All dates used in the files that comprise the ESRD Population PUF are "relative dates" which are calculated as relative days from the patient ***Reference Date*** (year the patient entered the ESRD Program). The PUF will not include any elements of dates except for year. Relative dates will be shown as number of the number of Study Days elapsed since the ***Reference Date***. The example patient below is meant to provide a more detailed understanding.

Example

A patient is admitted to the ESRD Program by the physician who attests to their condition on ***04/15/2015*** and completes a CMS-2728 (End Stage Renal Disease Medical Evidence Report -- Medicare Entitlement and/or Patient Registration) form with signature. For Medicare coverage purposes, the date on the CMS-2728 is the date the patient entered the ESRD Program, so the ***Reference Date*** for this patient is ***04/15/2015.*** All clinical lab collection dates for this patient will be reported as ***Study Days*** -- the number of days relative to ***04/15/2015***. The ***Reference Date*** itself is never reported in the PUF. See an example in Table 1.

Table 1: Relative Study Date Example

Clinical HD Adequacy De-identified Dates Data Element	Description of Event	Original Dates not included in the PUF files	Elapsed Study Days	Format
PAT_ID	Patient ID from REMIS-internal system generated ID	….	….	Number
….	Patient entered the ESRD Program	04/15/2015	0	Number
CLNCL_MO_YR	Clinical Reporting Date	89	16	Number
HD_KTV_COLL_DATE	Kt/V Hemodialysis Collection Date	05/02/2015	17	Number
HD_KTV_COLL_DATE	Kt/V Hemodialysis Collection Date	05/25/2015	40	Number

In the example above, the Clinical Month Year (***CLNCL_MO_YR***) value of ***89*** represents a calendar date of ***May/2015*** which will be reported as 16 ***Study Days*** after the patient's ***Reference Date***. The de-identified dates reported as ***Study Days*** have an implied order of event sequence showing the first Kt/V Hemodialysis collection date (***HD_KTV_COLL_DATE***) occurred 17 ***Study Days*** after the patient's ***Reference Date*** while the second event occurred 40 ***Study Days*** after the ***Reference Date***. Researchers can determine that 23 days elapsed between the first collection date and the subsequent collection date without knowing the month and/or day that the events occurred.

Validation of Data

Data submitted by ESRD facilities and support organizations into the REMIS and CROWNWeb applications and associated databases are considered valid because they have passed multiple business rule and format edits. This data is extracted for the PUF as submitted and not changed, updated or transformed from the

original source databases except to remove certain identifier data elements (e.g. date of birth) as described above in order to meet HIPAA Privacy Rules for the de-identification process.

Sample Use Cases

Use Case 1[1] Association of Albumin Levels and Mortality in HD Patients

Description: A longitudinal clinical cohort study to determine the association between time-varying serum albumin level and the mortality rate in hemodialysis patients using the ESRD Population PUF. This study looks at adult ESRD hemodialysis patients, stratified by their primary diagnoses to determine the association between serum albumin levels and mortality rates.

Background: Low levels of serum albumin are indicative of protein energy wasting in dialysis patients and a strong predictor of the mortality risk in the ESRD population. There is supporting evidence that the factors causing low albumin levels may be associated with high mortality and morbidity in dialysis patients. Hence, regular monitoring of serum albumin levels can be useful for predicting outcomes in dialysis patients.

Table 2: Patient Table

Patient Table
Pat_ID
Age_Range
DOD
Gender
Primdiag

Table 3: Clinical Mineral Metabolism

Patient Table
Pat_ID
Modality: Where the Modality =HD
Clncl_month_yr
Albumin

[1] Chen, JB., Cheng, BC., Yang, CH. et al. BMC Nephrol (2016) 17: 117. https://doi.org/10.1186/s12882-016-0332-5

Figure 1: Joining Patient Table to Clinical Mineral Metabolism Table

Pat_ID	Age_Range	Gender	DOD	PrimDiag
100	55 to 64	Male	1989	Diabetes with renal manifestations Type 2
107	45 to 54	Female	1994	Hypertension: Unspecified with renal failure
2100	65 to 74	Female	2014	Glomerulonephritis (GN) (histologically not examined)
3149	45 to 54	Male	2015	Diabetes with renal manifestations Type 2

Pat_ID	Prov_Nbr	Clncl_Mo_Yr	Modality	Albumin
100	070000011	187	HD	3.3
100	070000011	502	HD	4.5
100	800000113	548	HD	4.6
2100	0766622	65	HD	3.8
2100	0766622	190	HD	3.7
3149	920000116	12	HD	3.9
3149	920000116	168	HD	4.2
3149	920000116	195	HD	4.2

Use Case 2: Examining Multiple, Non-Transfer Admissions for ESRD Patients

Description: Use Case 2 entails examining multiple, non-transfer admissions for ESRD patients using the ESRD Population PUF.

Background: Multiple, non-transfer admissions for ESRD may be more common among specific population demographics. This study examines common causes of renal failure, and average time (in days) between admissions among a population of ESRD patients with multiple, non-transfer admissions, by the population demographic characteristics.

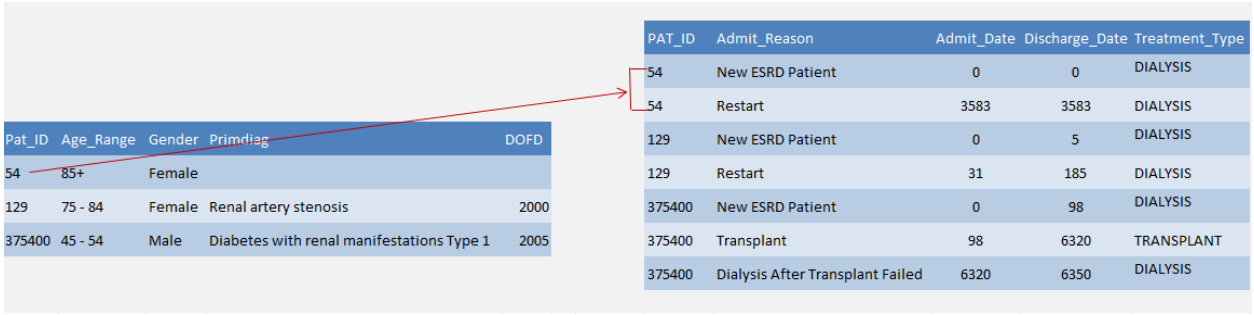
Table 4: Patient Table

Patient Table
Pat_ID
Age_Range
Gender
Primdiag

Table 5: Admit Treatment Table

Patient Table
Pat_ID
Admit_Reason: Filter on <> Transfer In
Admit_Date: Calculate time between admissions
Discharge_Date

Figure 2: Joining Patient Table to Admit Treatment Table



Pat_ID	Age_Range	Gender	Primdiag	DOFD
54	85+	Female		
129	75 - 84	Female	Renal artery stenosis	2000
375400	45 - 54	Male	Diabetes with renal manifestations Type 1	2005

PAT_ID	Admit_Reason	Admit_Date	Discharge_Date	Treatment_Type
54	New ESRD Patient	0	0	DIALYSIS
54	Restart	3583	3583	DIALYSIS
129	New ESRD Patient	0	5	DIALYSIS
129	Restart	31	185	DIALYSIS
375400	New ESRD Patient	0	98	DIALYSIS
375400	Transplant	98	6320	TRANSPLANT
375400	Dialysis After Transplant Failed	6320	6350	DIALYSIS