

Table IA. Demographic Summary

This table contains general descriptive information about the patients in the DEMOGRAPHIC table. These patients may or may not be represented in other CDM tables.

	N	%	Source table
Patients	966,870		DEM_L3_N
Age			DEM_L3_AGEYRSDIST1
Mean	45		
Median	45		
Age group			DEM_L3_AGEYRSDIST2
0-4	27,956	2.9	
5-14	99,826	10.3	
15-21	74,694	7.8	
22-64	518,621	53.6	
65+	245,773	25.4	
Missing	0	0.0	
Hispanic			DEM_L3_HISPDIST
N (No)	644,649	66.7	
Y (Yes)	20,064	2.1	
Missing or Refused	302,157	31.2	
Sex			DEM_L3_SEXDIST
F (Female)	491,410	50.8	
M (Male)	475,459	49.2	
Missing or Ambiguous	1	0.0	
Race			DEM_L3_RACEDIST
White	636,931	65.9	
Non-White	35,140	3.7	
Missing or Refused	294,799	30.5	

The four 'flavors of null' defined in the CDM are combined in the missing categories shown here, but details are available in the source tables.

Cell counts between 1 and the low-cell count threshold are displayed as BT and treated as missing.

Sums which include BT values are marked with an asterisk (*), either accompanying the sum of all above-threshold values or alone if all values are BT.

Age is calculated as current age or age at death if death date is known. If multiple death records exist, the earlier death date is used.

Percentages are sums of the percentages in the source table and are subject to rounding errors.

Race distribution is shown for all patients and for patients seen after Meaningful Use standards were widely implemented.

Table IA. Demographic Summary (continued)

This table contains general descriptive information about the patients in the DEMOGRAPHIC table. These patients may or may not be represented in other CDM tables.

	N	%	Source table
Race among patients with at least 1 encounter after December 2011			XTBL_L3_RACE_ENC
White	603,197	69.0	
Non-White	33,198	3.8	
Missing or Refused	237,988	27.2	
Gender Identity			DEM_L3_GENDERDIST
GQ (Genderqueer/Non-Binary)	0	0.0	
M (Man)	0	0.0	
W (Woman)	0	0.0	
MU (Multiple gender categories), SE (Something else), TF (Transgender female/Trans woman/Male-to-female), or TM (Transgender male/Trans man/Female-to-male)	0	0.0	
Missing or Refused	966,870	100.0	
Sexual Orientation			DEM_L3_ORIENTDIST
Bisexual	0	0.0	
Gay	0	0.0	
Lesbian	0	0.0	
Queer	0	0.0	
Straight	0	0.0	
AS (Asexual), MU (Multiple sexual orientations), SE (Something else), QS (Questioning)	0	0.0	
Missing or Refused	966,870	100.0	

The four 'flavors of null' defined in the CDM are combined in the missing categories shown here, but details are available in the source tables.

Cell counts between 1 and the low-cell count threshold are displayed as BT and treated as missing.

Sums which include BT values are marked with an asterisk (*), either accompanying the sum of all above-threshold values or alone if all values are BT.

Age is calculated as current age or age at death if death date is known. If multiple death records exist, the earlier death date is used.

Percentages are sums of the percentages in the source table and are subject to rounding errors.

Race distribution is shown for all patients and for patients seen after Meaningful Use standards were widely implemented.

Table IB. Potential Pools of Patients

This table illustrates the number of patients meeting different inclusion criteria and supports Data Check 3.04 (less than 50% of patients with encounters have DIAGNOSIS records) and Data Check 3.05 (less than 50% of patients with encounters have PROCEDURES records). Data check exceptions are highlighted in red and must be corrected.

Metric	Metric Description	Result	Source table
Potential pool of patients for observational studies	Number of unique patients with at least 1 ED, EI, IP, OS, or AV encounter within the past 5 years	679,793	ENC_L3_DASH2
Potential pool of patients for trials	Number of unique patients with at least 1 ED, EI, IP, OS, or AV encounter within the past 1 year	341,359	ENC_L3_DASH2
Potential pool of patients for studies requiring data on diagnoses, vital measures and (a) medications or (b) medications and lab results	Number of unique patients with at least 1 encounter and DIAGNOSIS and VITAL records within the past 5 years	499,198	XTBL_L3_DASH1
	Number of unique patients with at least 1 encounter and DIAGNOSIS, VITAL, and PRESCRIBING or DISPENSING records within the past 5 years	461,154	XTBL_L3_DASH2
	Number of unique patients with at least 1 encounter and DIAGNOSIS, VITAL, PRESCRIBING or DISPENSING, and LAB_RESULT_CM records within the past 5 years	460,216	XTBL_L3_DASH3
Patients with diagnosis data	Percentage of patients with encounters who have at least 1 diagnosis	98%	ENC_L3_N; DIA_L3_N
Patients with procedure data	Percentage of patients with encounters who have at least 1 procedure	84%	ENC_L3_N; PRO_L3_N

Cell counts between 1 and the low-cell count threshold are displayed as BT and treated as missing.

Sums which include BT values are marked with an asterisk (*), either accompanying the sum of all above-threshold values or alone if all values are BT.

Table IC. Height, Weight, and Body Mass Index

This table contains descriptive statistics and frequencies of VITAL measurements.

	Result	%	Source table
Height measurements			VIT_L3_HT_DIST
Records	6,226,535		
Height (inches), mean	65		
Height (inches), median	66		
Weight measurements			VIT_L3_WT_DIST
Records	11,260,683		
Weight (lbs.), mean	183		
Weight (lbs.), median	180		
Body Mass Index (BMI) measurements			VIT_L3_BMI
Records	6,126,883		
BMI <=25	1,977,436	32.3	
BMI 26-30	1,698,613	27.7	
BMI >=31	2,450,834	40.0	

Cell counts between 1 and the low-cell count threshold are displayed as BT and treated as missing.

Sums which include BT values are marked with an asterisk (*), either accompanying the sum of all above-threshold values or alone if all values are BT.

Table ID. Records, Patients, Encounters, and Date Ranges by Table

This table contains summary counts, date ranges, and distinct encounters where applicable for each characterized table. The total number of records in each table is also displayed. In most cases this will be identical to the number of characterized records, but may differ slightly because of the lookback date restriction.

Table	All Data	Characterized Data			Characterized Data Range			Source Tables
	Records	Records	Patients	Encounters	Field name	5th Percentile	95th Percentile	
DEMOGRAPHIC	966,870	966,870	966,870	---	BIRTH_DATE	1932_Oct	2012_Aug	DEM_L3_N; XTBL_L3_DATES
ENROLLMENT	966,868	966,868	966,868	---	ENR_START_DATE	2010_Jan	2018_Apr	ENR_L3_N; XTBL_L3_DATES
ENCOUNTER	66,662,590	66,662,590	956,583	66,662,590	ADMIT_DATE	2010_Jun	2019_Apr	ENC_L3_N; XTBL_L3_DATES
DIAGNOSIS	98,986,521	98,986,521	934,427	41,417,785	ADMIT_DATE	2010_May	2019_Apr	DIA_L3_N; XTBL_L3_DATES
PROCEDURES	53,609,004	53,609,004	806,224	26,238,704	ADMIT_DATE	2010_Jun	2019_Apr	PRO_L3_N; XTBL_L3_DATES
DEATH	65,774	65,774	65,774	---	DEATH_DATE	2010_Sep	2019_Apr	DEATH_L3_N; XTBL_L3_DATES
VITAL	12,713,955	12,713,955	648,554	12,132,928	MEASURE_DATE	2010_Jun	2019_Jun	VIT_L3_N; XTBL_L3_DATES
LAB_RESULT_CM	225,777,281	225,777,281	797,266	21,702,654	RESULT_DATE	2010_Jul	2019_Jun	LAB_L3_N; XTBL_L3_DATES
PRESCRIBING	77,303,389	77,303,389	796,553	18,130,077	RX_ORDER_DATE	2010_Aug	2019_May	PRES_L3_N; XTBL_L3_DATES
DISPENSING	4,909,610	4,909,610	295,989	---	DISPENSE_DATE	2011_Apr	2019_May	DISP_L3_N; XTBL_L3_DATES
CONDITION	16,272,953	16,272,953	762,745	8,731,161	REPORT_DATE	2010_Jun	2019_Apr	COND_L3_N; XTBL_L3_DATES
DEATH_CAUSE	55,584	55,584	14,982	---	---	---	---	DEATHC_L3_N
PROVIDER	103,996	103,996	---	---	---	---	---	PROV_L3_N
MED_ADMIN	1,041,598	1,041,598	123,029	468,251	MEDADMIN_START_DATE	2015_May	2019_Oct	MEDADM_L3_N; XTBL_L3_DATES
PCORNET_TRIAL	320	320	320	---	---	---	---	TRIAL_L3_N
LDS_ADDRESS_HISTORY	7,479,400	7,479,400	955,110	---	---	---	---	LDSADRS_L3_N
IMMUNIZATION	11,804,769	11,804,769	738,870	1,149,754	VX_RECORD_DATE			IMMUNE_L3_N; XTBL_L3_DATES

The source for the total number of records is the DATAMART_ALL table.

Cell counts between 1 and the low-cell count threshold are displayed as BT and treated as missing.

Records are the sum of ALL_N and NULL_N for the field which is the primary key for the table. Patients are the sum of DISTINCT_N and NULL_N for the PATID field. Encounters are the sum of DISTINCT_N and NULL_N for the ENCOUNTERID field.

Table IE. Records Per Table By Encounter Type

This table contains record counts by encounter type for the ENCOUNTER, DIAGNOSIS, and PROCEDURES tables.

Encounter Type	ENCOUNTER		DIAGNOSIS		PROCEDURES	
	N	%	N	%	N	%
AV (Ambulatory Visit)	42,846,033	64.3	85,292,213	86.2	49,319,785	92.0
ED (Emergency Dept)	1,592,511	2.4	3,742,535	3.8	2,014,832	3.8
EI (ED to IP Stay)	0		0		0	
IC (Institutional Professional Consult)	707,017	1.1	2,721,118	2.7	753,672	1.4
IP (Inpatient Hospital Stay)	823,215	1.2	4,985,450	5.0	1,009,383	1.9
IS (Non-acute Institutional Stay)	95,435	0.1	352,449	0.4	106,380	0.2
OA (Other Ambulatory Visit)	18,997,645	28.5	1,503,780	1.5	200,049	0.4
OS (Observation Stay)	38,726	0.1	388,976	0.4	204,903	0.4
Missing, NI, UN or OT	1,562,008	2.3	0		0	
Total	66,662,590		98,986,521		53,609,004	
Source table	ENC_L3_ENCTYPE		DIA_L3_ENCTYPE		PRO_L3_ENCTTPE	

Cell counts between 1 and the low-cell count threshold are displayed as BT and treated as missing.

Table IG. Lab Results For Selected Lab Tests

This table illustrates the number of records and number of unique patients for 30 high volume data curation lab groups, and the percentage of patients in the ENCOUNTER table who have these results. Although there is not a required relationship between the ENCOUNTER and LAB_RESULT_CM tables, patients with encounters are the most relevant denominator for this table. Version 3.2 of the data curation lab groups includes 490 concepts of interest to the Collaborative Research Groups (CRGs). Groups were constructed based on the LOINC attributes of COMPONENT, SYSTEM, and, if necessary, TIME, METHOD and CLASS. More information about the data curation lab groups is available on the Data Curation home page (<https://pcornet.imeetcentral.com/p/aQAAAAACjjsH>).

DC_LAB_GROUP	Records	Percentage of records in the LAB_RESULT_CM table with a LAB_LOINC code	Patients	Percentage of patients in the ENCOUNTER table	Source tables
ALBUMIN B/S/P	2,710,284	1.2	399,352	41.7	LAB_L3_DCGROUP;ENC_L3_N
ALP TOTAL	2,511,286	1.1	395,244	41.3	LAB_L3_DCGROUP;ENC_L3_N
ALT	2,876,670	1.3	410,473	42.9	LAB_L3_DCGROUP;ENC_L3_N
AST	2,684,271	1.2	403,299	42.2	LAB_L3_DCGROUP;ENC_L3_N
BASOPHILS ABSOLUTE	3,166,479	1.4	456,652	47.7	LAB_L3_DCGROUP;ENC_L3_N
BILIRUBIN TOTAL B/S/P	2,499,048	1.1	406,301	42.5	LAB_L3_DCGROUP;ENC_L3_N
BUN	4,915,476	2.2	473,060	49.5	LAB_L3_DCGROUP;ENC_L3_N
CALCIUM B/S/P	4,831,310	2.1	470,593	49.2	LAB_L3_DCGROUP;ENC_L3_N
CHLORIDE B/S/P	4,699,318	2.1	464,944	48.6	LAB_L3_DCGROUP;ENC_L3_N
CHOLESTEROL-LDL ABSOLUTE	1,700,863	0.8	349,924	36.6	LAB_L3_DCGROUP;ENC_L3_N
CREATININE B/S/P	5,137,140	2.3	481,324	50.3	LAB_L3_DCGROUP;ENC_L3_N
EGFR	4,892,037	2.2	439,496	45.9	LAB_L3_DCGROUP;ENC_L3_N
GLUCOSE B/S/P	6,002,015	2.7	488,369	51.1	LAB_L3_DCGROUP;ENC_L3_N
HEMATOCRIT	4,487,352	2.0	490,011	51.2	LAB_L3_DCGROUP;ENC_L3_N
HEMOGLOBIN A1C	908,143	0.4	147,487	15.4	LAB_L3_DCGROUP;ENC_L3_N
HEMOGLOBIN B/S/P	4,958,971	2.2	516,401	54.0	LAB_L3_DCGROUP;ENC_L3_N
INR	3,034,699	1.3	151,805	15.9	LAB_L3_DCGROUP;ENC_L3_N
LYMPHOCYTES ABSOLUTE	3,528,283	1.6	466,572	48.8	LAB_L3_DCGROUP;ENC_L3_N
MCH	4,169,330	1.8	485,856	50.8	LAB_L3_DCGROUP;ENC_L3_N
MCHC	4,195,905	1.9	485,854	50.8	LAB_L3_DCGROUP;ENC_L3_N
MCV	4,241,694	1.9	485,842	50.8	LAB_L3_DCGROUP;ENC_L3_N
MONOCYTES ABSOLUTE	3,488,874	1.5	466,596	48.8	LAB_L3_DCGROUP;ENC_L3_N
NEUTROPHILS ABSOLUTE	3,961,864	1.8	467,289	48.8	LAB_L3_DCGROUP;ENC_L3_N

Cell counts between 1 and the low-cell count threshold are displayed as BT and treated as missing.

Number of LAB_RESULT_CM records with a LAB_LOINC can be found in Table IVI.

Number of patients in the ENCOUNTER table is the DISTINCT_N for the PATID field in the ENC_L3_N query.

Table IG. Lab Results For Selected Lab Tests (continued)

This table illustrates the number of records and number of unique patients for 30 high volume data curation lab groups, and the percentage of patients in the ENCOUNTER table who have these results. Although there is not a required relationship between the ENCOUNTER and LAB_RESULT_CM tables, patients with encounters are the most relevant denominator for this table. Version 3.2 of the data curation lab groups includes 490 concepts of interest to the Collaborative Research Groups (CRGs). Groups were constructed based on the LOINC attributes of COMPONENT, SYSTEM, and, if necessary, TIME, METHOD and CLASS. More information about the data curation lab groups is available on the Data Curation home page (<https://pcorntet.imeetcentral.com/p/aQAAAAACjjsH>).

DC_LAB_GROUP	Records	Percentage of records in the LAB_RESULT_CM table with a LAB_LOINC code	Patients	Percentage of patients in the ENCOUNTER table	Source tables
PLATELETS	4,555,787	2.0	494,801	51.7	LAB_L3_DCGROUP;ENC_L3_N
POTASSIUM B/S/P	5,135,811	2.3	476,038	49.8	LAB_L3_DCGROUP;ENC_L3_N
RBC B/S/P	4,259,647	1.9	485,860	50.8	LAB_L3_DCGROUP;ENC_L3_N
RDW	4,269,427	1.9	485,601	50.8	LAB_L3_DCGROUP;ENC_L3_N
SODIUM B/S/P	4,935,142	2.2	475,614	49.7	LAB_L3_DCGROUP;ENC_L3_N
TOTAL CO2	4,678,410	2.1	464,256	48.5	LAB_L3_DCGROUP;ENC_L3_N
WBC B/S/P	4,334,585	1.9	490,276	51.3	LAB_L3_DCGROUP;ENC_L3_N

Cell counts between 1 and the low-cell count threshold are displayed as BT and treated as missing.

Number of LAB_RESULT_CM records with a LAB_LOINC can be found in Table IVI.

Number of patients in the ENCOUNTER table is the DISTINCT_N for the PATID field in the ENC_L3_N query.

PCORnet Empirical Data Curation Report Table of Contents

This report is derived from the results of the data curation query. This query package fully characterizes all CDM v5.1 tables. Please refer to the Work Plan for details about the source tables.

Report Run Date: 2020-01-18

Query Run Date: 2020-01-16

Maximum Table Refresh Date: 2019-12-05

Low Cell Count Threshold: 0

Query Package: DC V5.11

Lookback Date: 2000-01-16

CDM Version: 5.1

SAS_ETS licensed: Yes / SAS_ETS installed: No

Section	Table	Table Description	Data Check(s)
n/a	n/a	Data Check Exception Summary	n/a
Section I: Descriptive Information	Table IA	Demographic Summary	n/a
	Table IB	Potential Pools of Patients	3.04, 3.05
	Table IC	Height, Weight, and Body Mass Index (BMI)	n/a
	Table ID	Records, Patients, Encounters, and Date Ranges by Table	n/a
	Table IE	Records Per Table by Encounter Type	n/a
	Table IF	Date Obfuscation or Imputation	n/a
	Table IG	Lab Results For Selected Lab Tests	n/a
	Table IH	Patients with Selected Diagnoses	n/a
	Table II	Patients with Selected Procedures	n/a
	Chart IA	Trend in Vital Measures by Measurement Date, Past 5 Years	n/a
	Chart IB	Trend in Encounters by Admit Date and Encounter Type, Past 5 Years	n/a
	Chart IC	Trend in Institutional Encounters by Discharge Date and Encounter Type, Past 5 Years	n/a
	Chart ID	Trend in Laboratory Results by Result Date, Past 5 Years	n/a
	Chart IE	Trend in Prescribed Medications by Rx Order Date, Past 5 Years	n/a
	Chart IF	Trend in Dispensed Medications by Dispense Date, Past 5 Years	n/a
	Chart IG	Trend in Administered Medications by Start Date, Past 5 Years	n/a
	Chart IH	Trend in Condition Records by Report Date, Past 5 Years	n/a
	Chart II	Trend in Death Records by Death Date and Source, Past 5 Years	n/a
	Chart IJ	Trend in Immunization Records by Vx Record Date, Past 5 Years	n/a

REPORT_RUN_DATE is the date the EDC report was run. QUERY_RUN_DATE is the date the data curation query (xtbl portion if running the programs separately) was run. MAXIMUM_TABLE_REFRESH_DATE is derived from the HARVEST table. The LOW_CELL_COUNT_THRESHOLD is established by the user. The default value is 0; changes to this value are highlighted in orange and will be used to assess compliance with the terms of the PCRF Scope of Work. The LOOKBACK_DATE is the earliest date eligible for inclusion in the query results (see the Work Plan for details). CDM_VERSION is from the HARVEST table. The SAS_ETS module is used for Data Check 2.08.

PCORnet Empirical Data Curation Report Table of Contents (continued - page 2 of 3)

This report is derived from the results of the data curation query. This query package fully characterizes all CDM v5.1 tables. Please refer to the Work Plan for details about the source tables.

Section	Table	Table Description	Data Check(s)
Section II: Data Model Conformance	Table IIA	Primary Key Errors	1.05
	Table IIB	Values Outside of Common Data Model (CDM) Specifications	1.06
	Table IIC	Non-Permissible Missing Values	1.07
	Table IID	Diagnostic Errors	1.01, 1.02, 1.03, 1.04
	Table IIE	Orphan Records, Replication Errors and Encounter Duplication	1.08, 1.09, 1.10, 1.11, 1.12
	Table IIF	Potential Code Errors	1.13
Section III: Data Plausibility	Table IIIA	Future Dates	2.01
	Table IIIB	Records with Extreme Values	2.02
	Table IIIC	Illogical Dates	2.03
	Table IIID	Encounters Per Visit and Per Patient	2.04
	Table IIIE	Laboratory Result Specimen Source Discrepancies	2.05
	Table IIIF	Quantitative Lab Result Outliers, Selected Tests	2.06
	Table IIIG	Monthly Record Volume Outliers, Selected Domains	2.08
	Chart IIIA	Monthly Record Volume Outliers, Encounters	2.08
	Chart IIIB	Monthly Record Volume Outliers, Diagnoses	2.08
	Chart IIIC	Monthly Record Volume Outliers, Procedures	2.08
	Chart IIID	Monthly Record Volume Outliers, Vitals	2.08
	Chart IIIE	Monthly Record Volume Outliers, Prescribing	2.08
	Chart IIIF	Monthly Record Volume Outliers, Labs	2.08

REPORT_RUN_DATE is the date the EDC report was run. QUERY_RUN_DATE is the date the data curation query (xtbl portion if running the programs separately) was run. MAXIMUM_TABLE_REFRESH_DATE is derived from the HARVEST table. The LOW_CELL_COUNT_THRESHOLD is established by the user. The default value is 0; changes to this value are highlighted in orange and will be used to assess compliance with the terms of the PCRF Scope of Work. The LOOKBACK_DATE is the earliest date eligible for inclusion in the query results (see the Work Plan for details). CDM_VERSION is from the HARVEST table. The SAS_ETC module is used for Data Check 2.08.

PCORnet Empirical Data Curation Report Table of Contents (continued - page 3 of 3)

This report is derived from the results of the data curation query. This query package fully characterizes all CDM v5.1 tables. Please refer to the Work Plan for details about the source tables.

Section	Table	Table Description	Data Check(s)	
Section IV: Data Completeness and Plausibility	Table IVA	Diagnosis Records Per Encounter, Overall and by Encounter Type	3.01	
	Chart IVA	Diagnosis Records Per Encounter by Admit Date and Encounter Type, Past 5 Years	n/a	
	Table IVB	Procedure Records Per Encounter, Overall and by Encounter Type	3.02	
	Chart IVB	Procedure Records Per Encounter by Admit Date and Encounter Type, Past 5 Years	n/a	
	Table IVC	Missing or Unknown Values, Required Tables	3.03	
	Table IVD	Missing or Unknown Values, Optional Tables	3.03	
	Table IVE	Principal Diagnoses for Institutional Encounters	2.07, 3.06	
	Table IVF	Data Latency and Completeness of Encounter, Diagnosis and Procedure Data, Past 2 Years	3.07	
	Table IVG	Data Latency and Completeness of Vital, Prescription, and Lab Data, Past 2 Years	3.11	
	Table IVH	RXNORM Term Type Mapping, Overall and Past 5 Years	3.08	
	Table IVI	Laboratory Result Data Completeness, Overall and Past 5 Years	3.09, 3.10, 3.12	
	Table IVI_Ref	Laboratory Result Data Completeness Definitions	n/a	
	Section V: Data Persistence	Table VA	Changes in Tables	4.01
		Table VB	Changes in Selected Encounter Types and Domains	4.02
Table VC		Changes in Selected Code Types	4.03	

REPORT_RUN_DATE is the date the EDC report was run. QUERY_RUN_DATE is the date the data curation query (xtbl portion if running the programs separately) was run. MAXIMUM_TABLE_REFRESH_DATE is derived from the HARVEST table. The LOW_CELL_COUNT_THRESHOLD is established by the user. The default value is 0; changes to this value are highlighted in orange and will be used to assess compliance with the terms of the PCRF Scope of Work. The LOOKBACK_DATE is the earliest date eligible for inclusion in the query results (see the Work Plan for details). CDM_VERSION is from the HARVEST table. The SAS_ETTS module is used for Data Check 2.08.

Data Check Exception Summary

This table summarizes conformance with PCORnet Data Checks v7. Changes to the previous cycle's data checks are explained in the footnotes. Exceptions to required data checks are highlighted in red and must be corrected before returning results. Exceptions to investigative data checks are highlighted in blue and must be investigated and classified in the ETL ADD as an irremediable feature/limitation of source data, remediable, or warrants further investigation. Data check exceptions should not be classified as warrants further investigation for more than 1 data curation cycle. Network-wide results for PCORnet DataMarts approved as of Nov 7, 2019 are presented for comparison; a result of 10.3 means that 10.3% of DataMarts had exceptions to this data check.

Data Check	EDC Table	Data Check Description	Category	Type	Exception?	% DataMarts w/ exceptions
DC 1.01 ^a	Table IID	Required tables are not present	Data Model Conformance	Required	No	0
DC 1.02	Table IID	Required tables are not populated	Data Model Conformance	Required	No	0
DC 1.03 ^a	Table IID	Required fields are not present	Data Model Conformance	Required	No	0
DC 1.04 ^a	Table IID	Required fields do not conform to the data model specifications for data type, length, or name	Data Model Conformance	Required	No	0
DC 1.05 ^a	Table IIA	Tables have primary key definition errors	Data Model Conformance	Required	No	0
DC 1.06 ^d	Table IIB	Required fields contain values outside of data model specifications	Data Model Conformance	Required	No	0
DC 1.07 ^a	Table IIC	Required fields have non-permissible missing values	Data Model Conformance	Required	No	0
DC 1.08 ^a	Table IIE	Tables contain orphan PATIDs	Data Model Conformance	Required	No	0
DC 1.09 ^b	Table IIE	Tables contain orphan ENCOUNTERIDs	Data Model Conformance	Required	No	0
DC 1.10	Table IIE	Replication errors between the ENCOUNTER, PROCEDURES and DIAGNOSIS tables	Data Model Conformance	Required	No	0
DC 1.11 ^b	Table IIE	More than 5% of encounters are assigned to more than one patient	Data Model Conformance	Required	No	0
DC 1.12 ^b	Table IIE	Tables contain orphan PROVIDERIDs	Data Model Conformance	Required	No	0
DC 1.13 ^c	Table IIF	More than 5% of ICD, CPT, LOINC, RXCUI, or NDC codes do not conform to the expected length or content	Data Model Conformance	Required	No	0
DC 2.01 ^d	Table IIIA	More than 5% of records have future dates	Data Plausibility	Investigative	No	7.3
DC 2.02	Table IIIB	More than 10% of records fall into the lowest or highest categories of age, height, weight, diastolic blood pressure, systolic blood pressure or dispensed days supply	Data Plausibility	Investigative	No	8.7
DC 2.03 ^e	Table IIIC	More than 5% of patients have illogical date relationships	Data Plausibility	Investigative	No	22
DC 2.04	Table IIID	The average number of encounters per visit is > 2.0 for inpatient (IP), emergency department (ED), or ED to inpatient (EI) encounters	Data Plausibility	Investigative	No	0
DC 2.05	Table IIIE	More than 5% of results for selected laboratory tests do not have the appropriate specimen source	Data Plausibility	Investigative	No	5.3

(a) Added HASH_TOKEN, IMMUNIZATION and LDS_ADDRESS_HISTORY (b) Added IMMUNIZATION (c) Added IMMUNIZATION and CONDITION; ICD codes in OBS_GEN; NDC codes in PROCEDURES; and changed heuristic for ICD10 procedure codes to allow codes with 3-7 alphanumeric characters instead of the 7 required for billing. (d) Added IMMUNIZATION and LDS_ADDRESS_HISTORY (e) Added VX_RECORD_DATE and revised the logic for the PX_DATE comparison (f) New (g) Added VX_RECORD_DATE, DISPENSE_SOURCE, LAB_RESULT_SOURCE, PRO_SOURCE, VX_SOURCE (h) Added 5 year results (i) Added 5 year results and added additional "flavors of null" for SPECIMEN_SOURCE (SUB, SMPLS, and SPECIMEN) (j) Added Other Ambulatory encounters (k) Added distinct code count

Data Check Exception Summary (continued - page 2 of 4)

This table summarizes conformance with PCORnet Data Checks v7. Changes to the previous cycle's data checks are explained in the footnotes. Exceptions to required data checks are highlighted in red and must be corrected before returning results. Exceptions to investigative data checks are highlighted in blue and must be investigated and classified in the ETL ADD as an irremediable feature/limitation of source data, remediable, or warrants further investigation. Data check exceptions should not be classified as warrants further investigation for more than 1 data curation cycle. Network-wide results for PCORnet DataMarts approved as of Nov 7, 2019 are presented for comparison; a result of 10.3 means that 10.3% of DataMarts had exceptions to this data check.

Data Check	EDC Table	Data Check Description	Category	Type	Exception?	% DataMarts w/ exceptions
DC 2.06	Table IIIF	Median lab result values for selected laboratory tests are statistical or clinical outliers	Data Plausibility	Investigative	Yes	70
DC 2.07	Table IVE	The average number of principal diagnoses per encounter is above threshold [2.0 for inpatient (IP) and ED to inpatient (EI)]	Data Plausibility	Investigative	No	21
DC 2.08 ^f	Table IIIG	The monthly volume of encounter, diagnosis, procedure, vital, prescribing, or laboratory records is an outlier.	Data Plausibility	Investigative	No	42
DC 3.01	Table IVA	The average number of diagnoses records with known diagnosis types per encounter is below threshold [1.0 for ambulatory (AV), inpatient (IP), emergency department (ED), or ED to inpatient (EI) encounters]	Data Completeness	Investigative	No	4.3
DC 3.02	Table IVB	The average number of procedure records with known procedure types per encounter is below threshold [0.75 for ambulatory (AV) encounters, 0.75 for emergency department (ED) encounters, 1.00 for ED to inpatient (EI) encounters, and 1.00 for inpatient (IP) encounters]	Data Completeness	Investigative	No	15

(a) Added HASH_TOKEN, IMMUNIZATION and LDS_ADDRESS_HISTORY (b) Added IMMUNIZATION (c) Added IMMUNIZATION and CONDITION; ICD codes in OBS_GEN; NDC codes in PROCEDURES; and changed heuristic for ICD10 procedure codes to allow codes with 3-7 alphanumeric characters instead of the 7 required for billing. (d) Added IMMUNIZATION and LDS_ADDRESS_HISTORY (e) Added VX_RECORD_DATE and revised the logic for the PX_DATE comparison (f) New (g) Added VX_RECORD_DATE, DISPENSE_SOURCE, LAB_RESULT_SOURCE, PRO_SOURCE, VX_SOURCE (h) Added 5 year results (i) Added 5 year results and added additional “flavors of null” for SPECIMEN_SOURCE (SUB, SMPLS, and SPECIMEN) (j) Added Other Ambulatory encounters (k) Added distinct code count

Data Check Exception Summary (continued - page 3 of 4)

This table summarizes conformance with PCORnet Data Checks v7. Changes to the previous cycle's data checks are explained in the footnotes. Exceptions to required data checks are highlighted in red and must be corrected before returning results. Exceptions to investigative data checks are highlighted in blue and must be investigated and classified in the ETL ADD as an irremediable feature/limitation of source data, remediable, or warrants further investigation. Data check exceptions should not be classified as warrants further investigation for more than 1 data curation cycle. Network-wide results for PCORnet DataMarts approved as of Nov 7, 2019 are presented for comparison; a result of 10.3 means that 10.3% of DataMarts had exceptions to this data check.

Data Check	EDC Table	Data Check Description	Category	Type	Exception?	% DataMarts w/ exceptions
DC 3.03 ^g	Table IVC	More than 10% of records have missing or unknown values for the following fields: BIRTH_DATE, SEX, DISCHARGE_DISPOSITION (IP/EI encounters only), DISCHARGE_DATE (IP/EI encounters only), PX_DATE, RX_ORDER_DATE, DISPENSE_SUP, DX_ORIGIN, PX_SOURCE, VITAL_SOURCE, DEATH_SOURCE, CONDITION_SOURCE, RX_SOURCE, MEDADMIN_SOURCE, DIAGNOSIS.ENCOUNTERID, PROCEDURES.ENCOUNTERID, VITAL.ENCOUNTERID, MED_ADMIN.ENCOUNTERID, MED_ADMIN_CODE, MED_ADMIN_TYPE, OBS_CLIN_CODE, OBS_CLIN_TYPE, OBS_GEN_CODE, OBS_GEN_TYPE	Data Completeness	Investigative	Yes	74
DC 3.03 ^g	Table IVD	More than 10% of records have missing or unknown values for the following fields: BIRTH_DATE, SEX, DISCHARGE_DISPOSITION (IP/EI encounters only), DISCHARGE_DATE (IP/EI encounters only), PX_DATE, RX_ORDER_DATE, VX_RECORD_DATE, DISPENSE_SUP, DEATH_SOURCE, DISPENSE_SOURCE, CONDITION_SOURCE, DX_ORIGIN, LAB_RESULT_SOURCE, MEDADMIN_SOURCE, PRO_SOURCE, PX_SOURCE, RX_SOURCE, VITAL_SOURCE, VX_SOURCE, DIAGNOSIS.ENCOUNTERID, PROCEDURES.ENCOUNTERID, VITAL.ENCOUNTERID, MED_ADMIN.ENCOUNTERID, LAB_RESULT_CM.ENCOUNTERID, MED_ADMIN_CODE, MED_ADMIN_TYPE, OBS_CLIN_CODE, OBS_CLIN_TYPE, OBS_GEN_CODE, OBS_GEN_TYPE.	Data Completeness	Investigative	Yes	74
DC 3.04	Table IB	Less than 50% of patients with encounters have DIAGNOSIS records	Data Completeness	Required	No	0
DC 3.05	Table IB	Less than 50% of patients with encounters have PROCEDURES records	Data Completeness	Required	No	0
DC 3.06	Table IVE	More than 10% of IP (inpatient) or ED to inpatient (EI) encounters with any diagnosis don't have a principal diagnosis	Data Completeness	Investigative	Yes	24
DC 3.07	Table IVF	Encounters, diagnoses or procedures in an ambulatory (AV), emergency department (ED), ED to inpatient (EI), or inpatient (IP) setting are less than 75% complete three months prior to the current month	Data Completeness	Investigative	No	40
DC 3.08 ^h	Table IVH	Less than 80% of prescribing orders are mapped to a RXNORM_CUI which fully specifies the ingredient, strength and dose form	Data Completeness	Investigative	No	19

(a) Added HASH_TOKEN, IMMUNIZATION and LDS_ADDRESS_HISTORY (b) Added IMMUNIZATION (c) Added IMMUNIZATION and CONDITION; ICD codes in OBS_GEN; NDC codes in PROCEDURES; and changed heuristic for ICD10 procedure codes to allow codes with 3-7 alphanumeric characters instead of the 7 required for billing. (d) Added IMMUNIZATION and LDS_ADDRESS_HISTORY (e) Added VX_RECORD_DATE and revised the logic for the PX_DATE comparison (f) New (g) Added VX_RECORD_DATE, DISPENSE_SOURCE, LAB_RESULT_SOURCE, PRO_SOURCE, VX_SOURCE (h) Added 5 year results (i) Added 5 year results and added additional "flavors of null" for SPECIMEN_SOURCE (SUB, SMPLS, and SPECIMEN) (j) Added Other Ambulatory encounters (k) Added distinct code count

Data Check Exception Summary (continued - page 4 of 4)

This table summarizes conformance with PCORnet Data Checks v7. Changes to the previous cycle's data checks are explained in the footnotes. Exceptions to required data checks are highlighted in red and must be corrected before returning results. Exceptions to investigative data checks are highlighted in blue and must be investigated and classified in the ETL ADD as an irremediable feature/limitation of source data, remediable, or warrants further investigation. Data check exceptions should not be classified as warrants further investigation for more than 1 data curation cycle. Network-wide results for PCORnet DataMarts approved as of Nov 7, 2019 are presented for comparison; a result of 10.3 means that 10.3% of DataMarts had exceptions to this data check.

Data Check	EDC Table	Data Check Description	Category	Type	Exception?	% DataMarts w/ exceptions
DC 3.09	Table IVI	Less than 80% of laboratory results are mapped to LAB_LOINC	Data Completeness	Investigative	No	7.6
DC 3.10 ^h	Table IVI	Less than 80% of quantitative results for tests mapped to LAB_LOINC fully specify the normal range	Data Completeness	Investigative	Yes	50
DC 3.11	Table IVG	Vital, prescribing, or laboratory records are less than 75% complete three months prior to the current month.	Data Completeness	Investigative	No	43
DC 3.12 ⁱ	Table IVI	Less than 80% of quantitative results for tests mapped to LAB_LOINC fully specify the specimen source and result unit	Data Completeness	Investigative	Yes	64
DC 4.01 ^a	Table VA	More than a 5% decrease in the number of patients or records in a CDM table	Data Persistence	Investigative	Yes	20
DC 4.02 ^j	Table VB	More than a 5% decrease in the number of patients or records for diagnosis, procedures, labs or prescriptions during an ambulatory (AV), other ambulatory (OA), emergency department (ED), or inpatient (IP) encounter.	Data Persistence	Investigative	Yes	12
DC 4.03 ^k	Table VC	More than a 5% decrease in the number of records or distinct codes for ICD9 or ICD10 diagnosis or procedure codes or CPT/HCPCS procedure codes.	Data Persistence	Investigative	Yes	17

(a) Added HASH_TOKEN, IMMUNIZATION and LDS_ADDRESS_HISTORY (b) Added IMMUNIZATION (c) Added IMMUNIZATION and CONDITION; ICD codes in OBS_GEN; NDC codes in PROCEDURES; and changed heuristic for ICD10 procedure codes to allow codes with 3-7 alphanumeric characters instead of the 7 required for billing. (d) Added IMMUNIZATION and LDS_ADDRESS_HISTORY (e) Added VX_RECORD_DATE and revised the logic for the PX_DATE comparison (f) New (g) Added VX_RECORD_DATE, DISPENSE_SOURCE, LAB_RESULT_SOURCE, PRO_SOURCE, VX_SOURCE (h) Added 5 year results (i) Added 5 year results and added additional “flavors of null” for SPECIMEN_SOURCE (SUB, SMPLS, and SPECIMEN) (j) Added Other Ambulatory encounters (k) Added distinct code count

Table IA. Demographic Summary

This table contains general descriptive information about the patients in the DEMOGRAPHIC table. These patients may or may not be represented in other CDM tables.

	N	%	Source table
Patients	966,870		DEM_L3_N
Age			DEM_L3_AGEYRSDIST1
Mean	45		
Median	45		
Age group			DEM_L3_AGEYRSDIST2
0-4	27,956	2.9	
5-14	99,826	10.3	
15-21	74,694	7.8	
22-64	518,621	53.6	
65+	245,773	25.4	
Missing	0	0.0	
Hispanic			DEM_L3_HISPDIST
N (No)	644,649	66.7	
Y (Yes)	20,064	2.1	
Missing or Refused	302,157	31.2	
Sex			DEM_L3_SEXDIST
F (Female)	491,410	50.8	
M (Male)	475,459	49.2	
Missing or Ambiguous	1	0.0	
Race			DEM_L3_RACEDIST
White	636,931	65.9	
Non-White	35,140	3.7	
Missing or Refused	294,799	30.5	

The four 'flavors of null' defined in the CDM are combined in the missing categories shown here, but details are available in the source tables.

Cell counts between 1 and the low-cell count threshold are displayed as BT and treated as missing.

Sums which include BT values are marked with an asterisk (*), either accompanying the sum of all above-threshold values or alone if all values are BT.

Age is calculated as current age or age at death if death date is known. If multiple death records exist, the earlier death date is used.

Percentages are sums of the percentages in the source table and are subject to rounding errors.

Race distribution is shown for all patients and for patients seen after Meaningful Use standards were widely implemented.

Table IA. Demographic Summary (continued)

This table contains general descriptive information about the patients in the DEMOGRAPHIC table. These patients may or may not be represented in other CDM tables.

	N	%	Source table
Race among patients with at least 1 encounter after December 2011			XTBL_L3_RACE_ENC
White	603,197	69.0	
Non-White	33,198	3.8	
Missing or Refused	237,988	27.2	
Gender Identity			DEM_L3_GENDERDIST
GQ (Genderqueer/Non-Binary)	0	0.0	
M (Man)	0	0.0	
W (Woman)	0	0.0	
MU (Multiple gender categories), SE (Something else), TF (Transgender female/Trans woman/Male-to-female), or TM (Transgender male/Trans man/Female-to-male)	0	0.0	
Missing or Refused	966,870	100.0	
Sexual Orientation			DEM_L3_ORIENTDIST
Bisexual	0	0.0	
Gay	0	0.0	
Lesbian	0	0.0	
Queer	0	0.0	
Straight	0	0.0	
AS (Asexual), MU (Multiple sexual orientations), SE (Something else), QS (Questioning)	0	0.0	
Missing or Refused	966,870	100.0	

The four 'flavors of null' defined in the CDM are combined in the missing categories shown here, but details are available in the source tables.

Cell counts between 1 and the low-cell count threshold are displayed as BT and treated as missing.

Sums which include BT values are marked with an asterisk (*), either accompanying the sum of all above-threshold values or alone if all values are BT.

Age is calculated as current age or age at death if death date is known. If multiple death records exist, the earlier death date is used.

Percentages are sums of the percentages in the source table and are subject to rounding errors.

Race distribution is shown for all patients and for patients seen after Meaningful Use standards were widely implemented.

Table IB. Potential Pools of Patients

This table illustrates the number of patients meeting different inclusion criteria and supports Data Check 3.04 (less than 50% of patients with encounters have DIAGNOSIS records) and Data Check 3.05 (less than 50% of patients with encounters have PROCEDURES records). Data check exceptions are highlighted in red and must be corrected.

Metric	Metric Description	Result	Source table
Potential pool of patients for observational studies	Number of unique patients with at least 1 ED, EI, IP, OS, or AV encounter within the past 5 years	679,793	ENC_L3_DASH2
Potential pool of patients for trials	Number of unique patients with at least 1 ED, EI, IP, OS, or AV encounter within the past 1 year	341,359	ENC_L3_DASH2
Potential pool of patients for studies requiring data on diagnoses, vital measures and (a) medications or (b) medications and lab results	Number of unique patients with at least 1 encounter and DIAGNOSIS and VITAL records within the past 5 years	499,198	XTBL_L3_DASH1
	Number of unique patients with at least 1 encounter and DIAGNOSIS, VITAL, and PRESCRIBING or DISPENSING records within the past 5 years	461,154	XTBL_L3_DASH2
	Number of unique patients with at least 1 encounter and DIAGNOSIS, VITAL, PRESCRIBING or DISPENSING, and LAB_RESULT_CM records within the past 5 years	460,216	XTBL_L3_DASH3
Patients with diagnosis data	Percentage of patients with encounters who have at least 1 diagnosis	98%	ENC_L3_N; DIA_L3_N
Patients with procedure data	Percentage of patients with encounters who have at least 1 procedure	84%	ENC_L3_N; PRO_L3_N

Cell counts between 1 and the low-cell count threshold are displayed as BT and treated as missing.

Sums which include BT values are marked with an asterisk (*), either accompanying the sum of all above-threshold values or alone if all values are BT.

Table IC. Height, Weight, and Body Mass Index

This table contains descriptive statistics and frequencies of VITAL measurements.

	Result	%	Source table
Height measurements			VIT_L3_HT_DIST
Records	6,226,535		
Height (inches), mean	65		
Height (inches), median	66		
Weight measurements			VIT_L3_WT_DIST
Records	11,260,683		
Weight (lbs.), mean	183		
Weight (lbs.), median	180		
Body Mass Index (BMI) measurements			VIT_L3_BMI
Records	6,126,883		
BMI <=25	1,977,436	32.3	
BMI 26-30	1,698,613	27.7	
BMI >=31	2,450,834	40.0	

Cell counts between 1 and the low-cell count threshold are displayed as BT and treated as missing.

Sums which include BT values are marked with an asterisk (*), either accompanying the sum of all above-threshold values or alone if all values are BT.

Table ID. Records, Patients, Encounters, and Date Ranges by Table

This table contains summary counts, date ranges, and distinct encounters where applicable for each characterized table. The total number of records in each table is also displayed. In most cases this will be identical to the number of characterized records, but may differ slightly because of the lookback date restriction.

Table	All Data	Characterized Data			Characterized Data Range			Source Tables
	Records	Records	Patients	Encounters	Field name	5th Percentile	95th Percentile	
DEMOGRAPHIC	966,870	966,870	966,870	---	BIRTH_DATE	1932_Oct	2012_Aug	DEM_L3_N; XTBL_L3_DATES
ENROLLMENT	966,868	966,868	966,868	---	ENR_START_DATE	2010_Jan	2018_Apr	ENR_L3_N; XTBL_L3_DATES
ENCOUNTER	66,662,590	66,662,590	956,583	66,662,590	ADMIT_DATE	2010_Jun	2019_Apr	ENC_L3_N; XTBL_L3_DATES
DIAGNOSIS	98,986,521	98,986,521	934,427	41,417,785	ADMIT_DATE	2010_May	2019_Apr	DIA_L3_N; XTBL_L3_DATES
PROCEDURES	53,609,004	53,609,004	806,224	26,238,704	ADMIT_DATE	2010_Jun	2019_Apr	PRO_L3_N; XTBL_L3_DATES
DEATH	65,774	65,774	65,774	---	DEATH_DATE	2010_Sep	2019_Apr	DEATH_L3_N; XTBL_L3_DATES
VITAL	12,713,955	12,713,955	648,554	12,132,928	MEASURE_DATE	2010_Jun	2019_Jun	VIT_L3_N; XTBL_L3_DATES
LAB_RESULT_CM	225,777,281	225,777,281	797,266	21,702,654	RESULT_DATE	2010_Jul	2019_Jun	LAB_L3_N; XTBL_L3_DATES
PRESCRIBING	77,303,389	77,303,389	796,553	18,130,077	RX_ORDER_DATE	2010_Aug	2019_May	PRES_L3_N; XTBL_L3_DATES
DISPENSING	4,909,610	4,909,610	295,989	---	DISPENSE_DATE	2011_Apr	2019_May	DISP_L3_N; XTBL_L3_DATES
CONDITION	16,272,953	16,272,953	762,745	8,731,161	REPORT_DATE	2010_Jun	2019_Apr	COND_L3_N; XTBL_L3_DATES
DEATH_CAUSE	55,584	55,584	14,982	---	---	---	---	DEATHC_L3_N
PROVIDER	103,996	103,996	---	---	---	---	---	PROV_L3_N
MED_ADMIN	1,041,598	1,041,598	123,029	468,251	MEDADMIN_START_DATE	2015_May	2019_Oct	MEDADM_L3_N; XTBL_L3_DATES
PCORNET_TRIAL	320	320	320	---	---	---	---	TRIAL_L3_N
LDS_ADDRESS_HISTORY	7,479,400	7,479,400	955,110	---	---	---	---	LDSADRS_L3_N
IMMUNIZATION	11,804,769	11,804,769	738,870	1,149,754	VX_RECORD_DATE			IMMUNE_L3_N; XTBL_L3_DATES

The source for the total number of records is the DATAMART_ALL table.

Cell counts between 1 and the low-cell count threshold are displayed as BT and treated as missing.

Records are the sum of ALL_N and NULL_N for the field which is the primary key for the table. Patients are the sum of DISTINCT_N and NULL_N for the PATID field. Encounters are the sum of DISTINCT_N and NULL_N for the ENCOUNTERID field.

Table IE. Records Per Table By Encounter Type

This table contains record counts by encounter type for the ENCOUNTER, DIAGNOSIS, and PROCEDURES tables.

Encounter Type	ENCOUNTER		DIAGNOSIS		PROCEDURES	
	N	%	N	%	N	%
AV (Ambulatory Visit)	42,846,033	64.3	85,292,213	86.2	49,319,785	92.0
ED (Emergency Dept)	1,592,511	2.4	3,742,535	3.8	2,014,832	3.8
EI (ED to IP Stay)	0		0		0	
IC (Institutional Professional Consult)	707,017	1.1	2,721,118	2.7	753,672	1.4
IP (Inpatient Hospital Stay)	823,215	1.2	4,985,450	5.0	1,009,383	1.9
IS (Non-acute Institutional Stay)	95,435	0.1	352,449	0.4	106,380	0.2
OA (Other Ambulatory Visit)	18,997,645	28.5	1,503,780	1.5	200,049	0.4
OS (Observation Stay)	38,726	0.1	388,976	0.4	204,903	0.4
Missing, NI, UN or OT	1,562,008	2.3	0		0	
Total	66,662,590		98,986,521		53,609,004	
Source table	ENC_L3_ENCTYPE		DIA_L3_ENCTYPE		PRO_L3_ENCTTPE	

Cell counts between 1 and the low-cell count threshold are displayed as BT and treated as missing.

Table IF. Date Obfuscation or Imputation

This table contains information about the presence of date obfuscation or imputation. Imputed or obfuscated dates are important to consider when interpreting results. The use of date obfuscation is highlighted in orange and will be used to assess compliance with the terms of the PCRF Scope of Work.

Table	Field	DATE_MGMT	Source table
HARVEST	BIRTH_DATE_MGMT	01 (No imputation or obfuscation)	XTBL_L3_METADATA
HARVEST	ENR_START_DATE_MGMT	01 (No imputation or obfuscation)	XTBL_L3_METADATA
HARVEST	ENR_END_DATE_MGMT	01 (No imputation or obfuscation)	XTBL_L3_METADATA
HARVEST	ADMIT_DATE_MGMT	01 (No imputation or obfuscation)	XTBL_L3_METADATA
HARVEST	DISCHARGE_DATE_MGMT	01 (No imputation or obfuscation)	XTBL_L3_METADATA
HARVEST	PX_DATE_MGMT	01 (No imputation or obfuscation)	XTBL_L3_METADATA
HARVEST	RX_ORDER_DATE_MGMT	01 (No imputation or obfuscation)	XTBL_L3_METADATA
HARVEST	RX_START_DATE_MGMT	01 (No imputation or obfuscation)	XTBL_L3_METADATA
HARVEST	RX_END_DATE_MGMT	01 (No imputation or obfuscation)	XTBL_L3_METADATA
HARVEST	DISPENSE_DATE_MGMT	01 (No imputation or obfuscation)	XTBL_L3_METADATA
HARVEST	LAB_ORDER_DATE_MGMT	01 (No imputation or obfuscation)	XTBL_L3_METADATA
HARVEST	SPECIMEN_DATE_MGMT	01 (No imputation or obfuscation)	XTBL_L3_METADATA
HARVEST	RESULT_DATE_MGMT	01 (No imputation or obfuscation)	XTBL_L3_METADATA
HARVEST	MEASURE_DATE_MGMT	01 (No imputation or obfuscation)	XTBL_L3_METADATA
HARVEST	ONSET_DATE_MGMT	01 (No imputation or obfuscation)	XTBL_L3_METADATA
HARVEST	REPORT_DATE_MGMT	01 (No imputation or obfuscation)	XTBL_L3_METADATA
HARVEST	RESOLVE_DATE_MGMT	01 (No imputation or obfuscation)	XTBL_L3_METADATA
HARVEST	PRO_DATE_MGMT	01 (No imputation or obfuscation)	XTBL_L3_METADATA
HARVEST	DEATH_DATE_MGMT	01 (No imputation or obfuscation)	XTBL_L3_METADATA
HARVEST	MEDADMIN_START_DATE_MGMT	01 (No imputation or obfuscation)	XTBL_L3_METADATA
HARVEST	MEDADMIN_STOP_DATE_MGMT	01 (No imputation or obfuscation)	XTBL_L3_METADATA
HARVEST	OBSCLIN_DATE_MGMT	01 (No imputation or obfuscation)	XTBL_L3_METADATA
HARVEST	OBSGEN_DATE_MGMT	01 (No imputation or obfuscation)	XTBL_L3_METADATA
HARVEST	DX_DATE_MGMT	01 (No imputation or obfuscation)	XTBL_L3_METADATA
HARVEST	ADDRESS_PERIOD_START_MGMT	01 (No imputation or obfuscation)	XTBL_L3_METADATA
HARVEST	ADDRESS_PERIOD_END_MGMT	01 (No imputation or obfuscation)	XTBL_L3_METADATA
HARVEST	VX_RECORD_DATE_MGMT	01 (No imputation or obfuscation)	XTBL_L3_METADATA
HARVEST	VX_ADMIN_DATE_MGMT	01 (No imputation or obfuscation)	XTBL_L3_METADATA
HARVEST	VX_EXP_DATE_MGMT	01 (No imputation or obfuscation)	XTBL_L3_METADATA

Details regarding the type of obfuscation or imputation are available in the source data.

Table IG. Lab Results For Selected Lab Tests

This table illustrates the number of records and number of unique patients for 30 high volume data curation lab groups, and the percentage of patients in the ENCOUNTER table who have these results. Although there is not a required relationship between the ENCOUNTER and LAB_RESULT_CM tables, patients with encounters are the most relevant denominator for this table. Version 3.2 of the data curation lab groups includes 490 concepts of interest to the Collaborative Research Groups (CRGs). Groups were constructed based on the LOINC attributes of COMPONENT, SYSTEM, and, if necessary, TIME, METHOD and CLASS. More information about the data curation lab groups is available on the Data Curation home page (<https://pcornet.imeetcentral.com/p/aQAAAAACjjsH>).

DC_LAB_GROUP	Records	Percentage of records in the LAB_RESULT_CM table with a LAB_LOINC code	Patients	Percentage of patients in the ENCOUNTER table	Source tables
ALBUMIN B/S/P	2,710,284	1.2	399,352	41.7	LAB_L3_DCGROUP;ENC_L3_N
ALP TOTAL	2,511,286	1.1	395,244	41.3	LAB_L3_DCGROUP;ENC_L3_N
ALT	2,876,670	1.3	410,473	42.9	LAB_L3_DCGROUP;ENC_L3_N
AST	2,684,271	1.2	403,299	42.2	LAB_L3_DCGROUP;ENC_L3_N
BASOPHILS ABSOLUTE	3,166,479	1.4	456,652	47.7	LAB_L3_DCGROUP;ENC_L3_N
BILIRUBIN TOTAL B/S/P	2,499,048	1.1	406,301	42.5	LAB_L3_DCGROUP;ENC_L3_N
BUN	4,915,476	2.2	473,060	49.5	LAB_L3_DCGROUP;ENC_L3_N
CALCIUM B/S/P	4,831,310	2.1	470,593	49.2	LAB_L3_DCGROUP;ENC_L3_N
CHLORIDE B/S/P	4,699,318	2.1	464,944	48.6	LAB_L3_DCGROUP;ENC_L3_N
CHOLESTEROL-LDL ABSOLUTE	1,700,863	0.8	349,924	36.6	LAB_L3_DCGROUP;ENC_L3_N
CREATININE B/S/P	5,137,140	2.3	481,324	50.3	LAB_L3_DCGROUP;ENC_L3_N
EGFR	4,892,037	2.2	439,496	45.9	LAB_L3_DCGROUP;ENC_L3_N
GLUCOSE B/S/P	6,002,015	2.7	488,369	51.1	LAB_L3_DCGROUP;ENC_L3_N
HEMATOCRIT	4,487,352	2.0	490,011	51.2	LAB_L3_DCGROUP;ENC_L3_N
HEMOGLOBIN A1C	908,143	0.4	147,487	15.4	LAB_L3_DCGROUP;ENC_L3_N
HEMOGLOBIN B/S/P	4,958,971	2.2	516,401	54.0	LAB_L3_DCGROUP;ENC_L3_N
INR	3,034,699	1.3	151,805	15.9	LAB_L3_DCGROUP;ENC_L3_N
LYMPHOCYTES ABSOLUTE	3,528,283	1.6	466,572	48.8	LAB_L3_DCGROUP;ENC_L3_N
MCH	4,169,330	1.8	485,856	50.8	LAB_L3_DCGROUP;ENC_L3_N
MCHC	4,195,905	1.9	485,854	50.8	LAB_L3_DCGROUP;ENC_L3_N
MCV	4,241,694	1.9	485,842	50.8	LAB_L3_DCGROUP;ENC_L3_N
MONOCYTES ABSOLUTE	3,488,874	1.5	466,596	48.8	LAB_L3_DCGROUP;ENC_L3_N
NEUTROPHILS ABSOLUTE	3,961,864	1.8	467,289	48.8	LAB_L3_DCGROUP;ENC_L3_N

Cell counts between 1 and the low-cell count threshold are displayed as BT and treated as missing.

Number of LAB_RESULT_CM records with a LAB_LOINC can be found in Table IVI.

Number of patients in the ENCOUNTER table is the DISTINCT_N for the PATID field in the ENC_L3_N query.

Table IG. Lab Results For Selected Lab Tests (continued)

This table illustrates the number of records and number of unique patients for 30 high volume data curation lab groups, and the percentage of patients in the ENCOUNTER table who have these results. Although there is not a required relationship between the ENCOUNTER and LAB_RESULT_CM tables, patients with encounters are the most relevant denominator for this table. Version 3.2 of the data curation lab groups includes 490 concepts of interest to the Collaborative Research Groups (CRGs). Groups were constructed based on the LOINC attributes of COMPONENT, SYSTEM, and, if necessary, TIME, METHOD and CLASS. More information about the data curation lab groups is available on the Data Curation home page (<https://pcorntet.imeetcentral.com/p/aQAAAAACjjsH>).

DC_LAB_GROUP	Records	Percentage of records in the LAB_RESULT_CM table with a LAB_LOINC code	Patients	Percentage of patients in the ENCOUNTER table	Source tables
PLATELETS	4,555,787	2.0	494,801	51.7	LAB_L3_DCGROUP;ENC_L3_N
POTASSIUM B/S/P	5,135,811	2.3	476,038	49.8	LAB_L3_DCGROUP;ENC_L3_N
RBC B/S/P	4,259,647	1.9	485,860	50.8	LAB_L3_DCGROUP;ENC_L3_N
RDW	4,269,427	1.9	485,601	50.8	LAB_L3_DCGROUP;ENC_L3_N
SODIUM B/S/P	4,935,142	2.2	475,614	49.7	LAB_L3_DCGROUP;ENC_L3_N
TOTAL CO2	4,678,410	2.1	464,256	48.5	LAB_L3_DCGROUP;ENC_L3_N
WBC B/S/P	4,334,585	1.9	490,276	51.3	LAB_L3_DCGROUP;ENC_L3_N

Cell counts between 1 and the low-cell count threshold are displayed as BT and treated as missing.

Number of LAB_RESULT_CM records with a LAB_LOINC can be found in Table IVI.

Number of patients in the ENCOUNTER table is the DISTINCT_N for the PATID field in the ENC_L3_N query.

Table IH. Patients with Selected Diagnoses

This table illustrates the number of unique patients for 15 sentinel diagnoses, and the percentage of patients in the ENCOUNTER table who have these diagnoses. Diagnosis groups were defined using AHRQ's Clinical Classification Software (<https://www.hcup-us.ahrq.gov/toolssoftware/ccs/ccs.jsp>) for ICD9 and ICD10 diagnosis codes. These 15 diagnoses represent autoimmune diseases, cardiac diseases, diabetes, obesity, and conditions often diagnosed in childhood. These diagnoses are expected to be represented in most DataMarts.

DC_DX_GROUP	Patients	Percentage of patients in the ENCOUNTER table	Source tables
Acute myocardial infarction [CCS 100]	16,131	1.7	DIA_L3_DCGROUP;ENC_L3_N
Asthma [CCS 128]	71,098	7.4	DIA_L3_DCGROUP;ENC_L3_N
Attention-deficit conduct and disruptive behavior disorders [CCS 652]	48,370	5.1	DIA_L3_DCGROUP;ENC_L3_N
Cardiac dysrhythmias [CCS 106]	123,400	12.9	DIA_L3_DCGROUP;ENC_L3_N
Congestive heart failure; nonhypertensive [CCS 108]	39,933	4.2	DIA_L3_DCGROUP;ENC_L3_N
Cystic fibrosis [CCS 56]	184	0.0	DIA_L3_DCGROUP;ENC_L3_N
Developmental disorders [CCS 654]	25,635	2.7	DIA_L3_DCGROUP;ENC_L3_N
Diabetes mellitus with complications [CCS 50]	53,041	5.5	DIA_L3_DCGROUP;ENC_L3_N
Diabetes mellitus without complication [CCS 49]	151,411	15.8	DIA_L3_DCGROUP;ENC_L3_N
Disorders of lipid metabolism [CCS 53]	215,993	22.6	DIA_L3_DCGROUP;ENC_L3_N
Hypertension with complications and secondary hypertension [CCS 99]	44,420	4.6	DIA_L3_DCGROUP;ENC_L3_N
Multiple sclerosis [CCS 80]	3,034	0.3	DIA_L3_DCGROUP;ENC_L3_N
Otitis media and related conditions [CCS 92]	130,653	13.7	DIA_L3_DCGROUP;ENC_L3_N
Regional enteritis and ulcerative colitis [CCS 144]	6,494	0.7	DIA_L3_DCGROUP;ENC_L3_N
Rheumatoid arthritis and related disease [CCS 202]	11,936	1.2	DIA_L3_DCGROUP;ENC_L3_N

Cell counts between 1 and the low-cell count threshold are displayed as BT and treated as missing.

Number of patients in the ENCOUNTER table is the DISTINCT_N for the PATID field in the ENC_L3_N query.

Table II. Patients with Selected Procedures

This table illustrates the number of unique patients for 8 sentinel procedures, and the percentage of patients in the ENCOUNTER table who have these procedures. Procedure groups were defined using AHRQ's Clinical Classification Software (<https://www.hcup-us.ahrq.gov/toolssoftware/ccs/ccs.jsp>) for ICD9, ICD10, and CPT/HCPCS procedure codes. These 8 procedures represent cardiac procedures, orthopedic procedures, diagnostic imaging, and procedures common in pediatric populations. These procedures are expected to be represented in most DataMarts.

DC_PX_GROUP	Patients	Percentage of patients in the ENCOUNTER table	Source tables
Arthroplasty knee [CCS 152]	9,648	1.0	PRO_L3_DCGROUP;ENC_L3_N
Coronary artery bypass graft (CABG) [CCS 44]	1,980	0.2	PRO_L3_DCGROUP;ENC_L3_N
CT scan chest [CCS 178]	51,140	5.3	PRO_L3_DCGROUP;ENC_L3_N
Hip replacement, total and partial [CCS 153]	6,218	0.7	PRO_L3_DCGROUP;ENC_L3_N
Mammography [CCS 182]	2,347	0.2	PRO_L3_DCGROUP;ENC_L3_N
Myringotomy [CCS 23]	5,122	0.5	PRO_L3_DCGROUP;ENC_L3_N
Percutaneous transluminal coronary angioplasty (PTCA) [CCS 45]	8,838	0.9	PRO_L3_DCGROUP;ENC_L3_N
Tonsillectomy and/or adenoidectomy [CCS 30]	6,023	0.6	PRO_L3_DCGROUP;ENC_L3_N

Cell counts between 1 and the low-cell count threshold are displayed as BT and treated as missing.

Number of patients in the ENCOUNTER table is the DISTINCT_N for the PATID field in the ENC_L3_N query.

Chart IA. Trend In Vital Measures by Measurement Date, Past 5 Years

This chart illustrates relative changes over time in the number of records found in the VITAL table. For all charts, the X-axis is the 60 months prior to the maximum refresh date. Monthly record counts were standardized over the period shown to have a mean of 0 and a standard deviation of 1. The y-axis reflects the deviation in each month's count from the mean. A value above 0 indicates an above-average number of records; a value below 0 indicates a below-average number of records. Significant inflection points and other unexpected patterns should be investigated.

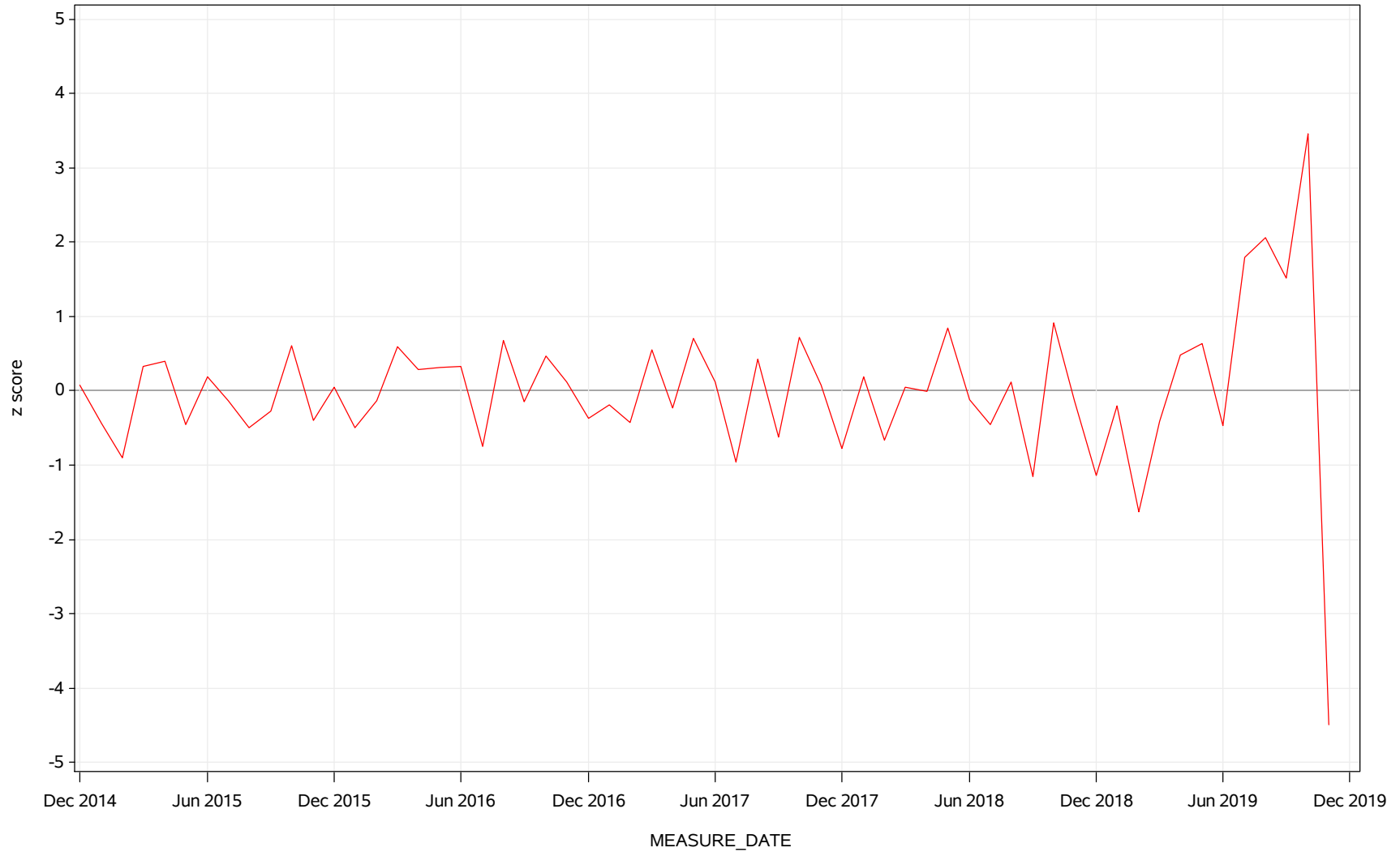


Chart IB. Trend in Encounters by Admit Date and Encounter Type, Past 5 Years

This 2 panel chart illustrates relative changes over time in the number of records per encounter type found in the ENCOUNTER table by admit date. The first panel includes AV, OA, IS IC, and OS encounter types; the second panel includes ED, EI and IP encounter types. For all charts, the X-axis is the 60 months prior to the maximum refresh date. Monthly record counts were standardized over the period shown to have a mean of 0 and a standard deviation of 1. The y-axis reflects the deviation in each month's count from the mean. A value above 0 indicates an above-average number of records; a value below 0 indicates a below-average number of records. Significant inflection points and other unexpected patterns should be investigated.

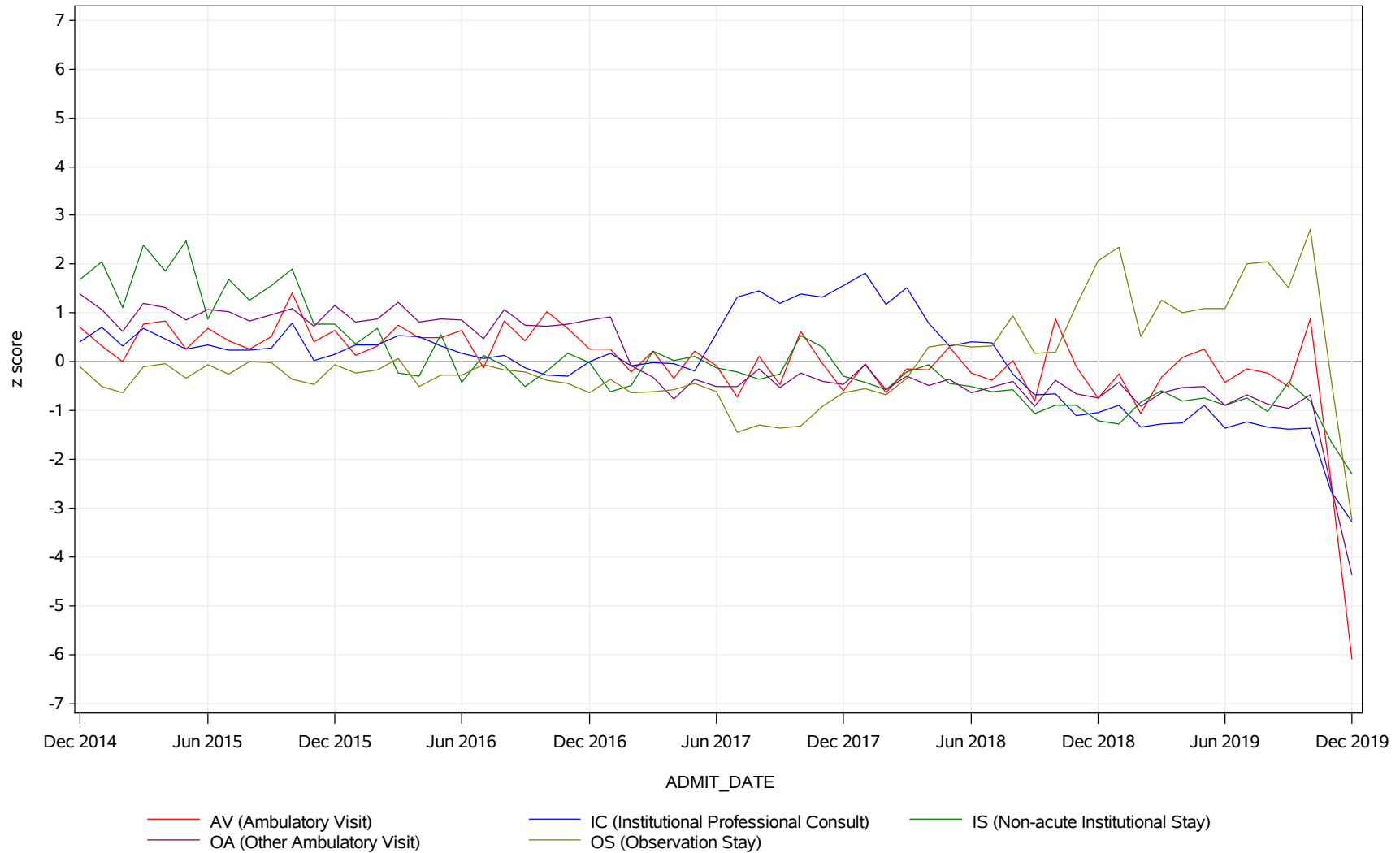


Chart IB. Trend in Encounters by Admit Date and Encounter Type, Past 5 Years (continued)

This 2 panel chart illustrates relative changes over time in the number of records per encounter type found in the ENCOUNTER table by admit date. The first panel includes AV, OA, IS IC, and OS encounter types; the second panel includes ED, EI and IP encounter types. For all charts, the X-axis is the 60 months prior to the maximum refresh date. Monthly record counts were standardized over the period shown to have a mean of 0 and a standard deviation of 1. The y-axis reflects the deviation in each month's count from the mean. A value above 0 indicates an above-average number of records; a value below 0 indicates a below-average number of records. Significant inflection points and other unexpected patterns should be investigated.

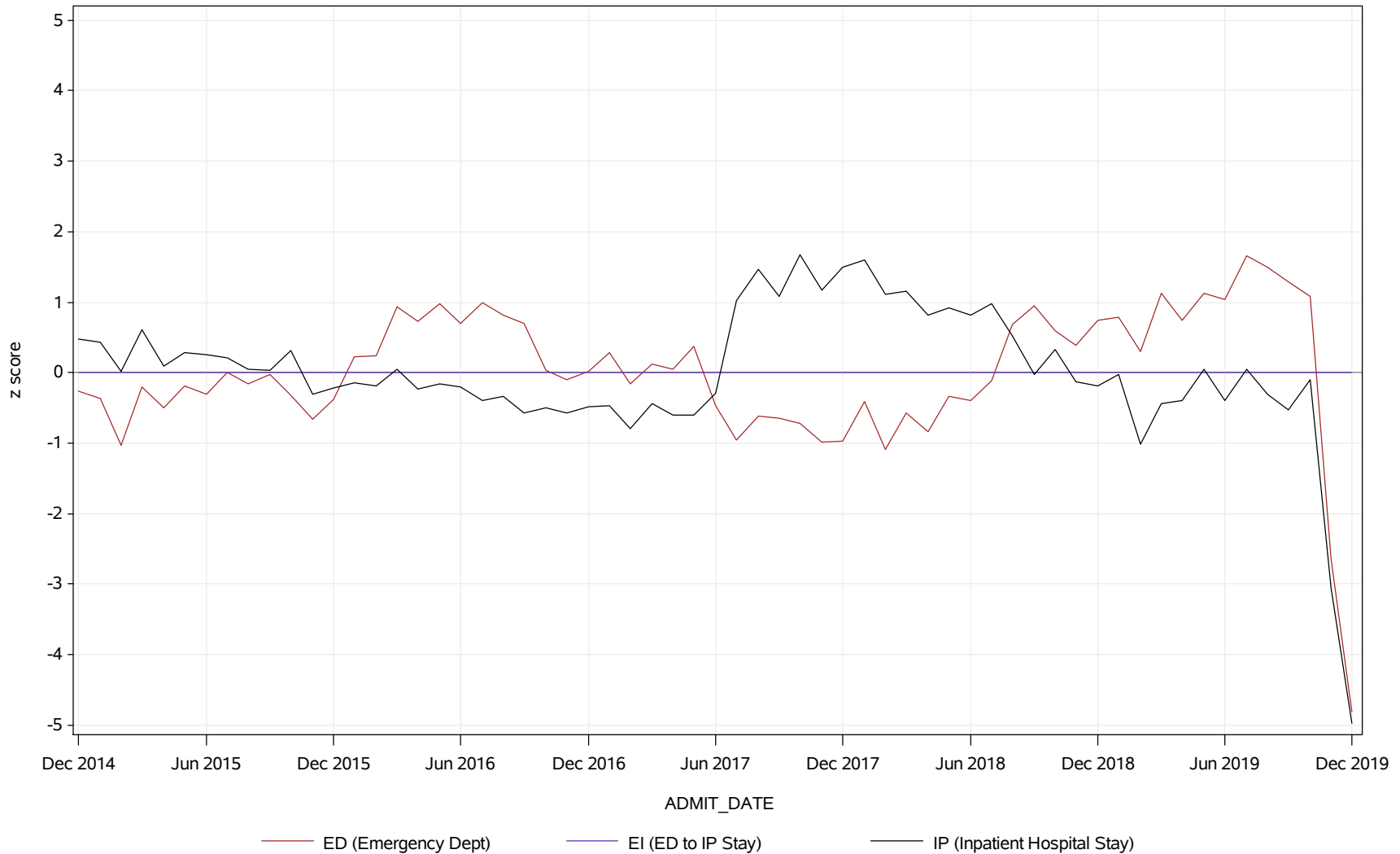


Chart IC. Trend in Institutional Encounters by Discharge Date and Encounter Type, Past 5 Years

This chart illustrates relative changes over time in the number of records for IP, IS, and EI encounter types found in the ENCOUNTER table by discharge date. For all charts, the X-axis is the 60 months prior to the maximum refresh date. Monthly record counts were standardized over the period shown to have a mean of 0 and a standard deviation of 1. The y-axis reflects the deviation in each month's count from the mean. A value above 0 indicates an above-average number of records; a value below 0 indicates a below-average number of records. Significant inflection points and other unexpected patterns should be investigated.

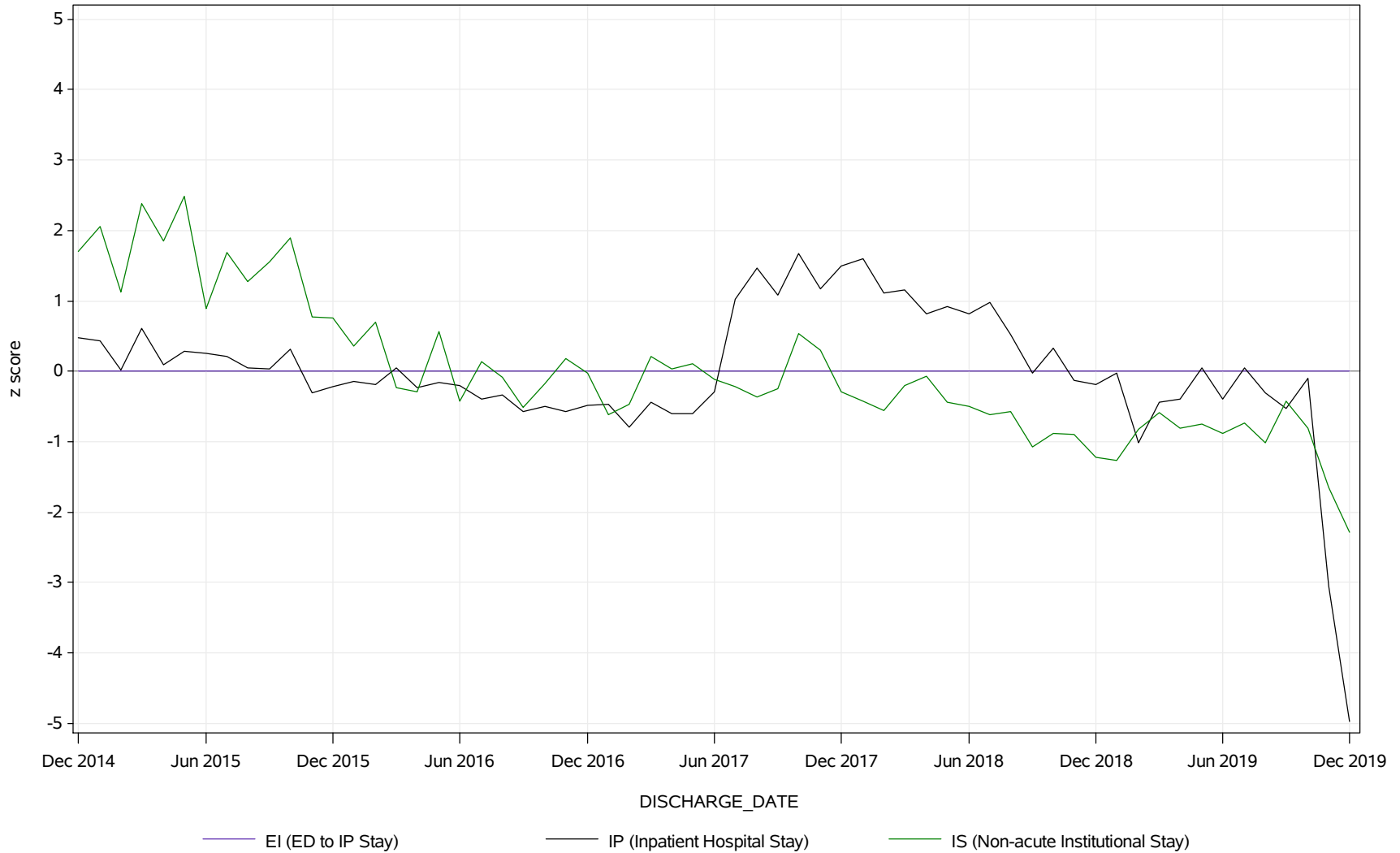


Chart ID. Trend in Lab Results By Result Date, Past 5 Years

This chart illustrates relative changes over time in the number of records found in the LAB_RESULT_CM table. For all charts, the X-axis is the 60 months prior to the maximum refresh date. Monthly record counts were standardized over the period shown to have a mean of 0 and a standard deviation of 1. The y-axis reflects the deviation in each month's count from the mean. A value above 0 indicates an above-average number of records; a value below 0 indicates a below-average number of records. Significant inflection points and other unexpected patterns should be investigated.

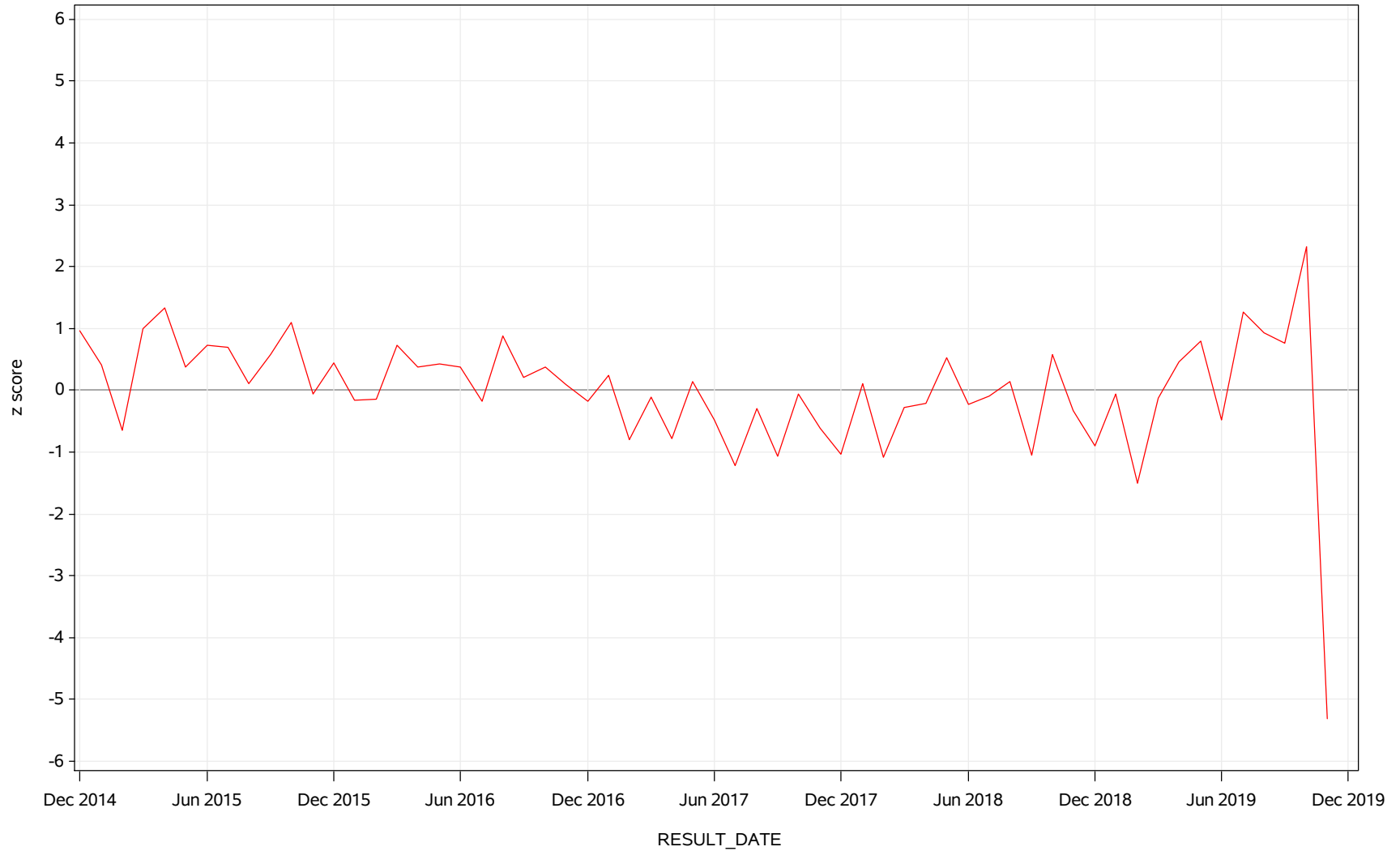


Chart IE. Trend in Prescribed Medications by Rx Order Date, Past 5 Years

This chart illustrates relative changes over time in the number of records found in the PRESCRIBING table. For all charts, the X-axis is the 60 months prior to the maximum refresh date. Monthly record counts were standardized over the period shown to have a mean of 0 and a standard deviation of 1. The y-axis reflects the deviation in each month's count from the mean. A value above 0 indicates an above-average number of records; a value below 0 indicates a below-average number of records. Significant inflection points and other unexpected patterns should be investigated. Significant inflection points and other unexpected patterns should be investigated.

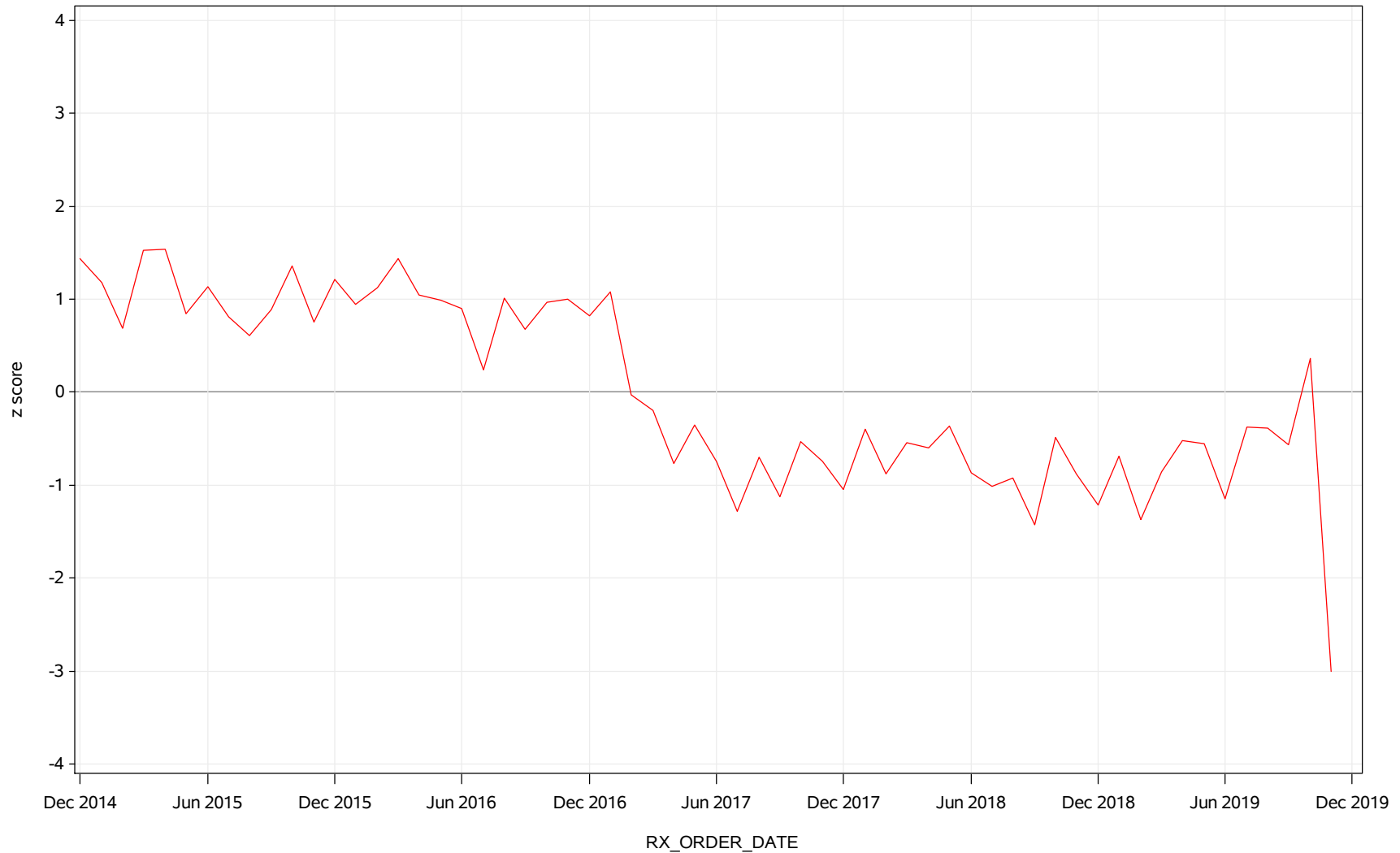


Chart IF. Trend in Dispensed Medications by Dispense Date, Past 5 Years

This chart illustrates relative changes over time in the number of records found in the DISPENSING table. For all charts, the X-axis is the 60 months prior to the maximum refresh date. Monthly record counts were standardized over the period shown to have a mean of 0 and a standard deviation of 1. The y-axis reflects the deviation in each month's count from the mean. A value above 0 indicates an above-average number of records; a value below 0 indicates a below-average number of records. Significant inflection points and other unexpected patterns should be investigated.

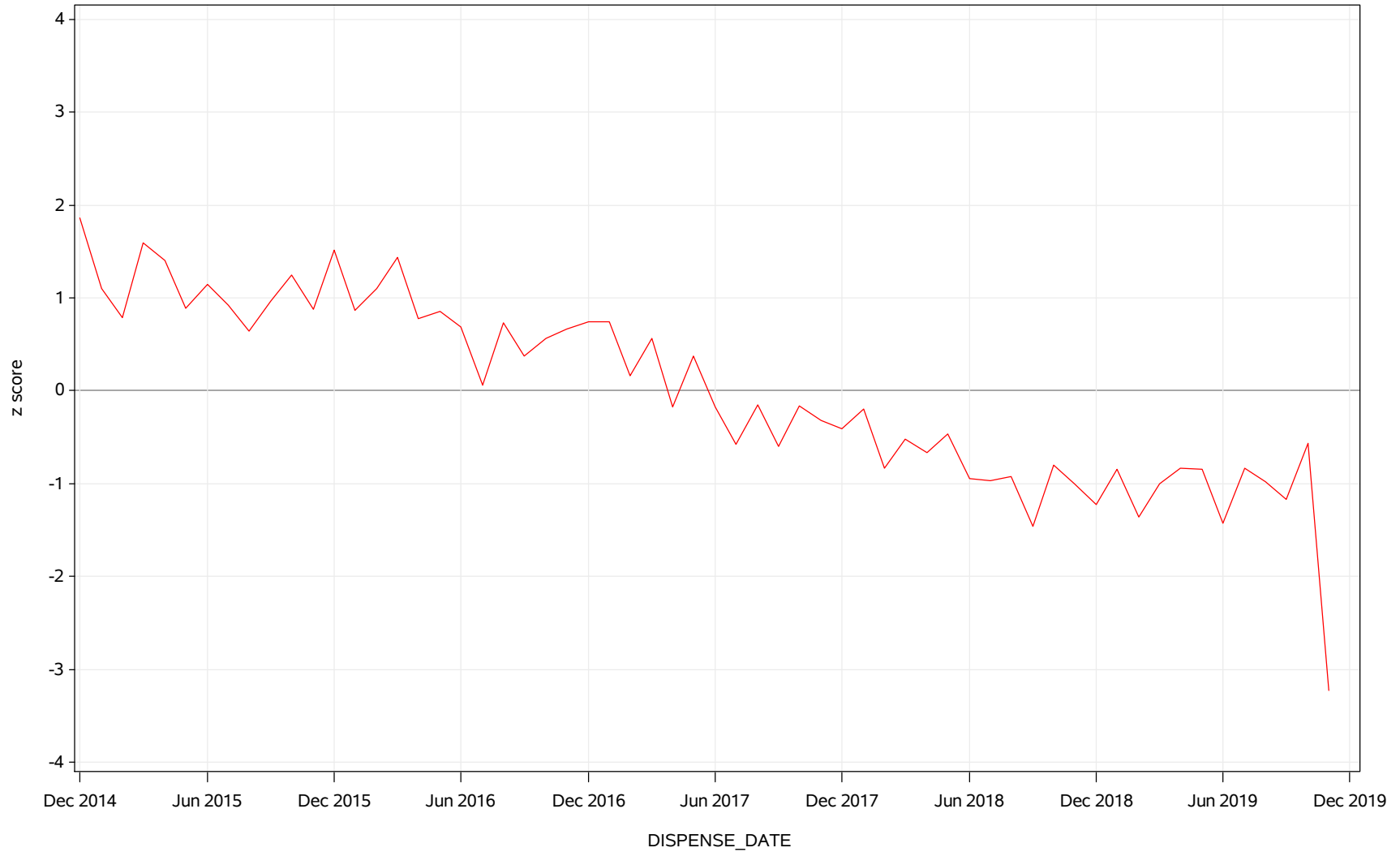


Chart IG. Trend in Administered Medications by Start Date, Past 5 Years

This chart illustrates relative changes over time in the number of records found in the MED_ADMIN table. For all charts, the X-axis is the 60 months prior to the maximum refresh date. Monthly record counts were standardized over the period shown to have a mean of 0 and a standard deviation of 1. The y-axis reflects the deviation in each month's count from the mean. A value above 0 indicates an above-average number of records; a value below 0 indicates a below-average number of records. Significant inflection points and other unexpected patterns should be investigated.

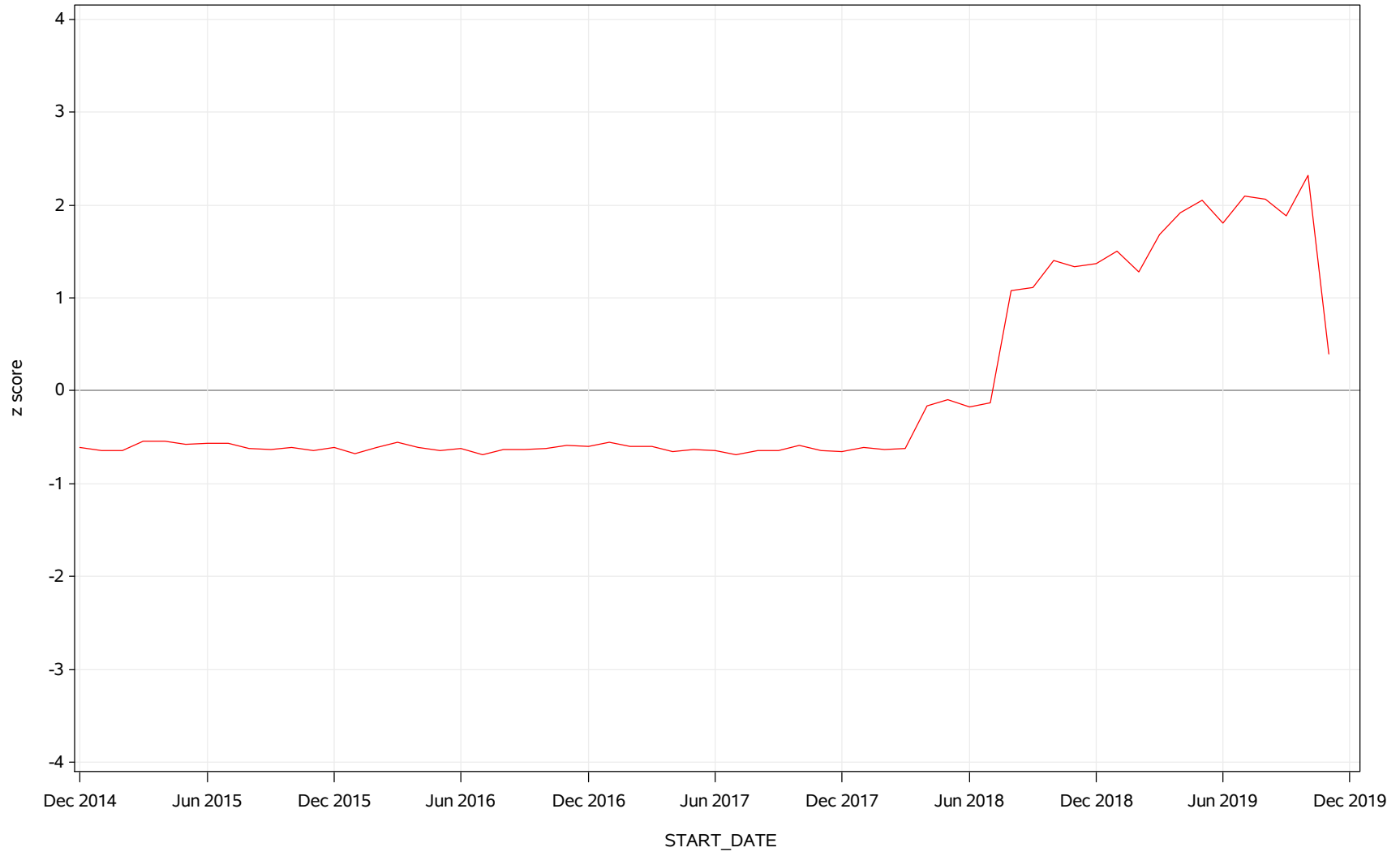


Chart IH. Trend in Condition Records by Report Date, Past 5 Years

This chart illustrates relative changes over time in the number of records found in the CONDITION table. For all charts, the X-axis is the 60 months prior to the maximum refresh date. Monthly record counts were standardized over the period shown to have a mean of 0 and a standard deviation of 1. The y-axis reflects the deviation in each month's count from the mean. A value above 0 indicates an above-average number of records; a value below 0 indicates a below-average number of records. Significant inflection points and other unexpected patterns should be investigated.

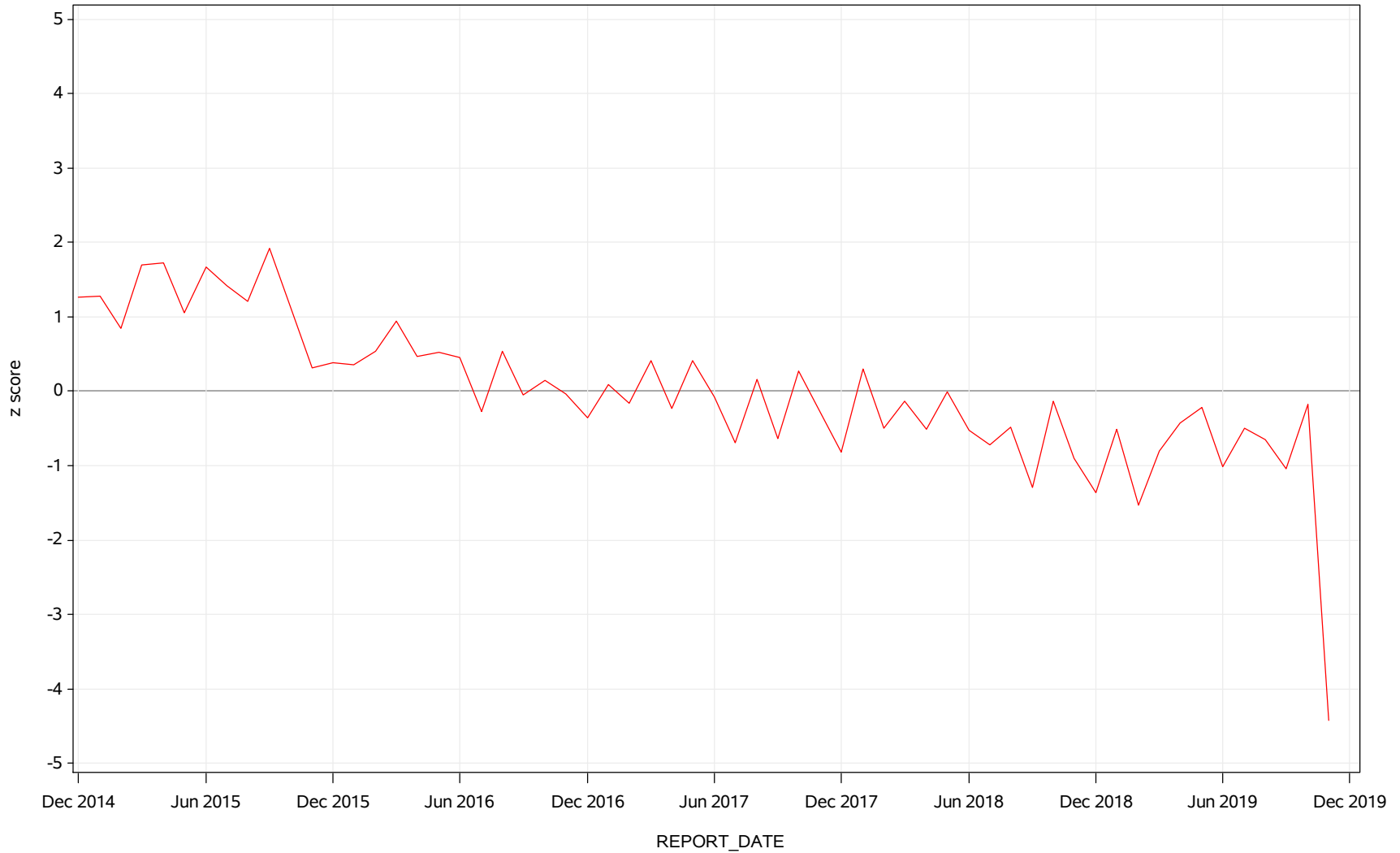


Chart II. Trend in Death Records by Death Date and Source, Past 5 Years

This chart illustrates relative changes over time in the number of records per death source found in the DEATH table. For all charts, the X-axis is the 60 months prior to the maximum refresh date. Monthly record counts were standardized over the period shown to have a mean of 0 and a standard deviation of 1. The y-axis reflects the deviation in each month's count from the mean. A value above 0 indicates an above-average number of records; a value below 0 indicates a below-average number of records. Significant inflection points and other unexpected patterns should be investigated.

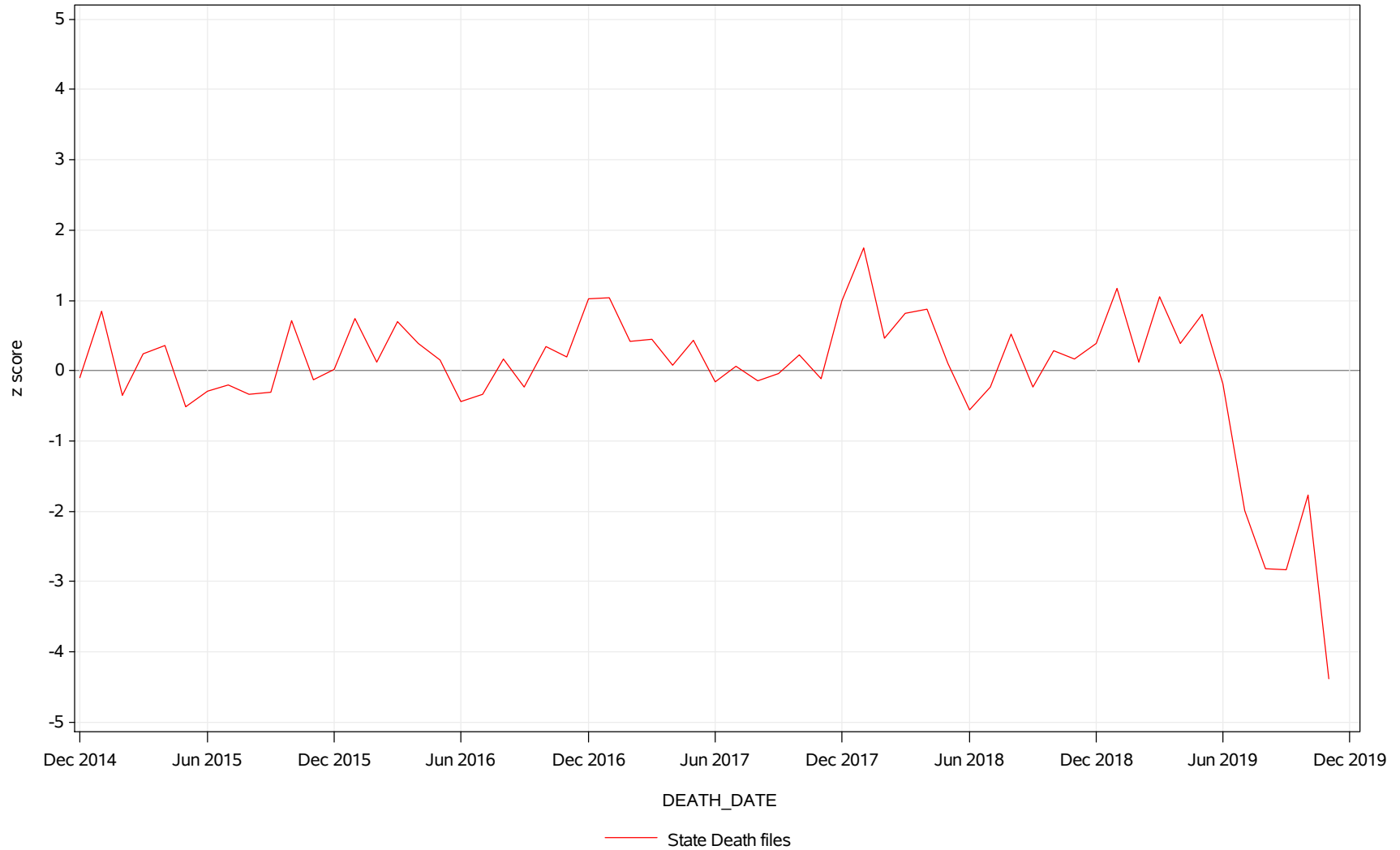


Table IIA. Primary Key Errors

This table shows the required primary key definitions and supports Data Check 1.05 (primary key definition errors). Data check exceptions are highlighted in red and must be corrected.

Table	CDM specifications for primary keys	Exception to specifications	Source table
DEMOGRAPHIC	PATID is unique	No	DEM_L3_N
ENROLLMENT	ENROLLID (concatenation of PATID + ENR_START_DATE+ENR_BASIS) is unique	No	ENR_L3_N
DEATH	DEATHID (concatenation of PATID and DEATH_SOURCE) is unique	No	DEATH_L3_N
ENCOUNTER	ENCOUNTERID is unique	No	ENC_L3_N
DIAGNOSIS	DIAGNOSISID is unique	No	DIA_L3_N
PROCEDURES	PROCEDURESID is unique	No	PRO_L3_N
VITAL	VITALID is unique	No	VIT_L3_N
PRESCRIBING	PRESCRIBINGID is unique	No	PRES_L3_N
DISPENSING	DISPENSINGID is unique	No	DISP_L3_N
LAB_RESULT_CM	LAB_RESULT_CM_ID is unique	No	LAB_L3_N
HARVEST	NETWORKID+DATAMARTID is unique	No	XTBL_L3_METADATA
CONDITION	CONDITIONID is unique	No	COND_L3_N
DEATH_CAUSE	DEATHCID (concatenation of PATID + DEATH_CAUSE + DEATH_CAUSE_CODE + DEATH_CAUSE_TYPE + DEATH_CAUSE_SOURCE) is unique	No	DEATHC_L3_N
PCORNET_TRIAL	TRIAL_KEY (concatenation of PATID + TRIALID + PARTICIPANTID) is unique	No	TRIAL_L3_N
PROVIDER	PROVIDERID is unique	No	PROV_L3_N
MED_ADMIN	MEDADMINID is unique	No	MEDADM_L3_N
LDS_ADDRESS_HISTORY	ADDRESSID is unique	No	LDSADRS_L3_N
IMMUNIZATION	IMMUNIZATIONID is unique	No	IMMUNE_L3_N

The data checking logic is that the count of all records (ALL_N) and distinct records for the variable of interest (DISTINCT_N) must match.

Table IIB. Values Outside of CDM Specifications

This table lists all fields with pre-defined value sets and supports Data Check 1.06 (required fields contain values outside of data model specifications). Data check exceptions are highlighted in red and must be corrected.

Table	Field	Number of records with values outside of specifications	Source table
DEMOGRAPHIC	SEX	0	DEM_L3_SEXDIST
DEMOGRAPHIC	HISPANIC	0	DEM_L3_HISPDIST
DEMOGRAPHIC	RACE	0	DEM_L3_RACEDIST
DEMOGRAPHIC	PAT_PREF_LANGUAGE_SPOKEN	0	DEM_L3_PATPREFLANG
DEMOGRAPHIC	GENDER_IDENTITY	0	DEM_L3_GENDERDIST
DEMOGRAPHIC	SEXUAL_ORIENTATION	0	DEM_L3_ORIENTDIST
DEATH	DEATH_DATE_IMPUTE	0	DEATH_L3_IMPUTE
DEATH	DEATH_SOURCE	0	DEATH_L3_SOURCE
DEATH	DEATH_MATCH_CONFIDENCE	0	DEATH_L3_MATCH
ENCOUNTER	ENC_TYPE	0	ENC_L3_ENCTYPE
ENCOUNTER	DISCHARGE_DISPOSITION	0	ENC_L3_DISDISP
ENCOUNTER	DISCHARGE_STATUS	0	ENC_L3_DISSTAT
ENCOUNTER	DRG_TYPE	0	ENC_L3_DRG_TYPE
ENCOUNTER	ADMITTING_SOURCE	0	ENC_L3_ADMSRC
ENCOUNTER	PAYER_TYPE_PRIMARY	0	ENC_L3_PAYERTYPE1
ENCOUNTER	PAYER_TYPE_SECONDARY	0	ENC_L3_PAYERTYPE2
ENCOUNTER	FACILITY_TYPE	0	ENC_L3_FACILITYTYPE
DIAGNOSIS	ENC_TYPE	0	DIA_L3_ENCTYPE
DIAGNOSIS	DX_TYPE	0	DIA_L3_DXTYPE_DXSOURCE
DIAGNOSIS	DX_SOURCE	0	DIA_L3_DXSOURCE
DIAGNOSIS	PDX	0	DIA_L3_PDX
DIAGNOSIS	DX_POA	0	DIA_L3_DXPOA
PROCEDURES	ENC_TYPE	0	PRO_L3_ENCTYPE
PROCEDURES	PX_TYPE	0	PRO_L3_PXTYPE_ENCTYPE
PROCEDURES	PX_SOURCE	0	PRO_L3_PXSOURCE
PROCEDURES	PPX	0	PRO_L3_PPX

Cell counts between 1 and the low-cell count threshold are displayed as BT and treated as missing.

NDC result is calculated as the difference between the total number of records with NDC codes minus the number of records with NDC codes in the HIPAA format (11-digit, no-dash).

Table IIB. Values Outside of CDM Specifications (continued - page 2 of 5)

This table lists all fields with pre-defined value sets and supports Data Check 1.06 (required fields contain values outside of data model specifications). Data check exceptions are highlighted in red and must be corrected.

Table	Field	Number of records with values outside of specifications	Source table
VITAL	VITAL_SOURCE	0	VIT_L3_VITAL_SOURCE
VITAL	BP_POSITION	0	VIT_L3_BP_POSITION_TYPE
VITAL	SMOKING	0	VIT_L3_SMOKING
VITAL	TOBACCO	0	VIT_L3_TOBACCO
VITAL	TOBACCO_TYPE	0	VIT_L3_TOBACCO_TYPE
PRESCRIBING	RX_BASIS	0	PRES_L3_BASIS
PRESCRIBING	RX_FREQUENCY	0	PRES_L3_FREQ
PRESCRIBING	RX_DOSE_FORM	0	PRES_L3_RXDOSEFORM
PRESCRIBING	RX_DOSE_ORDERED_UNIT	0	PRES_L3_RXDOSEODRUNIT
PRESCRIBING	RX_PRN_FLAG	0	PRES_L3_PRNFLAG
PRESCRIBING	RX_ROUTE	0	PRES_L3_ROUTE
PRESCRIBING	RX_SOURCE	0	PRES_L3_SOURCE
PRESCRIBING	RX_DISPENSE_AS_WRITTEN	0	PRES_L3_DISPASWRTN
DISPENSING	NDC	0	DISP_L3_N
DISPENSING	DISPENSE_DOSE_DISP_UNIT	0	DISP_L3_DOSEUNIT
DISPENSING	DISPENSE_ROUTE	0	DISP_L3_ROUTE
DISPENSING	DISPENSE_SOURCE	0	DISP_L3_SOURCE
LAB_RESULT_CM	LAB_LOINC_SOURCE	0	LAB_L3_LSOURCE
LAB_RESULT_CM	LAB_RESULT_SOURCE	0	LAB_L3_RSOURCE
LAB_RESULT_CM	SPECIMEN_SOURCE	0	LAB_L3_SOURCE
LAB_RESULT_CM	PRIORITY	0	LAB_L3_PRIORITY
LAB_RESULT_CM	RESULT_LOC	0	LAB_L3_LOC
LAB_RESULT_CM	LAB_PX_TYPE	0	LAB_L3_PX_TYPE
LAB_RESULT_CM	RESULT_QUAL	0	LAB_L3_QUAL
LAB_RESULT_CM	RESULT_MODIFIER	0	LAB_L3_MOD
LAB_RESULT_CM	NORM_MODIFIER_LOW	0	LAB_L3_LOW
LAB_RESULT_CM	NORM_MODIFIER_HIGH	0	LAB_L3_HIGH
LAB_RESULT_CM	ABN_IND	0	LAB_L3_ABN
LAB_RESULT_CM	RESULT_UNIT	0	LAB_L3_UNIT

Cell counts between 1 and the low-cell count threshold are displayed as BT and treated as missing.

NDC result is calculated as the difference between the total number of records with NDC codes minus the number of records with NDC codes in the HIPAA format (11-digit, no-dash).

Table IIB. Values Outside of CDM Specifications (continued - page 3 of 5)

This table lists all fields with pre-defined value sets and supports Data Check 1.06 (required fields contain values outside of data model specifications). Data check exceptions are highlighted in red and must be corrected.

Table	Field	Number of records with values outside of specifications	Source table
HARVEST	DATAMART_PLATFORM	0	XTBL_L3_METADATA
HARVEST	DATAMART_CLAIMS	0	XTBL_L3_METADATA
HARVEST	DATAMART_EHR	0	XTBL_L3_METADATA
HARVEST	BIRTH_DATE_MGMT	0	XTBL_L3_METADATA
HARVEST	ENR_START_DATE_MGMT	0	XTBL_L3_METADATA
HARVEST	ENR_END_DATE_MGMT	0	XTBL_L3_METADATA
HARVEST	ADMIT_DATE_MGMT	0	XTBL_L3_METADATA
HARVEST	DISCHARGE_DATE_MGMT	0	XTBL_L3_METADATA
HARVEST	PX_DATE_MGMT	0	XTBL_L3_METADATA
HARVEST	DX_DATE_MGMT	0	XTBL_L3_METADATA
HARVEST	RX_ORDER_DATE_MGMT	0	XTBL_L3_METADATA
HARVEST	RX_START_DATE_MGMT	0	XTBL_L3_METADATA
HARVEST	RX_END_DATE_MGMT	0	XTBL_L3_METADATA
HARVEST	DISPENSE_DATE_MGMT	0	XTBL_L3_METADATA
HARVEST	LAB_ORDER_DATE_MGMT	0	XTBL_L3_METADATA
HARVEST	SPECIMEN_DATE_MGMT	0	XTBL_L3_METADATA
HARVEST	RESULT_DATE_MGMT	0	XTBL_L3_METADATA
HARVEST	MEASURE_DATE_MGMT	0	XTBL_L3_METADATA
HARVEST	ONSET_DATE_MGMT	0	XTBL_L3_METADATA
HARVEST	REPORT_DATE_MGMT	0	XTBL_L3_METADATA
HARVEST	RESOLVE_DATE_MGMT	0	XTBL_L3_METADATA
HARVEST	PRO_DATE_MGMT	0	XTBL_L3_METADATA
HARVEST	DEATH_DATE_MGMT	0	XTBL_L3_METADATA
HARVEST	MEDADMIN_START_DATE_MGMT	0	XTBL_L3_METADATA
HARVEST	MEDADMIN_STOP_DATE_MGMT	0	XTBL_L3_METADATA
HARVEST	OBSCLIN_DATE_MGMT	0	XTBL_L3_METADATA
HARVEST	OBSGEN_DATE_MGMT	0	XTBL_L3_METADATA

Cell counts between 1 and the low-cell count threshold are displayed as BT and treated as missing.

NDC result is calculated as the difference between the total number of records with NDC codes minus the number of records with NDC codes in the HIPAA format (11-digit, no-dash).

Table IIB. Values Outside of CDM Specifications (continued - page 4 of 5)

This table lists all fields with pre-defined value sets and supports Data Check 1.06 (required fields contain values outside of data model specifications). Data check exceptions are highlighted in red and must be corrected.

Table	Field	Number of records with values outside of specifications	Source table
HARVEST	ADDRESS_PERIOD_START_MGMT	0	XTBL_L3_METADATA
HARVEST	ADDRESS_PERIOD_END_MGMT	0	XTBL_L3_METADATA
HARVEST	VX_RECORD_DATE_MGMT	0	XTBL_L3_METADATA
HARVEST	VX_ADMIN_DATE_MGMT	0	XTBL_L3_METADATA
HARVEST	VX_EXP_DATE_MGMT	0	XTBL_L3_METADATA
CONDITION	CONDITION_STATUS	0	COND_L3_STATUS
CONDITION	CONDITION_TYPE	0	COND_L3_TYPE
CONDITION	CONDITION_SOURCE	0	COND_L3_SOURCE
DEATH_CAUSE	DEATH_CAUSE_CODE	0	DEATHC_L3_CODE
DEATH_CAUSE	DEATH_CAUSE_SOURCE	0	DEATHC_L3_SOURCE
DEATH_CAUSE	DEATH_CAUSE_TYPE	0	DEATHC_L3_TYPE
DEATH_CAUSE	DEATH_CAUSE_CONFIDENCE	0	DEATHC_L3_CONF
PROVIDER	PROVIDER_SPECIALTY_PRIMARY	0	PROV_L3_SPECIALTY
PROVIDER	PROVIDER_SEX	0	PROV_L3_SEX
PROVIDER	PROVIDER_NPI_FLAG	0	PROV_L3_NPIFLAG
MED_ADMIN	MEDADMIN_DOSE_ADMIN_UNIT	0	MEDADM_L3_DOSEADMUNIT
MED_ADMIN	MEDADMIN_ROUTE	0	MEDADM_L3_ROUTE
MED_ADMIN	MEDADMIN_SOURCE	0	MEDADM_L3_SOURCE
MED_ADMIN	MEDADMIN_TYPE	0	MEDADM_L3_TYPE
LDS_ADDRESS_HISTORY	ADDRESS_USE	0	LDSADRS_L3_ADRSUSE
LDS_ADDRESS_HISTORY	ADDRESS_TYPE	0	LDSADRS_L3_ADRSTYPE
LDS_ADDRESS_HISTORY	ADDRESS_PREFERRED	0	LDSADRS_L3_ADRSPREF
LDS_ADDRESS_HISTORY	ADDRESS_STATE	0	LDSADRS_L3_ADRSSTATE

Cell counts between 1 and the low-cell count threshold are displayed as BT and treated as missing.

NDC result is calculated as the difference between the total number of records with NDC codes minus the number of records with NDC codes in the HIPAA format (11-digit, no-dash).

Table IIB. Values Outside of CDM Specifications (continued - page 5 of 5)

This table lists all fields with pre-defined value sets and supports Data Check 1.06 (required fields contain values outside of data model specifications). Data check exceptions are highlighted in red and must be corrected.

Table	Field	Number of records with values outside of specifications	Source table
IMMUNIZATION	VX_CODE_TYPE	0	IMMUNE_L3_CODE_CODETYPE
IMMUNIZATION	VX_STATUS	0	IMMUNE_L3_STATUS
IMMUNIZATION	VX_STATUS_REASON	0	IMMUNE_L3_STATUSREASON
IMMUNIZATION	VX_SOURCE	0	IMMUNE_L3_SOURCE
IMMUNIZATION	VX_DOSE_UNIT	0	IMMUNE_L3_DOSEUNIT
IMMUNIZATION	VX_ROUTE	0	IMMUNE_L3_ROUTE
IMMUNIZATION	VX_BODY_SITE	0	IMMUNE_L3_BODYSITE
IMMUNIZATION	VX_MANUFACTURER	0	IMMUNE_L3_MANUFACTURER

Cell counts between 1 and the low-cell count threshold are displayed as BT and treated as missing.

NDC result is calculated as the difference between the total number of records with NDC codes minus the number of records with NDC codes in the HIPAA format (11-digit, no-dash).

Table IIC. Non-Permissible Missing Values

This table lists all fields which are required to be populated, as defined by the table constraints in the CDM, and supports Data Check 1.07 (required fields have non-permissible missing values). Data check exceptions are highlighted in red and must be corrected.

Table	Field	Number of records with missing values	Source table
DEMOGRAPHIC	PATID	0	DEM_L3_N
ENROLLMENT	PATID	0	ENR_L3_N
ENROLLMENT	ENR_START_DATE	0	XTBL_L3_DATES
ENROLLMENT	ENR_BASIS	0	ENR_L3_BASEDIST
DEATH	PATID	0	DEATH_L3_N
DEATH	DEATH_SOURCE	0	DEATH_L3_SOURCE
ENCOUNTER	PATID	0	ENC_L3_N
ENCOUNTER	ENCOUNTERID	0	ENC_L3_N
ENCOUNTER	ADMIT_DATE	0	XTBL_L3_DATES
ENCOUNTER	ENC_TYPE	0	ENC_L3_ENCTYPE
DIAGNOSIS	DIAGNOSISID	0	DIA_L3_N
DIAGNOSIS	PATID	0	DIA_L3_N
DIAGNOSIS	DX	0	DIA_L3_DX
DIAGNOSIS	DX_TYPE	0	DIA_L3_DXTYPE_ENCTYPE
DIAGNOSIS	DX_SOURCE	0	DIA_L3_DXSOURCE
PROCEDURES	PROCEDURESID	0	PRO_L3_N
PROCEDURES	PATID	0	PRO_L3_N
PROCEDURES	PX	0	PRO_L3_PX
PROCEDURES	PX_TYPE	0	PRO_L3_PXTYPE_ENCTYPE
VITAL	PATID	0	VIT_L3_N
VITAL	VITALID	0	VIT_L3_N
VITAL	MEASURE_DATE	0	XTBL_L3_DATES
VITAL	VITAL_SOURCE	0	VIT_L3_VITAL_SOURCE
PRESCRIBING	PRESCRIBINGID	0	PRES_L3_N
PRESCRIBING	PATID	0	PRES_L3_N

Cell counts between 1 and the low-cell count threshold are displayed as BT and treated as missing.

Sums which include BT values are marked with an asterisk (*), either next to the sum of the above-threshold values or alone if all values are BT.

Table IIC. Non-Permissible Missing Values (continued - page 2 of 3)

This table lists all fields which are required to be populated, as defined by the table constraints in the CDM, and supports Data Check 1.07 (required fields have non-permissible missing values). Data check exceptions are highlighted in red and must be corrected.

Table	Field	Number of records with missing values	Source table
DISPENSING	DISPENSINGID	0	DISP_L3_N
DISPENSING	PATID	0	DISP_L3_N
DISPENSING	DISPENSE_DATE	0	XTBL_L3_DATES
DISPENSING	NDC	0	DISP_L3_NDC
LAB_RESULT_CM	LAB_RESULT_CM_ID	0	LAB_L3_N
LAB_RESULT_CM	PATID	0	LAB_L3_N
LAB_RESULT_CM	RESULT_DATE	0	XTBL_L3_DATES
HARVEST	NETWORKID	0	XTBL_L3_METADATA
HARVEST	DATAMARTID	0	XTBL_L3_METADATA
CONDITION	PATID	0	COND_L3_N
CONDITION	CONDITIONID	0	COND_L3_N
CONDITION	CONDITION	0	COND_L3_N
CONDITION	CONDITION_TYPE	0	COND_L3_TYPE
CONDITION	CONDITION_SOURCE	0	COND_L3_SOURCE
DEATH_CAUSE	PATID	0	DEATHC_L3_N
DEATH_CAUSE	DEATH_CAUSE	0	DEATHC_L3_N
DEATH_CAUSE	DEATH_CAUSE_CODE	0	DEATHC_L3_CODE
DEATH_CAUSE	DEATH_CAUSE_TYPE	0	DEATHC_L3_TYPE
DEATH_CAUSE	DEATH_CAUSE_SOURCE	0	DEATHC_L3_SOURCE
PCORNET_TRIAL	PATID	0	TRIAL_L3_N
PCORNET_TRIAL	TRIALID	0	TRIAL_L3_N
PCORNET_TRIAL	PARTICIPANTID	0	TRIAL_L3_N
PROVIDER	PROVIDERID	0	PROV_L3_N
MED_ADMIN	PATID	0	MEDADM_L3_N
MED_ADMIN	MEDADMINID	0	MEDADM_L3_N
MED_ADMIN	MEDADMIN_START_DATE	0	XTBL_L3_DATES

Cell counts between 1 and the low-cell count threshold are displayed as BT and treated as missing.

Sums which include BT values are marked with an asterisk (*), either next to the sum of the above-threshold values or alone if all values are BT.

Table IIC. Non-Permissible Missing Values (continued - page 3 of 3)

This table lists all fields which are required to be populated, as defined by the table constraints in the CDM, and supports Data Check 1.07 (required fields have non-permissible missing values). Data check exceptions are highlighted in red and must be corrected.

Table	Field	Number of records with missing values	Source table
LDS_ADDRESS_HISTORY	PATID	0	LDSADRS_L3_N
LDS_ADDRESS_HISTORY	ADDRESSID	0	LDSADRS_L3_N
LDS_ADDRESS_HISTORY	ADDRESS_USE	0	LDSADRS_L3_ADRSUSE
LDS_ADDRESS_HISTORY	ADDRESS_TYPE	0	LDSADRS_L3_ADRSTYPE
LDS_ADDRESS_HISTORY	ADDRESS_PREFERRED	0	LDSADRS_L3_ADRSPREF
IMMUNIZATION	PATID	0	IMMUNE_L3_N
IMMUNIZATION	IMMUNIZATIONID	0	IMMUNE_L3_N
IMMUNIZATION	VX_CODE	0	IMMUNE_L3_CODE_CODETYPE
IMMUNIZATION	VX_CODE_TYPE	0	IMMUNE_L3_CODE_CODETYPE
IMMUNIZATION	VX_STATUS	0	IMMUNE_L3_STATUS

Cell counts between 1 and the low-cell count threshold are displayed as BT and treated as missing.

Sums which include BT values are marked with an asterisk (*), either next to the sum of the above-threshold values or alone if all values are BT.

Table IID. Diagnostic Errors

This table illustrates exceptions to Data Checks 1.01, 1.02, 1.03, and 1.04 (described below). Data check exceptions are highlighted in red and must be corrected.

Data Check	Data Check Description	Exception	Table(s)	Field(s)	Source table(s)
1.01	Required tables are not present	Required table is not present. All tables must be present in an instantiation of the CDM	None	n/a	DATAMART_ALL
1.02	Expected tables are not populated	Table expected to be populated (DEMOGRAPHIC, ENROLLMENT, ENCOUNTER, DIAGNOSIS, PROCEDURES, and HARVEST) is not populated	None	n/a	DATAMART_ALL
1.03	Required fields are not present	Required numeric field is not present	None	None	DATAMART_ALL
1.03	Required fields are not present	Required character field is not present	None	None	DATAMART_ALL
1.04	Required fields do not conform to data model specifications for data type, length, or name	Required character field is numeric	None	None	DATAMART_ALL
1.04	Required fields do not conform to data model specifications for data type, length, or name	Required numeric field is character	None	None	DATAMART_ALL
1.04	Required fields do not conform to data model specifications for data type, length, or name	Required field is present but of unexpected length	None	None	DATAMART_ALL

Diagnostic checks are performed on all CDM tables and fields. Diagnostic checks ensure conformance to CDM v4.1 table structure, variable names, SAS variable lengths, and SAS data types.

Table IIE. Orphan Records, Replication Errors and Encounter Duplication

This table illustrates exceptions to Data Checks 1.08 (tables contain orphan PATIDs), 1.09 (tables contain orphan ENCOUNTERIDs), 1.10 (replication errors between the ENCOUNTER, PROCEDURES and DIAGNOSIS tables), 1.11 (more than 5% of encounters are assigned to more than one patient), and 1.12 (tables contain orphan PROVIDERIDs). This table will be blank if there are no exceptions. Orphan PATIDs are not present in the DEMOGRAPHIC table. Orphan ENCOUNTERIDs are not present in the ENCOUNTER table. Orphan PROVIDERIDs are not present in the PROVIDER table. Replication errors are ENCOUNTERIDs in the DIAGNOSIS or PROCEDURES table where the encounter type or admit date does not match the corresponding value in the ENCOUNTER table. Data check exceptions are highlighted in red and must be corrected.

Data Check	Data Check Description	Exception	Table(s)	Field(s)	Count	%	Source table(s)
1.08	Orphan PATIDs	Orphan PATID(S) in the CONDITION, DIAGNOSIS, DEATH, DEATH_CAUSE, DISPENSING, ENCOUNTER, ENROLLMENT, HASH_TOKEN, IMMUNIZATION, LAB_RESULT_CM, LDS_ADDRESS_HISTORY, MED_ADMIN, OBS_CLIN, OBS_GEN, PCORNET_TRIAL, PRESCRIBING, PROCEDURES, PRO_CM, or VITAL table					XTBL_L3_MISMATCH; ENR_L3_N; ENC_L3_N; DIA_L3_N; PRO_L3_N; VIT_L3_N; LAB_L3_N; PRES_L3_N; DISP_L3_N; DEATH_L3_N; DEATHC_L3_N; COND_L3_N; PROCM_L3_N; TRIAL_L3_N; OBSCLIN_L3_N; OBSGEN_L3_N; MEDADM_L3_N; HASH_L3_N; IMMUNE_L3_N; LDSADRS_L3_N

For Data Checks 1.08, 1.09, 1.11 and 1.12, the count is the number of distinct IDs with an exception. For Data Check 1.10, the count is the number of records with replication errors. Percent is calculated as the number of distinct IDs with exceptions (shown) divided by the total distinct IDs in the table (not shown). For example, if the DIAGNOSIS table has 2 orphan PATIDs (source table: XTBL_L3_MISMATCH) and 100 distinct PATIDs (source table: DEM_L3_N), the percentage of distinct IDs is 2.0%. Source tables include all potential source tables for the data check calculations.

Table IIE. Orphan Records, Replication Errors and Encounter Duplication (continued - page 2 of 3)

This table illustrates exceptions to Data Checks 1.08 (tables contain orphan PATIDs), 1.09 (tables contain orphan ENCOUNTERIDs), 1.10 (replication errors between the ENCOUNTER, PROCEDURES and DIAGNOSIS tables), 1.11 (more than 5% of encounters are assigned to more than one patient), and 1.12 (tables contain orphan PROVIDERIDs). This table will be blank if there are no exceptions. Orphan PATIDs are not present in the DEMOGRAPHIC table. Orphan ENCOUNTERIDs are not present in the ENCOUNTER table. Orphan PROVIDERIDs are not present in the PROVIDER table. Replication errors are ENCOUNTERIDs in the DIAGNOSIS or PROCEDURES table where the encounter type or admit date does not match the corresponding value in the ENCOUNTER table. Data check exceptions are highlighted in red and must be corrected.

Data Check	Data Check Description	Exception	Table(s)	Field(s)	Count	%	Source table(s)
1.09	Orphan ENCOUNTERIDs	Orphan ENCOUNTERID(S) in the CONDITION, DIAGNOSIS, IMMUNIZATION, LAB_RESULT_CM, MED_ADMIN, OBS_CLIN, OBS_GEN, PRESCRIBING, PROCEDURES, PRO_CM, or VITAL table					XTBL_L3_MISMATCH; DIA_L3_N; PRO_L3_N; VIT_L3_N; PRES_L3_N; LAB_L3_N; PROCML3_N; COND_L3_N; OBSCLIN_L3_N; OBSGEN_L3_N; MEDADM_L3_N; IMMUNE_L3_N
1.10	Replication errors	Replication error(s) in ENC_TYPE or ADMIT_DATE in the DIAGNOSIS or PROCEDURES table				n/a n/a n/a	XTBL_L3_MISMATCH

For Data Checks 1.08, 1.09, 1.11 and 1.12, the count is the number of distinct IDs with an exception. For Data Check 1.10, the count is the number of records with replication errors. Percent is calculated as the number of distinct IDs with exceptions (shown) divided by the total distinct IDs in the table (not shown). For example, if the DIAGNOSIS table has 2 orphan PATIDs (source table: XTBL_L3_MISMATCH) and 100 distinct PATIDs (source table: DEM_L3_N), the percentage of distinct IDs is 2.0%. Source tables include all potential source tables for the data check calculations.

Table IIE. Orphan Records, Replication Errors and Encounter Duplication (continued - page 3 of 3)

This table illustrates exceptions to Data Checks 1.08 (tables contain orphan PATIDs), 1.09 (tables contain orphan ENCOUNTERIDs), 1.10 (replication errors between the ENCOUNTER, PROCEDURES and DIAGNOSIS tables), 1.11 (more than 5% of encounters are assigned to more than one patient), and 1.12 (tables contain orphan PROVIDERIDs). This table will be blank if there are no exceptions. Orphan PATIDs are not present in the DEMOGRAPHIC table. Orphan ENCOUNTERIDs are not present in the ENCOUNTER table. Orphan PROVIDERIDs are not present in the PROVIDER table. Replication errors are ENCOUNTERIDs in the DIAGNOSIS or PROCEDURES table where the encounter type or admit date does not match the corresponding value in the ENCOUNTER table. Data check exceptions are highlighted in red and must be corrected.

Data Check	Data Check Description	Exception	Table(s)	Field(s)	Count	%	Source table(s)
1.11	More than 5% of encounters are assigned to more than one patient	An ENCOUNTERID in the CONDITION, DIAGNOSIS, ENCOUNTER, IMMUNIZATION, LAB_RESULT_CM, MED_ADMIN, OBS_CLIN, OBS_GEN, PRESCRIBING, PROCEDURES, PRO_CM, or VITAL table is associated with more than 1 PATID in the same table					XTBL_L3_NON_UNIQUE; COND_L3_N; DIA_L3_N; ENC_L3_N; LAB_L3_N; PRES_L3_N; PRO_L3_N; VIT_L3_N; PROCM_L3_N; OBSCLIN_L3_N; OBSGEN_L3_N; MEDADM_L3_N; IMMUNE_L3_N
1.12	Orphan PROVIDERIDs	Orphan PROVIDERID(S) in the ENCOUNTER, DIAGNOSIS, IMMUNIZATION, MED_ADMIN, OBS_CLIN, OBS_GEN, PRESCRIBING, or PROCEDURES table.					XTBL_L3_MISMATCH; DIA_L3_N; ENC_L3_N; MEDADM_L3_N; OBSCLIN_L3_N; OBSGEN_L3_N; PRES_L3_N; PRO_L3_N; IMMUNE_L3_N

For Data Checks 1.08, 1.09, 1.11 and 1.12, the count is the number of distinct IDs with an exception. For Data Check 1.10, the count is the number of records with replication errors. Percent is calculated as the number of distinct IDs with exceptions (shown) divided by the total distinct IDs in the table (not shown). For example, if the DIAGNOSIS table has 2 orphan PATIDs (source table: XTBL_L3_MISMATCH) and 100 distinct PATIDs (source table: DEM_L3_N), the percentage of distinct IDs is 2.0%. Source tables include all potential source tables for the data check calculations.

Table IIF. Potential Code Errors

This table illustrates exceptions to Data Check 1.13 (more than 5% of ICD, CPT, LOINC, RXCUI, or NDC codes do not conform to the expected length or content). Results will be displayed for the code types observed in the data. Data check exceptions are highlighted in red and must be corrected. The CDM provides guidance on addressing potential errors (see General Implementation Guidance issue #5).

Table	Code Type	Bad Records	Records	Percent
DIAGNOSIS	09	0	66,225,738	0.00
DIAGNOSIS	10	0	38,711,015	0.00
PROCEDURES	09	0	535,105	0.00
PROCEDURES	10	0	160,396	0.00
PROCEDURES	CH	0	65,889,405	0.00
CONDITION	09	0	8,804,495	0.00
CONDITION	10	0	7,468,458	0.00
PRESCRIBING	RX	0	77,303,389	0.00
DISPENSING	ND	0	4,909,610	0.00
MED_ADMIN	RX	0	1,041,598	0.00
LAB_RESULT_CM	LC	0	225,777,281	0.00

Table IIIA. Future Dates

This table includes most but not all CDM date fields and supports Data Check 2.01 (more than 5% of records have future dates). Future dates are calculated as those with dates occurring after the maximum refresh date. Future dates may be attributable to data entry errors in the source data or ETL errors such as incorrect REFRESH dates or including scheduled appointments in the ENCOUNTER table. Data check exceptions are highlighted in blue and should be investigated and explained in the ETL ADD.

Table	Field	Records with future dates			Source table(s)
		Numerator	Denominator	%	
DEMOGRAPHIC	BIRTH_DATE	0	966,870	0.00	XTBL_L3_DATES
ENROLLMENT	ENR_START_DATE	0	966,868	0.00	XTBL_L3_DATES
DEATH	DEATH_DATE	0	65,767	0.00	XTBL_L3_DATES
ENCOUNTER	ADMIT_DATE	0	66,662,590	0.00	XTBL_L3_DATES
ENCOUNTER	DISCHARGE_DATE	0	66,630,867	0.00	XTBL_L3_DATES
DIAGNOSIS	ADMIT_DATE	0	98,986,521	0.00	XTBL_L3_DATES
DIAGNOSIS	DX_DATE	0	98,986,521	0.00	XTBL_L3_DATES
PROCEDURES	ADMIT_DATE	0	53,609,004	0.00	XTBL_L3_DATES
PROCEDURES	PX_DATE	0	53,609,004	0.00	XTBL_L3_DATES
VITAL	MEASURE_DATE	0	12,713,955	0.00	XTBL_L3_DATES
PRESCRIBING	RX_ORDER_DATE	0	77,303,389	0.00	XTBL_L3_DATES
PRESCRIBING	RX_START_DATE	0	77,303,389	0.00	XTBL_L3_DATES
PRESCRIBING	RX_END_DATE	0	77,303,389	0.00	XTBL_L3_DATES
DISPENSING	DISPENSE_DATE	0	4,909,610	0.00	XTBL_L3_DATES
LAB_RESULT_CM	LAB_ORDER_DATE	0	225,777,281	0.00	XTBL_L3_DATES
LAB_RESULT_CM	SPECIMEN_DATE	0	225,777,281	0.00	XTBL_L3_DATES
LAB_RESULT_CM	RESULT_DATE	0	225,777,281	0.00	XTBL_L3_DATES
HARVEST	REFRESH_MAX	0			XTBL_L3_METADATA
CONDITION	REPORT_DATE	0	16,272,953	0.00	XTBL_L3_DATES
CONDITION	RESOLVE_DATE	0	16,272,953	0.00	XTBL_L3_DATES
CONDITION	ONSET_DATE	0	0		XTBL_L3_DATES
MED_ADMIN	MEDADMIN_START_DATE	0	1,041,598	0.00	XTBL_L3_DATES
MED_ADMIN	MEDADMIN_STOP_DATE	0	1,041,598	0.00	XTBL_L3_DATES

The numerator is the number of records with dates after the maximum DataMart refresh date (see REFRESH_MAX in the XTBL_L3_METADATA table). The denominator is the number of records with a populated date.

Cell counts between 1 and the low-cell count threshold are displayed as BT and treated as missing.

Table IIIA. Future Dates (continued)

This table includes most but not all CDM date fields and supports Data Check 2.01 (more than 5% of records have future dates). Future dates are calculated as those with dates occurring after the maximum refresh date. Future dates may be attributable to data entry errors in the source data or ETL errors such as incorrect REFRESH dates or including scheduled appointments in the ENCOUNTER table. Data check exceptions are highlighted in blue and should be investigated and explained in the ETL ADD.

Table	Field	Records with future dates			Source table(s)
		Numerator	Denominator	%	
LDS_ADDRESS_HISTORY	ADDRESS_PERIOD_START	0	7,479,400	0.00	XTBL_L3_DATES
LDS_ADDRESS_HISTORY	ADDRESS_PERIOD_END	1	5,879,606	0.00	XTBL_L3_DATES
IMMUNIZATION	VX_RECORD_DATE	0	0		XTBL_L3_DATES
IMMUNIZATION	VX_ADMIN_DATE	0	11,768,946	0.00	XTBL_L3_DATES

The numerator is the number of records with dates after the maximum DataMart refresh date (see REFRESH_MAX in the XTBL_L3_METADATA table). The denominator is the number of records with a populated date.

Cell counts between 1 and the low-cell count threshold are displayed as BT and treated as missing.

Table IIIB. Records With Extreme Values

This table supports Data Check 2.02 (more than 10% of records fall into the lowest or highest categories of age, height, weight, diastolic blood pressure, systolic blood pressure, or dispensed days supply). A high percentage of records in these categories may signal incorrect measurement units. Exceptions for blood pressure measures are expected for primarily pediatric populations. Data check exceptions are highlighted in blue and should be investigated and explained in the ETL ADD.

Table	Field	Data Check Parameters		Records	Records with values in the lowest category		Records with values in the highest category		Media n	Source table
		Low	High		N	%	N	%		
DEMOGRAPHIC	AGE (derived from BIRTH_DATE)	<0 yrs.	>89 yrs.	966,870	0	0.0	11,869	1.2	45	DEM_L3_AGEYRSDIST2; DEM_L3_AGEYRSDIST1
VITAL	HT	<0 inches	>=95 inches	6,226,535	0	0.0	759	0.0	66	VIT_L3_HT; VIT_L3_HT_DIST
VITAL	WT	<0 lbs.	>350 lbs.	11,260,683	0	0.0	104,473	0.9	180	VIT_L3_WT; VIT_L3_WT_DIST
VITAL	DIASTOLIC	<40 mgHg	>120 mgHg	12,713,955	63,686	0.5	6,097	0.0	n/a	VIT_L3_DIASTOLIC
VITAL	SYSTOLIC	<40 mgHg	>210 mgHg	12,713,955	2,473	0.0	5,597	0.0	n/a	VIT_L3_SYSTOLIC
DISPENSING	DISPENSE_SUP_GROUP	<1 day	>90 days	4,909,610	126	0.0	11,866	0.2	n/a	DISP_L3_SUPDIST2

Table excludes records with values outside of CDM specifications.

Cell counts between 1 and the low-cell count threshold are displayed as BT and treated as missing. Sums which include BT values are marked with an asterisk (*), either accompanying the sum of all above-threshold values or alone if all values are BT.

Table IIIC. Illogical Dates

This table shows the number of patients who have dates of service in the DEATH, DISPENSING, ENCOUNTER, LAB_RESULT_CM, PRESCRIBING, PROCEDURES, VITAL, MED_ADMIN, OBS_CLIN, or OBS_GEN table which occur before their birth date or after their death date, and the number of patients who have procedure dates occurring before the admit date for the same encounter in the PROCEDURES table. The table shows these numbers as a percentage of patients in the ENCOUNTER table. Although some patients may not be included in the ENCOUNTER table, patients with encounters are the most relevant denominator for this table. This table supports Data Check 2.03 (more than 5% of patients have illogical date relationships). Data check exceptions are highlighted in blue and should be investigated and explained in the ETL ADD.

DATE_COMPARISON	Patients	Percentage of total patients in the ENCOUNTER table	Source tables
ADMIT_DATE < BIRTH_DATE	0	0.0	XTBL_L3_DATE_LOGIC;ENC_L3_N
DISCHARGE_DATE < BIRTH_DATE	0	0.0	XTBL_L3_DATE_LOGIC;ENC_L3_N
PX_DATE < BIRTH_DATE	0	0.0	XTBL_L3_DATE_LOGIC;ENC_L3_N
MEASURE_DATE < BIRTH_DATE	0	0.0	XTBL_L3_DATE_LOGIC;ENC_L3_N
DISPENSE_DATE < BIRTH_DATE	0	0.0	XTBL_L3_DATE_LOGIC;ENC_L3_N
RX_START_DATE < BIRTH_DATE	0	0.0	XTBL_L3_DATE_LOGIC;ENC_L3_N
RESULT_DATE < BIRTH_DATE	0	0.0	XTBL_L3_DATE_LOGIC;ENC_L3_N
DEATH_DATE < BIRTH_DATE	0	0.0	XTBL_L3_DATE_LOGIC;ENC_L3_N
MEDADMIN_START_DATE < BIRTH_DATE	0	0.0	XTBL_L3_DATE_LOGIC;ENC_L3_N
OBSCLIN_DATE < BIRTH_DATE	0	0.0	XTBL_L3_DATE_LOGIC;ENC_L3_N
OBSGEN_DATE < BIRTH_DATE	0	0.0	XTBL_L3_DATE_LOGIC;ENC_L3_N
VX_RECORD_DATE < BIRTH_DATE	0	0.0	XTBL_L3_DATE_LOGIC;ENC_L3_N
ADMIT_DATE > DEATH_DATE	23	0.0	XTBL_L3_DATE_LOGIC;ENC_L3_N
DISCHARGE_DATE > DEATH_DATE	23	0.0	XTBL_L3_DATE_LOGIC;ENC_L3_N
PX_DATE > DEATH_DATE	5	0.0	XTBL_L3_DATE_LOGIC;ENC_L3_N
MEASURE_DATE > DEATH_DATE	2	0.0	XTBL_L3_DATE_LOGIC;ENC_L3_N
DISPENSE_DATE > DEATH_DATE	2	0.0	XTBL_L3_DATE_LOGIC;ENC_L3_N
RX_START_DATE > DEATH_DATE	19	0.0	XTBL_L3_DATE_LOGIC;ENC_L3_N
RESULT_DATE > DEATH_DATE	1	0.0	XTBL_L3_DATE_LOGIC;ENC_L3_N
MEDADMIN_START_DATE > DEATH_DATE	0	0.0	XTBL_L3_DATE_LOGIC;ENC_L3_N
OBSCLIN_DATE > DEATH_DATE	0	0.0	XTBL_L3_DATE_LOGIC;ENC_L3_N
OBSGEN_DATE > DEATH_DATE	0	0.0	XTBL_L3_DATE_LOGIC;ENC_L3_N
VX_RECORD_DATE > DEATH_DATE	0	0.0	XTBL_L3_DATE_LOGIC;ENC_L3_N

Cell counts between 1 and the low-cell count threshold are displayed as BT and treated as missing.

Sums which include BT values are marked with an asterisk (*), either next to the sum of the above-threshold values or alone if all values are BT.

Table IIIC. Illogical Dates (continued)

This table shows the number of patients who have dates of service in the DEATH, DISPENSING, ENCOUNTER, LAB_RESULT_CM, PRESCRIBING, PROCEDURES, VITAL, MED_ADMIN, OBS_CLIN, or OBS_GEN table which occur before their birth date or after their death date, and the number of patients who have procedure dates occurring before the admit date for the same encounter in the PROCEDURES table. The table shows these numbers as a percentage of patients in the ENCOUNTER table. Although some patients may not be included in the ENCOUNTER table, patients with encounters are the most relevant denominator for this table. This table supports Data Check 2.03 (more than 5% of patients have illogical date relationships). Data check exceptions are highlighted in blue and should be investigated and explained in the ETL ADD.

DATE_COMPARISON	Patients	Percentage of total patients in the ENCOUNTER table	Source tables
ADMIT_DATE > DISCHARGE_DATE	0	0.0	XTBL_L3_DATE_LOGIC;ENC_L3_N
PX_DATE is More Than 5 Days Prior To The ADMIT_DATE	0	0.0	XTBL_L3_DATE_LOGIC;ENC_L3_N
PX_DATE is More Than 5 Days After The DISCHARGE_DATE	0	0.0	XTBL_L3_DATE_LOGIC;ENC_L3_N

Cell counts between 1 and the low-cell count threshold are displayed as BT and treated as missing.

Sums which include BT values are marked with an asterisk (*), either next to the sum of the above-threshold values or alone if all values are BT.

Table IIID. Encounters Per Visit and Per Patient

This table shows the number of encounters, patients, encounters per patient, visits (unique encounters per patient, provider, encounter type, and day), and encounters per visit by encounter type. These data support Data Check 2.04 (the average number of encounters per visit is > 2.0 for inpatient (IP), emergency department (ED), or ED to inpatient (EI) encounters). A high number of encounters per visit may signal potential redundancy or duplication. Data check exceptions are highlighted in blue and should be investigated and explained in the ETL ADD.

Encounter Type	Encounters	Patients	Encounters per Patient	Encounters with known PROVIDERID	Visit (unique combinations of PATID, ENC_TYPE, ADMIT_DATE, and PROVIDERID)	Encounters with known PROVIDERID per visit	Source table
AV (Ambulatory Visit)	42,846,033	908,636	47.2	42,846,033	41,331,501	1.04	ENC_L3_ENCTYPE
ED (Emergency Dept)	1,592,511	281,706	5.7	1,592,511	1,337,910	1.19	ENC_L3_ENCTYPE
EI (ED to IP Stay)	0	0		0	0		ENC_L3_ENCTYPE
IC (Institutional Professional Consult)	707,017	90,191	7.8	707,017	686,843	1.03	ENC_L3_ENCTYPE
IP (Inpatient Hospital Stay)	823,215	206,778	4.0	823,215	816,762	1.01	ENC_L3_ENCTYPE
IS (Non-acute Institutional Stay)	95,435	9,788	9.8	95,435	95,354	1.00	ENC_L3_ENCTYPE
OA (Other Ambulatory Visit)	18,997,645	641,330	29.6	18,997,645	15,885,444	1.20	ENC_L3_ENCTYPE
OS (Observation Stay)	38,726	27,634	1.4	38,726	37,799	1.02	ENC_L3_ENCTYPE
Missing, NI, UN or OT	1,562,008	345,030	4.5	1,562,008	1,562,008	1.00	ENC_L3_ENCTYPE

Cell counts between 1 and the low-cell count threshold are displayed as BT and treated as missing.

Sums which include BT values are marked with an asterisk (*), either next to the sum of the above-threshold values or alone if all values are BT.

Table III.E. Specimen Source Discrepancies, Selected Tests

This table compares the actual and reference specimen source categories for selected LAB_LOINC codes. The reference specimen source for each LAB_LOINC code is based on the system attribute assigned by LOINC®. For this comparison, specimen source values for blood, serum and plasma are considered equivalent. The table includes 1 row for matched and unmatched values for the actual category and reference category. Each row shows the number of LAB_LOINC codes, number of records, and percent of all records included in the comparison. LAB_LOINC counts are a sum of distinct LAB_LOINC codes per combination of actual and reference specimen source. These data support Data Check 2.05 (more than 5% of results for selected laboratory tests do not have the appropriate specimen source). Data check exceptions occur when the actual and reference categories do not match and the percentage of records exceeds 5%. Data check exceptions are highlighted in blue and should be investigated and explained in the ETL ADD.

Actual Specimen Source Category	Reference Specimen Source Category	LAB_LOINC codes	Records	Percent of records	Source table
Blood, Plasma, Platelet Poor Plasma, Serum, or Serum/Plasma	Blood, Plasma, Platelet Poor Plasma, Serum, or Serum/Plasma	519	130,485,839	87.41	LAB_L3_LOINC_SOURCE
	Other	133	279,538	0.19	LAB_L3_LOINC_SOURCE
Cerebrospinal Fluid	Cerebrospinal Fluid	10	42,642	0.03	LAB_L3_LOINC_SOURCE
	Other	145	19,143	0.01	LAB_L3_LOINC_SOURCE
Urine	Urine	112	18,241,894	12.22	LAB_L3_LOINC_SOURCE
	Other	202	137,483	0.09	LAB_L3_LOINC_SOURCE
Body Fluid	Body Fluid	30	73,443	0.05	LAB_L3_LOINC_SOURCE
	Other	132	5,238	0.00	LAB_L3_LOINC_SOURCE
Stool	Stool	0	0		LAB_L3_LOINC_SOURCE
	Other	0	0		LAB_L3_LOINC_SOURCE
Cervix/Vagina	Cervix/Vagina	0	0		LAB_L3_LOINC_SOURCE
	Other	7	1,229	0.00	LAB_L3_LOINC_SOURCE
Tissue	Tissue	0	0		LAB_L3_LOINC_SOURCE
	Other	0	0		LAB_L3_LOINC_SOURCE
Any of the above			149,286,449		LAB_L3_LOINC_SOURCE

This comparison includes LOINC codes where the specimen source could be directly mapped to a CDM-defined value and were either among the most common 2000 tests (defined by LOINC) or were included in the data curation lab groups. The reference specimen source for each LAB_LOINC code is available in the lab_loinc_ref table, which was created from the 2010-2018 Unified Medical Language System® (UMLS®) release files and the PCORnet_CDM_ValueSet_ReferenceFile_v1.4.

CDM values for the categories shown above are as follows: Blood, Plasma, Platelet Poor Plasma, Serum, or Serum/Plasma includes BLD, BLD.DOT, BLD_TISS, BLD_TISS^DONOR, BLDA, BLDC, BLDCO, BLDCOA, BLDCOV, BLDV, BLD^DONOR, BLD^FETUS, RBC, RBC^DONOR, WBC, SER_PLAS, SER_PLAS_BLD, SER_PLAS^DONOR, SER_PLAS.ULTRACENTRIFUGATE, SER, SER^DONOR, PLAS, PPP, PPP_BLD, and PPP^FETUS. Cerebrospinal Fluid includes CSF. Urine includes URINE and URINE_SED. Body Fluid includes AMNIO_FLD, BODY_FLD, GAST_FLD, PERICARD_FLD, PERITON_FLD, PLR_FLD, SALIVA, SEMEN, SPUTUM, SWEAT, and SYNV_FLD. Stool includes STOOL. Cervix/Vagina includes CVX and CVX_VAG. Tissue includes TISS and TISS.FNA.

*Sum includes BT values.

Table IIIF. Lab Result Outliers, Selected Tests

This table includes PCORnet and DataMart results for selected LOINC codes[^]. These data support Data Check 2.06 (median lab result values for selected laboratory tests are statistical or clinical outliers). Results are displayed for lab groups which include at least one LOINC code that is a statistical or clinical outlier. Data check exceptions are highlighted in blue and should be investigated and explained in the ETL ADD. The contents of the table are as follows:

- a) Data Curation Normal Range: For non-pediatric DataMarts, this represents the range of normal lab test results based on the lowest value of the lower bounds and highest value of the upper bounds from three different clinical lab information sources for the example unit of measure shown here. For pediatric DataMarts, this information was provided by Cincinnati Children's Hospital.
- b) PCORnet Results: Displays the Median, the Outlier Lower Bound (25th percentile-3*interquartile range), the Outlier Upper Bound (75th percentile+3*interquartile range), and indicates whether the Median is Outside Normal Range.
- c) DataMart Results: Displays the N (number of records) and the Median and indicates whether the Median is a Statistical Outlier, Outside Normal Range, and/or a Clinical Outlier. A Statistical Outlier is defined as a DataMart Median below the Outlier Lower Bound or above the Outlier Upper Bound. A Clinical Outlier is defined as a DataMart Median outside of the normal range and a PCORnet Median within the normal range.

Data Curation Lab Group	LAB LOINC	EX_UCUM_UNITS	Data Curation Normal Range		PCORnet Results				DataMart Results				
			Clinical Lower Bound	Clinical Upper Bound	Median	Outlier Lower Bound	Outlier Upper Bound	Outside Normal Range	N	Median	Statistical Outlier	Outside Normal Range	Clinical Outlier
GLUCOSE B/S/P	2339-0	MG/DL	60.0	126.0	127.0	54.0	201.0	Yes	82020	106.0	No	No	No
GLUCOSE B/S/P	2345-7	MG/DL	60.0	126.0	106.0	88.0	123.0	No	5021491	107.0	No	No	No
GLUCOSE B/S/P	32016-8	MG/DL	60.0	126.0	121.5	0.0	285.0	No	823189	147.0	No	Yes	Yes
GLUCOSE B/S/P	41652-9	MG/DL	60.0	126.0	125.0	32.0	207.0	No	75315	136.0	No	Yes	Yes
HEMATOCRIT	20570-8	%	34.9	50.4	34.3	17.7	52.0	Yes	4484481	38.5	No	No	No
HEMATOCRIT	4544-3	%	34.9	50.4	35.8	29.8	41.7	No	2871	29.0	Yes	Yes	Yes

[^]Source tables are LAB_L3_LOINC_RESULT_NUM and Q2_STAT_DLG_LOINC.

[^]LOINC codes belonging to the top 30 data curation lab groups (see Table 1G) for which PCORnet results are available. PCORnet results were calculated using Cycle 6 results for all DataMarts approved as of Aug 1, 2019, and were calculated separately for pediatric DataMarts and non-pediatric DataMarts. Results were calculated for LOINC codes that had at least 1,000 records occurring in at least 2 pediatric DataMarts, or at least 10,000 records occurring in at least 5 non-pediatric DataMarts. Pediatric results are available for 53 LOINC codes; non-pediatric results are available for 101 LOINC codes.

Table IIIG. Monthly Record Volume Outliers, Selected Domains

This table displays data check exceptions for Data Check 2.08 (the monthly volume of encounter, diagnosis, procedure, vital, prescribing, or laboratory records is an outlier). The Evaluation Window[^] is shown in the table. Outliers are identified by computing a Difference Ratio^{^^} of each month compared to statistics of the previous 12 months. An outlier is defined as 0 records or a Difference Ratio of -7.0 or less. Months are excluded from the data check calculation if (a) the average record count during the previous 12 months was <500 or (b) the standard deviation during the previous 12 months was 0. Data check exceptions are graphically compared to monthly record volumes in Charts IIIA through Chart IIIF. Data check exceptions are highlighted in blue and should be investigated and explained in the ETL ADD.

Table	Evaluation Window [^]	Encounter Type	Exception Month	Records	Average Records in the Previous 12 Months	Standard Deviation in the Previous 12 Months	Difference Ratio ^{^^}	Source table
	Not eligible (SAS_ETS is required)							

[^] The Evaluation Window is determined as follows. The most recent twelve months are excluded because they are already evaluated by the data latency checks (Data Checks 3.07 and 3.11). Months prior to the 5th percentile for the Encounter Admission Date, Rx Order Date, Lab Result Date, and Vital Measurement Date are excluded since volumes may be unstable in these early months.

^{^^} The Difference Ratio is calculated as the volume in the current month minus the average volume in the previous 12 months divided by the standard deviation during the previous 12 months.

Encounter types of interest are AV (Ambulatory Visit), ED (Emergency Department), EI (Emergency Department Admit to Inpatient Hospital Stay), and IP (Inpatient Hospital Stay).

Table IVA. Diagnosis Records Per Encounter and Per Patient, Overall and by Encounter Type

This table supports Data Check 3.01 (the average number of diagnoses records with known diagnosis types per encounter is below threshold [1.0 for ambulatory (AV), inpatient (IP), emergency department (ED), or ED to inpatient (EI) encounters]). Data check exceptions are highlighted in blue and should be investigated and explained in the ETL ADD.

Encounter Type	DIAGNOSIS records	DIAGNOSIS records with known DX_TYPE	ENCOUNTER records	Diagnosis records per encounter	Diagnosis records with known DX_TYPE per encounter	Source table
AV (Ambulatory Visit)	85,292,213	85,292,213	42,846,033	1.99	1.99	DIA_L3_ENCTYPE; DIA_L3_DXTYPE_ENCTYPE; ENC_L3_ENCTYPE
ED (Emergency Dept)	3,742,535	3,742,535	1,592,511	2.35	2.35	DIA_L3_ENCTYPE; DIA_L3_DXTYPE_ENCTYPE; ENC_L3_ENCTYPE
EI (ED to IP Stay)	0	0	0			DIA_L3_ENCTYPE; DIA_L3_DXTYPE_ENCTYPE; ENC_L3_ENCTYPE
IC (Institutional Professional Consult)	2,721,118	2,721,118	707,017	3.85	3.85	DIA_L3_ENCTYPE; DIA_L3_DXTYPE_ENCTYPE; ENC_L3_ENCTYPE
IP (Inpatient Hospital Stay)	4,985,450	4,985,450	823,215	6.06	6.06	DIA_L3_ENCTYPE; DIA_L3_DXTYPE_ENCTYPE; ENC_L3_ENCTYPE
IS (Non-acute Institutional Stay)	352,449	352,449	95,435	3.69	3.69	DIA_L3_ENCTYPE; DIA_L3_DXTYPE_ENCTYPE; ENC_L3_ENCTYPE
OA (Other Ambulatory Visit)	1,503,780	1,503,780	18,997,645	0.08	0.08	DIA_L3_ENCTYPE; DIA_L3_DXTYPE_ENCTYPE; ENC_L3_ENCTYPE
OS (Observation Stay)	388,976	388,976	38,726	10.04	10.04	DIA_L3_ENCTYPE; DIA_L3_DXTYPE_ENCTYPE; ENC_L3_ENCTYPE
Missing, NI, UN or OT	0	0	1,562,008	0.00	0.00	DIA_L3_ENCTYPE; DIA_L3_DXTYPE_ENCTYPE; ENC_L3_ENCTYPE
Total	98,986,521	98,986,521	66,662,590	1.48	1.48	DIA_L3_ENCTYPE; DIA_L3_DXTYPE_ENCTYPE; ENC_L3_ENCTYPE

Cell counts between 1 and the low-cell count threshold are displayed as BT and treated as missing.

Sums which include BT values are marked with an asterisk (*), either next to the sum of the above-threshold values or alone if all values are below threshold.

The denominator comes from the ENCOUNTER table; results will be inflated if there are a high number of orphan ENCOUNTERIDs in the DIAGNOSIS table.

Chart IVA. Diagnosis Records Per Encounter by Admit Date and Encounter Type, Past 5 Years

This chart displays changes over time in the number of diagnosis codes per encounter in the ENCOUNTER table. For all charts, the X-axis is the 60 months prior to the maximum refresh date. Significant inflection points and other unexpected patterns should be investigated.

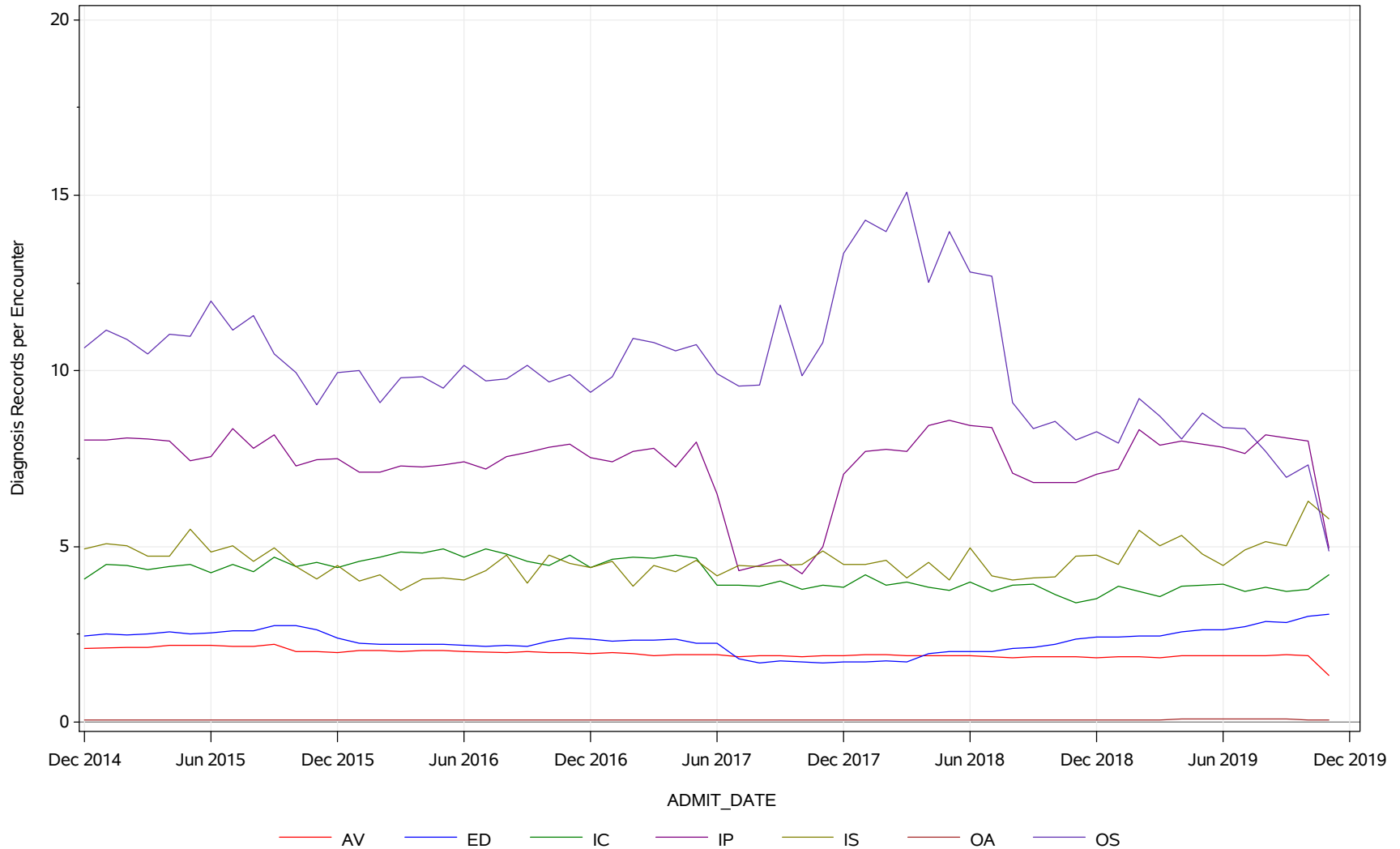


Table IVB. Procedure Records Per Encounter and Per Patient, Overall and by Encounter Type

This table supports Data Check 3.02 (the average number of procedure records with known procedure types per encounter is below threshold [0.75 for ambulatory (AV) encounters, 0.75 for emergency department (ED) encounters, 1.00 for ED to inpatient (EI) encounters, and 1.00 for inpatient (IP) encounters]). Data check exceptions are highlighted in blue and should be investigated and explained in the ETL ADD.

Encounter Type	PROCEDURES records	PROCEDURES records with known PX_TYPE	ENCOUNTER records	Procedures records per encounter	Procedures records with known PX_TYPE per encounter	Source table
AV (Ambulatory Visit)	49,319,785	49,319,785	42,846,033	1.15	1.15	PRO_L3_ENCTYPE; PRO_L3_PXTYPE_ENCTYPE; ENC_L3_ENCTYPE
ED (Emergency Dept)	2,014,832	2,014,832	1,592,511	1.27	1.27	PRO_L3_ENCTYPE; PRO_L3_PXTYPE_ENCTYPE; ENC_L3_ENCTYPE
EI (ED to IP Stay)	0	0	0			PRO_L3_ENCTYPE; PRO_L3_PXTYPE_ENCTYPE; ENC_L3_ENCTYPE
IC (Institutional Professional Consult)	753,672	753,672	707,017	1.07	1.07	PRO_L3_ENCTYPE; PRO_L3_PXTYPE_ENCTYPE; ENC_L3_ENCTYPE
IP (Inpatient Hospital Stay)	1,009,383	1,009,383	823,215	1.23	1.23	PRO_L3_ENCTYPE; PRO_L3_PXTYPE_ENCTYPE; ENC_L3_ENCTYPE
IS (Non-acute Institutional Stay)	106,380	106,380	95,435	1.11	1.11	PRO_L3_ENCTYPE; PRO_L3_PXTYPE_ENCTYPE; ENC_L3_ENCTYPE
OA (Other Ambulatory Visit)	200,049	200,049	18,997,645	0.01	0.01	PRO_L3_ENCTYPE; PRO_L3_PXTYPE_ENCTYPE; ENC_L3_ENCTYPE
OS (Observation Stay)	204,903	204,903	38,726	5.29	5.29	PRO_L3_ENCTYPE; PRO_L3_PXTYPE_ENCTYPE; ENC_L3_ENCTYPE
Missing, NI, UN or OT	0	0	1,562,008	0.00	0.00	PRO_L3_ENCTYPE; PRO_L3_PXTYPE_ENCTYPE; ENC_L3_ENCTYPE
Total	53,609,004	53,609,004	66,662,590	0.80	0.80	PRO_L3_ENCTYPE; PRO_L3_PXTYPE_ENCTYPE; ENC_L3_ENCTYPE

Cell counts between 1 and the low-cell count threshold are displayed as BT and treated as missing.

Sums which include BT values are marked with an asterisk (*), either next to the sum of the above-threshold values or alone if all values are below threshold.

The denominator comes from the ENCOUNTER table; results will be inflated if there are a high number of orphan ENCOUNTERIDs in the PROCEDURES table.

Chart IVB. Procedure Records Per Encounter by Admit Date and Encounter Type, Past 5 Years

This chart displays changes over time in the number of procedure codes per encounter in the ENCOUNTER table. For all charts, the X-axis is the 60 months prior to the maximum refresh date. Significant inflection points and other unexpected patterns should be investigated.

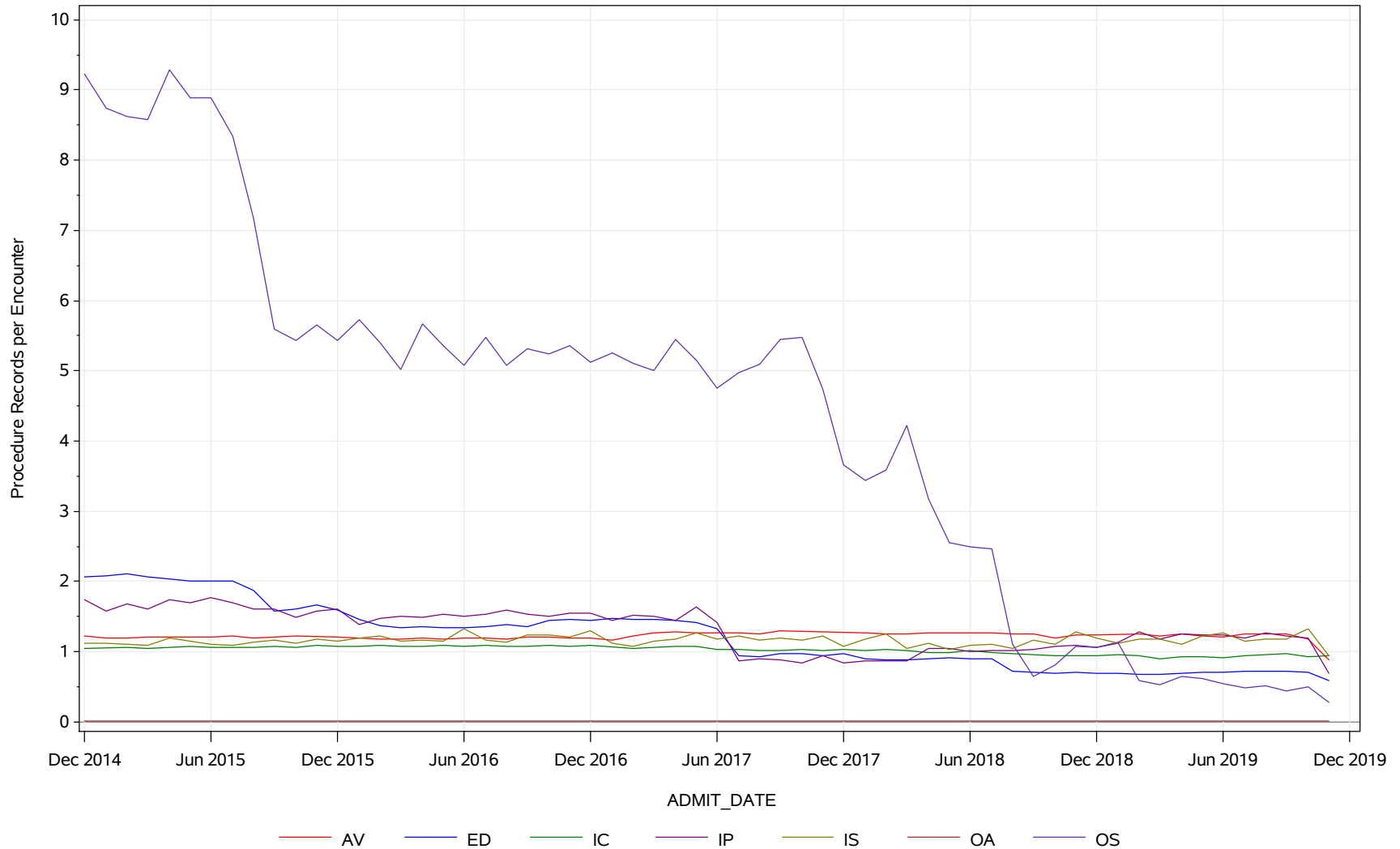


Table IVC. Missing or Unknown Values, Required Tables

This table includes fields in the DEMOGRAPHIC, ENROLLMENT, ENCOUNTER, DIAGNOSIS, and PROCEDURES tables which are included in the query results and are not required to be populated. The table depicts the percentage of records with missing or unknown values. Results support Data Check 3.03 (more than 10% of records have missing or unknown values) for the following fields: BIRTH_DATE, SEX, DISCHARGE_DATE (IP/EI encounters only), DISCHARGE_DISPOSITION (IP/EI encounters only), DX_ORIGIN, DIAGNOSIS.ENCOUNTERID, PX_DATE, PX_SOURCE, and PROCEDURES.ENCOUNTERID. Data check exceptions are highlighted in blue and should be investigated and explained in the ETL ADD. Network-wide results for PCORnet DataMarts approved as of Nov 7, 2019 are presented for comparison; a result of 20.0 means that half of the DataMarts had values below 20%.

Table	Field	Encounter Type Constraint	Records with missing, NI, UN, or OT values				Source table
			Numerator	Denominator	%	50th Percentile	
DEMOGRAPHIC	BIRTH_DATE		0	966,870		0.0	XTBL_L3_DATES
DEMOGRAPHIC	BIRTH_TIME		0	966,870		0.0	XTBL_L3_TIMES
DEMOGRAPHIC	SEX		1	966,870	0.0	0.0	DEM_L3_SEXDIST
DEMOGRAPHIC	HISPANIC		302,157	966,870	31.3	15.5	DEM_L3_HISPDIST
DEMOGRAPHIC	RACE		294,799	966,870	30.5	19.6	DEM_L3_RACEDIST
DEMOGRAPHIC	GENDER_IDENTITY		966,870	966,870	100.0	100.0	DEM_L3_GENDERDIST
DEMOGRAPHIC	SEXUAL_ORIENTATION		966,870	966,870	100.0	100.0	DEM_L3_ORIENTDIST
DEMOGRAPHIC	PAT_PREF_LANGUAGE_SPOKEN		286,613	966,870	29.6	29.7	DEM_L3_PATPREFLANG
ENROLLMENT	ENR_END_DATE		0	966,868		0.0	XTBL_L3_DATES
ENROLLMENT	CHART		0	966,868		0.0	ENR_L3_CHART
ENCOUNTER	ADMIT_TIME		0	66,662,590		0.0	XTBL_L3_TIMES
ENCOUNTER	DISCHARGE_DATE	IP or EI	0	823,215		0.0	ENC_L3_ENCTYPE_DDATE_YM
ENCOUNTER	DISCHARGE_TIME		31,723	66,662,590	0.0	25.4	XTBL_L3_TIMES
ENCOUNTER	ENC_TYPE		1,562,008	66,662,590	2.3	0.9	ENC_L3_ENCTYPE
ENCOUNTER	PROVIDERID		0	66,662,590		1.8	ENC_L3_N
ENCOUNTER	FACILITYID		118,617	66,662,590	0.2	0.0	ENC_L3_N
ENCOUNTER	DISCHARGE_DISPOSITION	IP or EI	581,284	823,215	70.6	0.9	ENC_L3_ENCTYPE_DISDISP
ENCOUNTER	DISCHARGE_STATUS	IP or EI	588,056	823,215	71.4	7.8	ENC_L3_ENCTYPE_DISSTAT

Table excludes records with values outside of CDM specifications.

The four 'flavors of null' defined in the CDM are combined here but details are available in the source tables.

The denominator is derived from the applicable _N query (ALL_N + NULL_N for TAG=PATID) or the specific query for measures with encounter type constraints.

Sums which include BT values are marked with an asterisk (*), either next to the sum of the above-threshold values or alone if all values are BT.

Table IVC. Missing or Unknown Values, Required Tables (continued)

This table includes fields in the DEMOGRAPHIC, ENROLLMENT, ENCOUNTER, DIAGNOSIS, and PROCEDURES tables which are included in the query results and are not required to be populated. The table depicts the percentage of records with missing or unknown values. Results support Data Check 3.03 (more than 10% of records have missing or unknown values) for the following fields: BIRTH_DATE, SEX, DISCHARGE_DATE (IP/EI encounters only), DISCHARGE_DISPOSITION (IP/EI encounters only), DX_ORIGIN, DIAGNOSIS.ENCOUNTERID, PX_DATE, PX_SOURCE, and PROCEDURES.ENCOUNTERID. Data check exceptions are highlighted in blue and should be investigated and explained in the ETL ADD. Network-wide results for PCORnet DataMarts approved as of Nov 7, 2019 are presented for comparison; a result of 20.0 means that half of the DataMarts had values below 20%.

Table	Field	Encounter Type Constraint	Records with missing, NI, UN, or OT values				Source table
			Numerator	Denominator	%	50th Percentile	
ENCOUNTER	DRG	IP or EI	603,672	823,215	73.3	20.0	ENC_L3_ENCTYPE_DRG
ENCOUNTER	ADMITTING_SOURCE	IP or EI	665,460	823,215	80.8	56.0	ENC_L3_ENCTYPE_ADMSRC
ENCOUNTER	PAYER_TYPE_PRIMARY		13,526,500	66,662,590	20.3	62.2	ENC_L3_PAYERTYPE1
ENCOUNTER	PAYER_TYPE_SECONDARY		50,587,438	66,662,590	75.9	97.9	ENC_L3_PAYERTYPE2
ENCOUNTER	FACILITY_TYPE		66,662,590	66,662,590	100.0	100.0	ENC_L3_FACILITYTYPE
DIAGNOSIS	DX_DATE		0	98,986,521		100.0	XTBL_L3_DATES
DIAGNOSIS	DX_TYPE		0	98,986,521		0.0	DIA_L3_DXTYPE_DXSOURCE
DIAGNOSIS	DX_SOURCE		12,768,438	98,986,521	12.9	0.0	DIA_L3_DXSOURCE
DIAGNOSIS	DX_ORIGIN		0	98,986,521		0.0	DIA_L3_ORIGIN
DIAGNOSIS	PDX	IP or EI	3,372,848	4,985,450	67.7	0.0	DIA_L3_PDX_ENCTYPE
DIAGNOSIS	ENCOUNTERID		0	98,986,521		0.0	DIA_L3_N
DIAGNOSIS	DX_POA		89,517,111	98,986,521	90.4	100.0	DIA_L3_DXPOA
DIAGNOSIS	PROVIDERID		0	98,986,521		0.2	DIA_L3_N
PROCEDURES	PX_DATE		0	53,609,004		0.0	XTBL_L3_DATES
PROCEDURES	PX_TYPE		0	53,609,004		0.0	PRO_L3_PXTYPE_ENCTYPE
PROCEDURES	PX_SOURCE		67,057	53,609,004	0.1	0.0	PRO_L3_PXSOURCE
PROCEDURES	ENCOUNTERID		0	53,609,004		0.0	PRO_L3_N
PROCEDURES	PPX		53,609,004	53,609,004	100.0	100.0	PRO_L3_PPX
PROCEDURES	PROVIDERID		0	53,609,004		0.2	PRO_L3_N

Table excludes records with values outside of CDM specifications.

The four 'flavors of null' defined in the CDM are combined here but details are available in the source tables.

The denominator is derived from the applicable _N query (ALL_N + NULL_N for TAG=PATID) or the specific query for measures with encounter type constraints.

Sums which include BT values are marked with an asterisk (*), either next to the sum of the above-threshold values or alone if all values are BT.

Table IVD. Missing or Unknown Values, Optional Tables

This table includes fields in the VITAL, DEATH, LAB_RESULT_CM, PRESCRIBING, DISPENSING, CONDITION, DEATH_CAUSE, PRO_CM, PROVIDER, MED_ADMIN, OBS_CLIN and OBS_GEN tables which are included in the query results and are not required to be populated. Vital measures (e.g. HEIGHT) are not included because the table structure does not support missingness assessment. The table depicts the percentage of records with missing or unknown values. Results support Data Check 3.03 (more than 10% of records have missing or unknown values) for the following fields: RX_ORDER_DATE, VX_RECORD_DATE, DISPENSE_SUP, DEATH_SOURCE, DISPENSE_SOURCE, CONDITION_SOURCE, DX_ORIGIN, LAB_RESULT_SOURCE, MEDADMIN_SOURCE, PRO_SOURCE, PX_SOURCE, RX_SOURCE, VITAL_SOURCE, VX_SOURCE, DIAGNOSIS.ENCOUNTERID, PROCEDURES.ENCOUNTERID, VITAL.ENCOUNTERID, MED_ADMIN.ENCOUNTERID, LAB_RESULT_CM.ENCOUNTERID, MED_ADMIN_CODE, MED_ADMIN_TYPE, OBS_CLIN_CODE, OBS_CLIN_TYPE, OBS_GEN_CODE, OBS_GEN_TYPE. Exceptions are highlighted in blue and should be investigated and explained in the ETL ADD.

Network-wide results for PCORnet DataMarts approved as of Nov 7, 2019 is presented for comparison; a result of 20.0 means that half of the DataMarts had values below 20%.

Table	Field	Records with missing, NI, UN, or OT values				Source table
		Numerator	Denominator	%	50th Percentile	
VITAL	ENCOUNTERID	0	12,713,955		0.0	VIT_L3_N
VITAL	VITAL_SOURCE	0	12,713,955		0.0	VIT_L3_VITAL_SOURCE
VITAL	MEASURE_TIME	0	12,713,955		0.0	XTBL_L3_TIMES
DEATH	DEATH_DATE_IMPUTE	65,774	65,774	100.0	0.0	DEATH_L3_IMPUTE
DEATH	DEATH_MATCH_CONFIDENCE	65,774	65,774	100.0	100.0	DEATH_L3_MATCH
DEATH	DEATH_DATE	0	65,774		0.0	XTBL_L3_DATES
DEATH	DEATH_SOURCE	0	65,774		0.0	DEATH_L3_SOURCE
LAB_RESULT_CM	ENCOUNTERID	0	225,777,281		0.0	LAB_L3_N
LAB_RESULT_CM	SPECIMEN_SOURCE	72,353,648	225,777,281	32.0	15.4	LAB_L3_SOURCE
LAB_RESULT_CM	LAB_LOINC	0	225,777,281		0.0	LAB_L3_LOINC
LAB_RESULT_CM	PRIORITY	225,777,281	225,777,281	100.0	100.0	LAB_L3_PRIORITY
LAB_RESULT_CM	RESULT_LOC	225,777,281	225,777,281	100.0	0.0	LAB_L3_LOC
LAB_RESULT_CM	LAB_PX_TYPE	0	225,777,281		98.7	LAB_L3_PX_TYPE
LAB_RESULT_CM	LAB_PX	0	225,777,281		99.1	LAB_L3_PX_PXTYPE
LAB_RESULT_CM	LAB_ORDER_DATE	0	225,777,281		0.0	XTBL_L3_DATES
LAB_RESULT_CM	SPECIMEN_DATE	0	225,777,281		0.0	XTBL_L3_DATES
LAB_RESULT_CM	SPECIMEN_TIME	0	225,777,281		0.1	XTBL_L3_TIMES
LAB_RESULT_CM	RESULT_TIME	0	225,777,281		0.0	XTBL_L3_TIMES

Table excludes records with values outside of CDM specifications.

The four 'flavors of null' defined in the CDM are combined here but details are available in the source tables.

The denominator is derived from the applicable _N query (ALL_N + NULL_N for TAG=PATID).

Sums which include BT values are marked with an asterisk (*), either next to the sum of the above-threshold values or alone if all values are BT.

Table IVD. Missing or Unknown Values, Optional Tables (continued - page 2 of 6)

This table includes fields in the VITAL, DEATH, LAB_RESULT_CM, PRESCRIBING, DISPENSING, CONDITION, DEATH_CAUSE, PRO_CM, PROVIDER, MED_ADMIN, OBS_CLIN and OBS_GEN tables which are included in the query results and are not required to be populated. Vital measures (e.g. HEIGHT) are not included because the table structure does not support missingness assessment. The table depicts the percentage of records with missing or unknown values. Results support Data Check 3.03 (more than 10% of records have missing or unknown values) for the following fields: RX_ORDER_DATE, VX_RECORD_DATE, DISPENSE_SUP, DEATH_SOURCE, DISPENSE_SOURCE, CONDITION_SOURCE, DX_ORIGIN, LAB_RESULT_SOURCE, MEDADMIN_SOURCE, PRO_SOURCE, PX_SOURCE, RX_SOURCE, VITAL_SOURCE, VX_SOURCE, DIAGNOSIS.ENCOUNTERID, PROCEDURES.ENCOUNTERID, VITAL.ENCOUNTERID, MED_ADMIN.ENCOUNTERID, LAB_RESULT_CM.ENCOUNTERID, MED_ADMIN_CODE, MED_ADMIN_TYPE, OBS_CLIN_CODE, OBS_CLIN_TYPE, OBS_GEN_CODE, OBS_GEN_TYPE. Exceptions are highlighted in blue and should be investigated and explained in the ETL ADD.

Network-wide results for PCORnet DataMarts approved as of Nov 7, 2019 is presented for comparison; a result of 20.0 means that half of the DataMarts had values below 20%.

Table	Field	Records with missing, NI, UN, or OT values				Source table
		Numerator	Denominator	%	50th Percentile	
LAB_RESULT_CM	RESULT_QUAL	205,562,725	225,777,281	91.0	100.0	LAB_L3_QUAL
LAB_RESULT_CM	RESULT_NUM	20,214,556	225,777,281	9.0	7.4	LAB_L3_LOINC_RESULT_NUM
LAB_RESULT_CM	RESULT_MODIFIER	0	225,777,281		0.0	LAB_L3_MOD
LAB_RESULT_CM	RESULT_UNIT	49,298,314	225,777,281	21.8	26.2	LAB_L3_UNIT
LAB_RESULT_CM	NORM_MODIFIER_LOW	67,373,924	225,777,281	29.8	24.8	LAB_L3_LOW
LAB_RESULT_CM	NORM_MODIFIER_HIGH	67,373,924	225,777,281	29.8	25.9	LAB_L3_HIGH
LAB_RESULT_CM	ABN_IND	225,777,281	225,777,281	100.0	75.3	LAB_L3_ABN
LAB_RESULT_CM	LAB_LOINC_SOURCE	72,353,648	225,777,281	32.0	0.0	LAB_L3_LSOURCE
LAB_RESULT_CM	LAB_RESULT_SOURCE	0	225,777,281		0.0	LAB_L3_RSOURCE
PRESCRIBING	ENCOUNTERID	0	77,303,389		0.0	PRES_L3_N
PRESCRIBING	RX_PROVIDERID	0	77,303,389		0.2	PRES_L3_N
PRESCRIBING	RX_ORDER_DATE	0	77,303,389		0.0	XTBL_L3_DATES
PRESCRIBING	RX_ORDER_TIME	0	77,303,389		0.0	XTBL_L3_TIMES
PRESCRIBING	RX_START_DATE	0	77,303,389		0.0	XTBL_L3_DATES
PRESCRIBING	RX_END_DATE	0	77,303,389		16.1	XTBL_L3_DATES
PRESCRIBING	RX_DAYS_SUPPLY	53,094,808	77,303,389	68.7	100.0	PRES_L3_SUPDIST2
PRESCRIBING	RX_FREQUENCY	77,303,389	77,303,389	100.0	61.4	PRES_L3_FREQ
PRESCRIBING	RX_BASIS	65,235,789	77,303,389	84.4	0.0	PRES_L3_BASIS

Table excludes records with values outside of CDM specifications.

The four 'flavors of null' defined in the CDM are combined here but details are available in the source tables.

The denominator is derived from the applicable _N query (ALL_N + NULL_N for TAG=PATID).

Sums which include BT values are marked with an asterisk (*), either next to the sum of the above-threshold values or alone if all values are BT.

Table IVD. Missing or Unknown Values, Optional Tables (continued - page 3 of 6)

This table includes fields in the VITAL, DEATH, LAB_RESULT_CM, PRESCRIBING, DISPENSING, CONDITION, DEATH_CAUSE, PRO_CM, PROVIDER, MED_ADMIN, OBS_CLIN and OBS_GEN tables which are included in the query results and are not required to be populated. Vital measures (e.g. HEIGHT) are not included because the table structure does not support missingness assessment. The table depicts the percentage of records with missing or unknown values. Results support Data Check 3.03 (more than 10% of records have missing or unknown values) for the following fields: RX_ORDER_DATE, VX_RECORD_DATE, DISPENSE_SUP, DEATH_SOURCE, DISPENSE_SOURCE, CONDITION_SOURCE, DX_ORIGIN, LAB_RESULT_SOURCE, MEDADMIN_SOURCE, PRO_SOURCE, PX_SOURCE, RX_SOURCE, VITAL_SOURCE, VX_SOURCE, DIAGNOSIS.ENCOUNTERID, PROCEDURES.ENCOUNTERID, VITAL.ENCOUNTERID, MED_ADMIN.ENCOUNTERID, LAB_RESULT_CM.ENCOUNTERID, MED_ADMIN_CODE, MED_ADMIN_TYPE, OBS_CLIN_CODE, OBS_CLIN_TYPE, OBS_GEN_CODE, OBS_GEN_TYPE. Exceptions are highlighted in blue and should be investigated and explained in the ETL ADD.

Network-wide results for PCORnet DataMarts approved as of Nov 7, 2019 is presented for comparison; a result of 20.0 means that half of the DataMarts had values below 20%.

Table	Field	Records with missing, NI, UN, or OT values				Source table
		Numerator	Denominator	%	50th Percentile	
PRESCRIBING	RXNORM_CUI	0	77,303,389		4.2	PRES_L3_RXCUI
PRESCRIBING	RX_QUANTITY	15,911,517	77,303,389	20.6	57.4	PRES_L3_RXQTY_DIST
PRESCRIBING	RX_REFILLS	77,303,389	77,303,389	100.0	53.0	PRES_L3_RXREFILL_DIST
PRESCRIBING	RX_DOSE_ORDERED	77,303,389	77,303,389	100.0	48.9	PRES_L3_RXDOSEODR_DIST
PRESCRIBING	RX_DOSE_ORDERED_UNIT	77,303,389	77,303,389	100.0	79.0	PRES_L3_RXDOSEODRUNIT
PRESCRIBING	RX_DOSE_FORM	63,978,496	77,303,389	82.8	100.0	PRES_L3_RXDOSEFORM
PRESCRIBING	RX_PRN_FLAG	0	77,303,389		9.4	PRES_L3_PRNFLAG
PRESCRIBING	RX_ROUTE	71,064,613	77,303,389	91.9	80.1	PRES_L3_ROUTE
PRESCRIBING	RX_SOURCE	0	77,303,389		0.0	PRES_L3_SOURCE
PRESCRIBING	RX_DISPENSE_AS_WRITTEN	77,226,453	77,303,389	99.9	100.0	PRES_L3_DISPASWRTN
DISPENSING	PRESCRIBINGID	4,909,610	4,909,610	100.0	100.0	DISP_L3_N
DISPENSING	DISPENSE_SUP	0	4,909,610		0.0	DISP_L3_SUPDIST2
DISPENSING	DISPENSE_DOSE_DISP	4,909,610	4,909,610	100.0	100.0	DISP_L3_DOSE_DIST
DISPENSING	DISPENSE_DOSE_DISP_UNIT	4,909,610	4,909,610	100.0	100.0	DISP_L3_DOSEUNIT
DISPENSING	DISPENSE_ROUTE	4,909,610	4,909,610	100.0	100.0	DISP_L3_ROUTE
DISPENSING	DISPENSE_SOURCE	0	4,909,610		0.0	DISP_L3_SOURCE
CONDITION	ENCOUNTERID	0	16,272,953		21.9	COND_L3_N
CONDITION	REPORT_DATE	0	16,272,953		0.0	XTBL_L3_DATES

Table excludes records with values outside of CDM specifications.

The four 'flavors of null' defined in the CDM are combined here but details are available in the source tables.

The denominator is derived from the applicable _N query (ALL_N + NULL_N for TAG=PATID).

Sums which include BT values are marked with an asterisk (*), either next to the sum of the above-threshold values or alone if all values are BT.

Table IVD. Missing or Unknown Values, Optional Tables (continued - page 4 of 6)

This table includes fields in the VITAL, DEATH, LAB_RESULT_CM, PRESCRIBING, DISPENSING, CONDITION, DEATH_CAUSE, PRO_CM, PROVIDER, MED_ADMIN, OBS_CLIN and OBS_GEN tables which are included in the query results and are not required to be populated. Vital measures (e.g. HEIGHT) are not included because the table structure does not support missingness assessment. The table depicts the percentage of records with missing or unknown values. Results support Data Check 3.03 (more than 10% of records have missing or unknown values) for the following fields: RX_ORDER_DATE, VX_RECORD_DATE, DISPENSE_SUP, DEATH_SOURCE, DISPENSE_SOURCE, CONDITION_SOURCE, DX_ORIGIN, LAB_RESULT_SOURCE, MEDADMIN_SOURCE, PRO_SOURCE, PX_SOURCE, RX_SOURCE, VITAL_SOURCE, VX_SOURCE, DIAGNOSIS.ENCOUNTERID, PROCEDURES.ENCOUNTERID, VITAL.ENCOUNTERID, MED_ADMIN.ENCOUNTERID, LAB_RESULT_CM.ENCOUNTERID, MED_ADMIN_CODE, MED_ADMIN_TYPE, OBS_CLIN_CODE, OBS_CLIN_TYPE, OBS_GEN_CODE, OBS_GEN_TYPE. Exceptions are highlighted in blue and should be investigated and explained in the ETL ADD.

Network-wide results for PCORnet DataMarts approved as of Nov 7, 2019 is presented for comparison; a result of 20.0 means that half of the DataMarts had values below 20%.

Table	Field	Records with missing, NI, UN, or OT values				Source table
		Numerator	Denominator	%	50th Percentile	
CONDITION	RESOLVE_DATE	0	16,272,953		88.5	XTBL_L3_DATES
CONDITION	ONSET_DATE	16,272,953	16,272,953	100.0	100.0	XTBL_L3_DATES
CONDITION	CONDITION_STATUS	0	16,272,953		0.0	COND_L3_STATUS
CONDITION	CONDITION_SOURCE	0	16,272,953		0.0	COND_L3_SOURCE
DEATH_CAUSE	DEATH_CAUSE_CONFIDENCE	1,156	55,584	2.1	2.1	DEATHC_L3_CONF
PROVIDER	PROVIDER_NPI	42,296	103,996	40.7	67.3	PROV_L3_N
PROVIDER	PROVIDER_NPI_FLAG	0	103,996		0.0	PROV_L3_NPIFLAG
PROVIDER	PROVIDER_SPECIALTY_PRIMARY	101,809	103,996	97.9	48.9	PROV_L3_SPECIALTY
PROVIDER	PROVIDER_SEX	49,056	103,996	47.2	68.1	PROV_L3_SEX
MED_ADMIN	ENCOUNTERID	0	1,041,598		0.0	MEDADM_L3_N
MED_ADMIN	PRESCRIBINGID	0	1,041,598		100.0	MEDADM_L3_N
MED_ADMIN	MEDADMIN_PROVIDERID	0	1,041,598		0.1	MEDADM_L3_N
MED_ADMIN	MEDADMIN_DOSE_ADMIN	1,041,598	1,041,598	100.0	2.7	MEDADM_L3_DOSEADM
MED_ADMIN	MEDADMIN_DOSE_ADMIN_UNIT	1,041,598	1,041,598	100.0	25.5	MEDADM_L3_DOSEADMUNIT
MED_ADMIN	MEDADMIN_ROUTE	961,484	1,041,598	92.3	18.4	MEDADM_L3_ROUTE
MED_ADMIN	MEDADMIN_SOURCE	0	1,041,598		0.0	MEDADM_L3_SOURCE
MED_ADMIN	MEDADMIN_CODE	0	1,041,598		2.9	MEDADM_L3_CODE_TYPE
MED_ADMIN	MEDADMIN_TYPE	0	1,041,598		0.0	MEDADM_L3_TYPE

Table excludes records with values outside of CDM specifications.

The four 'flavors of null' defined in the CDM are combined here but details are available in the source tables.

The denominator is derived from the applicable _N query (ALL_N + NULL_N for TAG=PATID).

Sums which include BT values are marked with an asterisk (*), either next to the sum of the above-threshold values or alone if all values are BT.

Table IVD. Missing or Unknown Values, Optional Tables (continued - page 5 of 6)

This table includes fields in the VITAL, DEATH, LAB_RESULT_CM, PRESCRIBING, DISPENSING, CONDITION, DEATH_CAUSE, PRO_CM, PROVIDER, MED_ADMIN, OBS_CLIN and OBS_GEN tables which are included in the query results and are not required to be populated. Vital measures (e.g. HEIGHT) are not included because the table structure does not support missingness assessment. The table depicts the percentage of records with missing or unknown values. Results support Data Check 3.03 (more than 10% of records have missing or unknown values) for the following fields: RX_ORDER_DATE, VX_RECORD_DATE, DISPENSE_SUP, DEATH_SOURCE, DISPENSE_SOURCE, CONDITION_SOURCE, DX_ORIGIN, LAB_RESULT_SOURCE, MEDADMIN_SOURCE, PRO_SOURCE, PX_SOURCE, RX_SOURCE, VITAL_SOURCE, VX_SOURCE, DIAGNOSIS.ENCOUNTERID, PROCEDURES.ENCOUNTERID, VITAL.ENCOUNTERID, MED_ADMIN.ENCOUNTERID, LAB_RESULT_CM.ENCOUNTERID, MED_ADMIN_CODE, MED_ADMIN_TYPE, OBS_CLIN_CODE, OBS_CLIN_TYPE, OBS_GEN_CODE, OBS_GEN_TYPE. Exceptions are highlighted in blue and should be investigated and explained in the ETL ADD.

Network-wide results for PCORnet DataMarts approved as of Nov 7, 2019 is presented for comparison; a result of 20.0 means that half of the DataMarts had values below 20%.

Table	Field	Records with missing, NI, UN, or OT values				Source table
		Numerator	Denominator	%	50th Percentile	
LDS_ADDRESS	ADDRESS_USE	0	7,479,400		---	LDSADRS_L3_ADRSUSE
LDS_ADDRESS	ADDRESS_TYPE	0	7,479,400		---	LDSADRS_L3_ADRSTYPE
LDS_ADDRESS	ADDRESS_PREFERRED	0	7,479,400		---	LDSADRS_L3_ADRSPREF
LDS_ADDRESS	ADDRESS_CITY	3	7,479,400	0.0	---	LDSADRS_L3_ADRSCITY
LDS_ADDRESS	ADDRESS_STATE	0	7,479,400		---	LDSADRS_L3_ADRSSTATE
LDS_ADDRESS	ADDRESS_ZIP5	0	7,479,400		---	LDSADRS_L3_ADRSZIP5
LDS_ADDRESS	ADDRESS_ZIP9	7,479,400	7,479,400	100.0	---	LDSADRS_L3_ADRSZIP9
LDS_ADDRESS	ADDRESS_PERIOD_START	0	7,479,400		---	XTBL_L3_DATES
LDS_ADDRESS	ADDRESS_PERIOD_END	1,599,794	7,479,400	21.4	---	XTBL_L3_DATES
IMMUNIZATION	ENCOUNTERID	10,047,454	11,804,769	85.1	6.7	IMMUNE_L3_N
IMMUNIZATION	PROCEDURESID	11,325,680	11,804,769	95.9	100.0	IMMUNE_L3_N
IMMUNIZATION	VX_PROVIDERID	0	11,804,769		67.5	IMMUNE_L3_N
IMMUNIZATION	VX_RECORD_DATE	11,804,769	11,804,769	100.0	0.0	XTBL_L3_DATES
IMMUNIZATION	VX_ADMIN_DATE	35,823	11,804,769	0.3	0.0	XTBL_L3_DATES
IMMUNIZATION	VX_CODE_TYPE	0	11,804,769		0.0	IMMUNE_L3_CODETYPE
IMMUNIZATION	VX_STATUS	0	11,804,769		0.0	IMMUNE_L3_STATUS
IMMUNIZATION	VX_STATUS_REASON	11,804,769	11,804,769	100.0	100.0	IMMUNE_L3_STATUSREASON
IMMUNIZATION	VX_SOURCE	0	11,804,769		0.0	IMMUNE_L3_SOURCE

Table excludes records with values outside of CDM specifications.

The four 'flavors of null' defined in the CDM are combined here but details are available in the source tables.

The denominator is derived from the applicable _N query (ALL_N + NULL_N for TAG=PATID).

Sums which include BT values are marked with an asterisk (*), either next to the sum of the above-threshold values or alone if all values are BT.

Table IVD. Missing or Unknown Values, Optional Tables (continued - page 6 of 6)

This table includes fields in the VITAL, DEATH, LAB_RESULT_CM, PRESCRIBING, DISPENSING, CONDITION, DEATH_CAUSE, PRO_CM, PROVIDER, MED_ADMIN, OBS_CLIN and OBS_GEN tables which are included in the query results and are not required to be populated. Vital measures (e.g. HEIGHT) are not included because the table structure does not support missingness assessment. The table depicts the percentage of records with missing or unknown values. Results support Data Check 3.03 (more than 10% of records have missing or unknown values) for the following fields: RX_ORDER_DATE, VX_RECORD_DATE, DISPENSE_SUP, DEATH_SOURCE, DISPENSE_SOURCE, CONDITION_SOURCE, DX_ORIGIN, LAB_RESULT_SOURCE, MEDADMIN_SOURCE, PRO_SOURCE, PX_SOURCE, RX_SOURCE, VITAL_SOURCE, VX_SOURCE, DIAGNOSIS.ENCOUNTERID, PROCEDURES.ENCOUNTERID, VITAL.ENCOUNTERID, MED_ADMIN.ENCOUNTERID, LAB_RESULT_CM.ENCOUNTERID, MED_ADMIN_CODE, MED_ADMIN_TYPE, OBS_CLIN_CODE, OBS_CLIN_TYPE, OBS_GEN_CODE, OBS_GEN_TYPE. Exceptions are highlighted in blue and should be investigated and explained in the ETL ADD.

Network-wide results for PCORnet DataMarts approved as of Nov 7, 2019 is presented for comparison; a result of 20.0 means that half of the DataMarts had values below 20%.

Table	Field	Records with missing, NI, UN, or OT values				Source table
		Numerator	Denominator	%	50th Percentile	
IMMUNIZATION	VX_DOSE	2,425,038	11,804,769	20.5	50.4	IMMUNE_L3_DOSE_DIST
IMMUNIZATION	VX_DOSE_UNIT	8,374,650	11,804,769	70.9	62.4	IMMUNE_L3_DOSEUNIT
IMMUNIZATION	VX_ROUTE	11,467,261	11,804,769	97.1	56.0	IMMUNE_L3_ROUTE
IMMUNIZATION	VX_BODY_SITE	4,628,881	11,804,769	39.2	60.6	IMMUNE_L3_BODYSITE
IMMUNIZATION	VX_MANUFACTURER	108	11,804,769	0.0	61.3	IMMUNE_L3_MANUFACTURER
IMMUNIZATION	VX_LOT_NUM	3,816,383	11,804,769	32.3	50.1	IMMUNE_L3_LOTNUM
IMMUNIZATION	VX_EXP_DATE	11,804,769	11,804,769	100.0	69.6	XTBL_L3_DATES

Table excludes records with values outside of CDM specifications.

The four 'flavors of null' defined in the CDM are combined here but details are available in the source tables.

The denominator is derived from the applicable _N query (ALL_N + NULL_N for TAG=PATID).

Sums which include BT values are marked with an asterisk (*), either next to the sum of the above-threshold values or alone if all values are BT.

Table IVE. Principal Diagnoses for Institutional Encounters

This table shows principal diagnosis code data availability for institutional encounters. Results support Data Check 3.06 (more than 10% of IP (inpatient) or ED to inpatient (EI) encounters with any diagnosis don't have a principal diagnosis) and Data Check 2.07 (the average number of principal diagnoses per encounter is above threshold [2.0 for inpatient (IP) and ED to inpatient (EI)]. For data check 3.06, exceptions are triggered when the percentage exceeds 10% or when 0 records have a principal diagnosis. Exceptions are highlighted in blue and should be investigated and explained in the ETL ADD.

Encounter Type	Distinct encounter IDs with a principal diagnosis	Distinct encounter IDs without a principal diagnosis	% of encounters without a principal diagnosis	Principal diagnoses	Principal diagnoses per encounter with any principal diagnosis	Source table
EI (ED to IP Stay)	0	0		0		DIA_L3_PDX_ENCTYPE; DIA_L3_PDXGRP_ENCTYPE
IP (Inpatient Hospital Stay)	523,134	259,840	33.2	886,104	1.7	DIA_L3_PDX_ENCTYPE; DIA_L3_PDXGRP_ENCTYPE
IS (Non-acute Institutional Stay)	93,716	1,228	1.3	105,517	1.1	DIA_L3_PDX_ENCTYPE; DIA_L3_PDXGRP_ENCTYPE
OS (Observation Stay)	35,955	1,102	3.0	55,748	1.6	DIA_L3_PDX_ENCTYPE; DIA_L3_PDXGRP_ENCTYPE

Table excludes records with values outside of CDM specifications.

Sums which include BT values are marked with an asterisk (*), either next to the sum of the above-threshold values or alone if all values are below threshold.

Percentage is calculated by dividing the number of encounters (i.e. distinct ENCOUNTERIDs) without a principal diagnosis by the total number of encounters in the DIAGNOSIS table.

The number of encounters without a principal diagnosis is obtained by identifying distinct encounterids which are not in a list of distinct encounterIDs where PDX=P.

Number of principal diagnoses per encounter with any principal diagnosis is calculated by dividing the number of principal diagnoses by the number of encounters with a principal diagnosis.

Table IVF. Data Latency and Completeness of Encounter, Diagnoses, and Procedures, Past 2 Years

This table includes ENCOUNTER, DIAGNOSIS, and PROCEDURES from the most recent 24 month period; month -0 is the month the data curation query was run. Data completeness is determined by comparing the actual volume to the expected volume in each month. Expected volume is determined by taking the average volume during the benchmark period of months -12 to month -23. Data completeness is reported as a percentage of the benchmark average. Temporal differences may be affected by data availability, ETL processes, date shifting, secular trends, and/or changes in data provenance.

These data support Data Check 3.07 (encounters, diagnoses, or procedures in an ambulatory (AV), emergency department (ED), ED to inpatient (EI), or inpatient (IP) setting are less than 75% complete three months prior to the current month). Data check exceptions occur if the month -3 result is <75% of the benchmark average or 0 records. Data check exceptions are highlighted in blue. Data check exceptions and unexpected results (e.g. significant discrepancies in data completeness between the tables) should be investigated and explained in the ETL ADD.

Month	Calendar Month	Ambulatory, ED, Inpatient or ED-to-Inpatient encounters		Ambulatory, ED, Inpatient or ED-to-Inpatient diagnoses		Ambulatory, ED, Inpatient or ED-to-Inpatient procedures	
		Records	Percent of benchmark average	Records	Percent of benchmark average	Records	Percent of benchmark average
Month -0	01/2020	0		0		0	
Month -1	12/2019	0		0		0	
Month -2	11/2019	192,673	60.4	270,247	43.1	167,065	42.7
Month -3	10/2019	375,303	117.7	753,427	120.3	437,780	111.8
Month -4	09/2019	304,555	95.5	619,563	98.9	372,608	95.2
Month -5	08/2019	319,235	100.1	646,677	103.2	392,538	100.2
Month -6	07/2019	324,666	101.8	652,810	104.2	398,111	101.7
Month -7	06/2019	308,909	96.9	615,720	98.3	367,136	93.8
Month -8	05/2019	344,135	107.9	690,008	110.1	413,813	105.7
Month -9	04/2019	334,235	104.8	661,707	105.6	409,167	104.5
Month -10	03/2019	313,879	98.4	612,641	97.8	376,500	96.1
Month -11	02/2019	273,109	85.6	535,610	85.5	332,758	85.0

Cell counts between 1 and the low-cell count threshold are displayed as BT and treated as missing.

Table IVF. Data Latency and Completeness of Encounter, Diagnoses, and Procedures, Past 2 Years (continued)

This table includes ENCOUNTER, DIAGNOSIS, and PROCEDURES from the most recent 24 month period; month -0 is the month the data curation query was run. Data completeness is determined by comparing the actual volume to the expected volume in each month. Expected volume is determined by taking the average volume during the benchmark period of months -12 to month -23. Data completeness is reported as a percentage of the benchmark average. Temporal differences may be affected by data availability, ETL processes, date shifting, secular trends, and/or changes in data provenance.

These data support Data Check 3.07 (encounters, diagnoses, or procedures in an ambulatory (AV), emergency department (ED), ED to inpatient (EI), or inpatient (IP) setting are less than 75% complete three months prior to the current month). Data check exceptions occur if the month -3 result is <75% of the benchmark average or 0 records. Data check exceptions are highlighted in blue. Data check exceptions and unexpected results (e.g. significant discrepancies in data completeness between the tables) should be investigated and explained in the ETL ADD.

Month	Calendar Month	Ambulatory, ED, Inpatient or ED-to-Inpatient encounters		Ambulatory, ED, Inpatient or ED-to-Inpatient diagnoses		Ambulatory, ED, Inpatient or ED-to-Inpatient procedures	
		Records	Percent of benchmark average	Records	Percent of benchmark average	Records	Percent of benchmark average
Benchmark Period							
Month -12	01/2019	316,751	99.3	622,632	99.4	385,126	98.3
Month -13	12/2018	292,098	91.6	565,718	90.3	353,803	90.4
Month -14	11/2018	323,725	101.5	628,375	100.3	391,879	100.1
Month -15	10/2018	374,540	117.5	719,585	114.9	441,984	112.9
Month -16	09/2018	289,470	90.8	560,502	89.5	354,168	90.4
Month -17	08/2018	331,316	103.9	633,841	101.2	407,889	104.2
Month -18	07/2018	309,149	96.9	610,131	97.4	386,875	98.8
Month -19	06/2018	315,570	99.0	629,562	100.5	393,510	100.5
Month -20	05/2018	342,852	107.5	685,354	109.4	428,587	109.4
Month -21	04/2018	318,210	99.8	634,088	101.2	398,990	101.9
Month -22	03/2018	319,700	100.3	635,387	101.4	394,576	100.8
Month -23	02/2018	293,276	92.0	592,539	94.6	361,665	92.4
Benchmark average		318,888		626,476		391,588	
Source table		ENC_L3_ ENCTYPE_ ADATE_YM		DIA_L3_ ENCTYPE_ ADATE_YM		PRO_L3_ ENCTTPE_ ADATE_YM	

Table IVG. Data Latency and Completeness of Vital, Prescribing, and Lab Data, Past 2 Years

This table includes VITAL, PRESCRIBING, and LAB_RESULT_CM data from the most recent 24 month period; month -0 is the month the data curation query was run. Data completeness is determined by comparing the actual volume to the expected volume in each month. Expected volume is determined by taking the average volume during the benchmark period of months -12 to month -23. Data completeness is reported as a percentage of the benchmark average. Temporal differences may be affected by data availability, ETL processes, date shifting, secular trends, and/or changes in data provenance.

These data support Data Check 3.11 (vital, prescribing, or laboratory records are less than 75% complete three months prior to the current month). Data check exceptions occur if the month -3 result is <75% of the benchmark average or 0 records. Data check exceptions are highlighted in blue. Data check exceptions and unexpected results should be investigated and explained in the ETL ADD.

Month	Calendar Month	Vitals		Prescriptions		Labs	
		Records	Percent of benchmark average	Records	Percent of benchmark average	Records	Percent of benchmark average
Month -0	01/2020	0		0		0	
Month -1	12/2019	0		0		0	
Month -2	11/2019	55,819	53.0	311,376	55.0	1,041,007	56.0
Month -3	10/2019	146,673	139.3	704,726	124.5	2,271,976	122.3
Month -4	09/2019	124,501	118.2	595,631	105.3	2,020,071	108.7
Month -5	08/2019	130,729	124.1	616,382	108.9	2,048,936	110.3
Month -6	07/2019	127,667	121.2	618,309	109.3	2,101,920	113.1
Month -7	06/2019	101,748	96.6	528,698	93.4	1,821,832	98.0
Month -8	05/2019	114,472	108.7	596,816	105.5	2,027,775	109.1
Month -9	04/2019	112,625	107.0	600,788	106.2	1,973,248	106.2
Month -10	03/2019	102,425	97.3	561,652	99.2	1,876,678	101.0
Month -11	02/2019	88,478	84.0	501,460	88.6	1,654,889	89.1

Cell counts between 1 and the low-cell count threshold are displayed as BT and treated as missing.

Table IVG. Data Latency and Completeness of Vital, Prescribing, and Lab Data, Past 2 Years (continued)

This table includes VITAL, PRESCRIBING, and LAB_RESULT_CM data from the most recent 24 month period; month -0 is the month the data curation query was run. Data completeness is determined by comparing the actual volume to the expected volume in each month. Expected volume is determined by taking the average volume during the benchmark period of months -12 to month -23. Data completeness is reported as a percentage of the benchmark average. Temporal differences may be affected by data availability, ETL processes, date shifting, secular trends, and/or changes in data provenance.

These data support Data Check 3.11 (vital, prescribing, or laboratory records are less than 75% complete three months prior to the current month). Data check exceptions occur if the month -3 result is <75% of the benchmark average or 0 records. Data check exceptions are highlighted in blue. Data check exceptions and unexpected results should be investigated and explained in the ETL ADD.

Month	Calendar Month	Vitals		Prescriptions		Labs	
		Records	Percent of benchmark average	Records	Percent of benchmark average	Records	Percent of benchmark average
Benchmark Period							
Month -12	01/2019	104,845	99.6	581,515	102.8	1,888,157	101.6
Month -13	12/2018	94,130	89.4	519,913	91.9	1,752,666	94.3
Month -14	11/2018	105,567	100.2	559,973	99.0	1,846,488	99.4
Month -15	10/2018	117,665	111.7	604,565	106.8	1,992,539	107.2
Month -16	09/2018	93,950	89.2	496,004	87.6	1,727,787	93.0
Month -17	08/2018	108,569	103.1	554,245	97.9	1,920,427	103.3
Month -18	07/2018	101,978	96.8	543,815	96.1	1,884,305	101.4
Month -19	06/2018	105,764	100.4	560,538	99.1	1,861,863	100.2
Month -20	05/2018	116,823	110.9	619,890	109.5	1,983,110	106.7
Month -21	04/2018	107,126	101.7	591,420	104.5	1,864,761	100.3
Month -22	03/2018	107,638	102.2	599,125	105.9	1,854,141	99.8
Month -23	02/2018	99,614	94.6	559,787	98.9	1,723,577	92.7
Benchmark average		105,306		565,899		1,858,318	
Source table		VIT_L3_ MDATE_YM		PRES_L3_ ODATE_YM		LAB_L3_ RDATE_YM	

Cell counts between 1 and the low-cell count threshold are displayed as BT and treated as missing.

Table IVH. RXNORM Term Type Mapping, Overall and Past 5 Years

This table shows the number of records in the PRESCRIBING table by RXNORM Term Type tier. The percentage mapped to Tier 1 brand drugs (a subset of the Tier 1 percentage) is also displayed. Guidance on mapping prescribing orders to RXNORM is provided in the CDM specifications. These data support Data Check 3.08 (less than 80% of prescribing orders are mapped to a RXNORM_CUI which fully specifies the ingredient, strength and dose form). Data check results are shown for all years and for the most recent 5 years. Data check exceptions occur if the Tier 1 percentage is <80% or the numerator is 0. Data check exceptions are highlighted in blue and should be investigated and explained in the ETL ADD. Data check exceptions for the most recent 5 year period will be used to assess compliance with the terms of the PCRF Scope of Work.

Period	Term Type Tier	Term Type Tier Description	Term Types	Numerator	Percentage	Source table
All years	Tier 1	RXNORM_CUI encodes ingredient(s), strength and dose form	SCD, SBD, BPCCK, and GPCK	74,220,675	96.01	PRES_L3_RXCUI_TIER
	Tier 1 brand	RXNORM_CUI encodes ingredient(s), strength, dose form and brand	SBD, BPCCK	95,984	0.12	PRES_L3_RXCUI
	Tier 2	RXNORM_CUI encodes ingredient(s) and potentially strength or dose form. Can still represent medications with multiple ingredients with a single RXCUI.	SBDF, SCDF, SBDG, SCDG, SBDC, BN, and MIN	45,503	0.06	PRES_L3_RXCUI_TIER
	Tier 3	Requires more than one RXNORM_CUI to represent medications with multiple ingredients.	SCDC, PIN, and IN	3,008,337	3.89	PRES_L3_RXCUI_TIER
	Tier 4	RXNORM_CUI does not encode any ingredient information.	DF and DFG	0		PRES_L3_RXCUI_TIER
	Unknown	RXNORM_CUI was not populated or could not be matched to the reference table	n/a	28,874	0.04	PRES_L3_RXCUI_TIER
Most Recent Five Years	Tier 1	RXNORM_CUI encodes ingredient(s), strength and dose form	SCD, SBD, BPCCK, and GPCK	37,035,329	96.29	PRES_L3_RXCUI_TIER_5Y

Cell counts between 1 and the low-cell count threshold are displayed as BT and treated as missing.

Term types were obtained from the rxnorm_cui_ref table. This table was created from the 2010-2019 Unified Medical Language System® (UMLS®) release files. It includes all RXCUIs from the RXNORM terminology (i.e., SAB=RXNORM) with one of the following term types: BN (Brand Name), BPCCK (Brand Name Pack), DF (Dose Form), DFG (Dose Form Group), GPCK (Generic Pack), IN (Ingredient), MIN (Multiple Ingredients), SBD (Semantic Branded Drug), SBDC (Semantic Branded Drug Component), SBDF (Semantic Branded Drug Form), SBDG (Semantic Branded Dose Form Group), SCD (Semantic Clinical Drug), SCDC (Semantic Clinical Drug Component), SCDF (Semantic Clinical Drug Form), SCDG (Semantic Clinical Dose Form Group), or PIN (Precise Ingredient).

Table IVI. Laboratory Result Data Completeness, Overall and Past 5 Years

This table shows the level of data completeness for LAB_RESULT_CM records and supports Data Check 3.09 (less than 80% of laboratory results are mapped to LAB_LOINC), Data Check 3.10 (less than 80% of quantitative results for tests mapped to LAB_LOINC fully specify the normal range), and Data Check 3.12 (less than 80% of quantitative results for tests mapped to LAB_LOINC fully specify the specimen source and result unit). Data check exceptions occur if the percentages are <80% or the numerator is 0. Data check results are shown for all years and for the most recent 5 years. Data check exceptions are highlighted in blue and should be investigated and explained in the ETL ADD. Data check exceptions for the most recent 5 year period will be used to assess compliance with the terms of the PCRf Scope of Work.

Period	Data Check	Description	Numerator	Denominator	Percentage	Source table
All years	n/a	Number of distinct LAB_LOINCs	2,179			LAB_L3_LOINC
	3.09	Results mapped to a known LAB_LOINC	225,777,281	225,777,281	100.00	LAB_L3_RECORDC; LAB_L3_N
	n/a	Results mapped to a known LAB_LOINC with a known result	225,777,281	225,777,281	100.00	LAB_L3_RECORDC
	n/a	Quantitative results	205,562,725			LAB_L3_RECORDC
	n/a	Quantitative results which specify the specimen source	136,374,416	205,562,725	66.34	LAB_L3_RECORDC
	n/a	Quantitative results which specify the result unit	172,486,404	205,562,725	83.91	LAB_L3_RECORDC
	3.10	Quantitative results which fully specify the normal range	158,366,128	205,562,725	77.04	LAB_L3_RECORDC
	3.12	Quantitative results which specify the specimen source and result unit	131,449,698	205,562,725	63.95	LAB_L3_RECORDC
Most Recent Five Years	3.10	Quantitative results which fully specify the normal range	75,240,322	101,803,562	73.91	LAB_L3_RECORDC_5Y
	3.12	Quantitative results which specify the specimen source and result unit	64,478,105	101,803,562	63.34	LAB_L3_RECORDC_5Y

The denominator for the percentage of results mapped to a known LAB_LOINC is the total number of records in the table (ALL_N + NULL_N for the LAB_RESULT_CM_ID field). Cell counts between 1 and the low-cell count threshold are displayed as BT and treated as missing. The criteria for each measurement are described in table IVI_REF.

Table IVI_Ref. Laboratory Result Data Completeness Definitions
This table shows the definitions for each of the numerator values in Table IVI.

Description	Source Table Row	LAB_LOINC criteria	RESULT_NUM criteria	RESULT_MODIFIER criteria	Other criteria
Number of distinct LAB_LOINCs		Not null			
Results mapped to a known LAB_LOINC	KNOWN_TEST	Not null			
Results mapped to a known LAB_LOINC with a known result	KNOWN_TEST_RESULT	Not null	Not null		RESULT_MODIFIER is in ('EQ','GE','GT','LE','LT') or RESULT_QUAL is in ("BORDERLINE", "POSITIVE", "NEGATIVE" or "UNDETERMINED")
Quantitative results	KNOWN_TEST_RESULT_NUM	Not null	Not null	'EQ','GE','GT','LE', or 'LT'	
Quantitative results which specify the specimen source	KNOWN_TEST_RESULT_NUM_SOURCE	Not null	Not null	'EQ','GE','GT','LE', or 'LT'	SPECIMEN_SOURCE is not in ('NI','UN','OT', 'SUB', 'SMPLS', 'SPECIMEN', null).
Quantitative results which specify the result unit	KNOWN_TEST_RESULT_NUM_UNIT	Not null	Not null	'EQ','GE','GT','LE', or 'LT'	RESULT_UNIT is not in ('NI','UN','OT', null)
Quantitative results which fully specify the normal range.	KNOWN_TEST_NUM_RESULT_RANGE	Not null	Not null	'EQ','GE','GT','LE', or 'LT'	[NORM_MODIFIER_LOW='EQ' and NORM_MODIFIER_HIGH='EQ' and NORM_RANGE_LOW is not null and NORM_RANGE_HIGH is not null] or [NORM_MODIFIER_LOW in ('GT','GE') and NORM_MODIFIER_HIGH='NO' and NORM_RANGE_LOW is not null and NORM_RANGE_HIGH is null] or [NORM_MODIFIER_HIGH in ('LE','LT') and NORM_MODIFIER_LOW='NO' and NORM_RANGE_HIGH is not null and NORM_RANGE_LOW is null].
Quantitative results which specify the specimen source and result unit	KNOWN_TEST_RESULT_NUM_SRCE_UNIT	Not null	Not null	'EQ','GE','GT','LE', or 'LT'	RESULT_UNIT is not a flavor of null and SPECIMEN_SOURCE is not in ('NI','UN','OT', 'SUB', 'SMPLS', 'SPECIMEN', null).

Table VA. Changes in Tables

This table shows changes in key DataMart attributes between the most recent approved DataMart refresh and the current DataMart refresh and supports Data Check 4.01 (more than a 5% decrease in the number of patients or records in a CDM table). Data check exceptions are highlighted in blue and should be investigated and explained in the ETL ADD.

Table	Records			Patients			Source table
	Previous Refresh	Current Refresh	Percent Change	Previous Refresh	Current Refresh	Percent Change	
DEMOGRAPHIC	955,965	966,870	1.1	955,965	966,870	1.1	DEM_L3_N
ENROLLMENT	955,965	966,868	1.1	955,965	966,868	1.1	ENR_L3_N
ENCOUNTER	65,812,602	66,662,590	1.3	943,471	956,583	1.4	ENC_L3_N
DIAGNOSIS	105,278,628	98,986,521	-6.0	926,525	934,427	0.9	DIA_L3_N
PROCEDURES	55,646,667	53,609,004	-3.7	796,258	806,224	1.3	PRO_L3_N
VITAL	10,485,308	12,713,955	21.3	678,147	648,554	-4.4	VIT_L3_N
DEATH	64,297	65,774	2.3	64,297	65,774	2.3	DEATH_L3_N
PRESCRIBING	51,171,314	77,303,389	51.1	775,396	796,553	2.7	PRES_L3_N
DISPENSING	4,363,713	4,909,610	12.5	291,898	295,989	1.4	DISP_L3_N
LAB_RESULT_CM	206,715,926	225,777,281	9.2	789,751	797,266	1.0	LAB_L3_N
CONDITION	15,911,312	16,272,953	2.3	755,748	762,745	0.9	COND_L3_N
DEATH_CAUSE	55,608	55,584	-0.0	14,989	14,982	-0.0	DEATHC_L3_N
PROVIDER	106,152	103,996	-2.0	---	---		PROV_L3_N
MED_ADMIN	1,387,172	1,041,598	-24.9	112,737	123,029	9.1	MEDADM_L3_N
IMMUNIZATION	11,307,158	11,804,769	4.4	731,791	738,870	1.0	IMMUNE_L3_N
LDS_ADDRESS_HISTORY	7,374,420	7,479,400	1.4	944,383	955,110	1.1	LDSADRS_L3_N

Cell counts between 1 and the low-cell count threshold are displayed as BT and treated as missing.

Table VB. Changes in Selected Encounter Types and Domains

This table shows changes in key DataMart attributes between the most recent approved DataMart refresh and the current DataMart refresh and supports Data Check 4.02 (more than a 5% decrease in the number of patients with diagnosis, procedures, labs or prescriptions during an AV, ED, or IP encounter). Data check exceptions are highlighted in blue and should be investigated and explained in the ETL ADD.

Encounter Type	Records			Patients			Source table
	Previous Refresh	Current Refresh	Percent Change	Previous Refresh	Current Refresh	Percent Change	
Ambulatory Visit (AV)		---			---		
Diagnoses	76,802,280	85,292,213	11.1	874,973	905,215	3.5	DIA_L3_ENCTYPE
Procedures	48,825,939	49,319,785	1.0	759,622	767,054	1.0	PRO_L3_ENCTYPE
Labs	120,174,856	139,472,403	16.1	729,413	701,737	-3.8	XTBL_L3_LAB_ENCTYPE
Prescriptions	31,165,757	45,704,016	46.6	738,445	651,855	-11.7	XTBL_L3_PRES_ENCTYPE
Emergency Department (ED)		---			---		
Diagnoses	2,278,866	3,742,535	64.2	234,592	280,414	19.5	DIA_L3_ENCTYPE
Procedures	1,854,602	2,014,832	8.6	225,068	237,985	5.7	PRO_L3_ENCTYPE
Labs	9,130,823	7,064,680	-22.6	110,982	98,772	-11.0	XTBL_L3_LAB_ENCTYPE
Prescriptions	450,700	611,007	35.6	89,392	91,304	2.1	XTBL_L3_PRES_ENCTYPE
Inpatient (IP)		---			---		
Diagnoses	13,916,699	4,985,450	-64.2	210,294	205,714	-2.2	DIA_L3_ENCTYPE
Procedures	4,235,130	1,009,383	-76.2	199,926	129,558	-35.2	PRO_L3_ENCTYPE
Labs	6,935,802	22,206,313	220.2	119,605	114,603	-4.2	XTBL_L3_LAB_ENCTYPE
Prescriptions	1,417,815	3,598,353	153.8	71,934	109,463	52.2	XTBL_L3_PRES_ENCTYPE
Other Ambulatory (OA)		---			---		
Diagnoses	12,279,299	1,503,780	-87.8	419,451	62,527	-85.1	DIA_L3_ENCTYPE
Procedures	729,090	200,049	-72.6	48,881	22,248	-54.5	PRO_L3_ENCTYPE
Labs	70,463,770	54,743,645	-22.3	401,629	358,214	-10.8	XTBL_L3_LAB_ENCTYPE
Prescriptions	18,130,556	26,862,957	48.2	535,467	571,297	6.7	XTBL_L3_PRES_ENCTYPE

Cell counts between 1 and the low-cell count threshold are displayed as BT and treated as missing.

Table VC. Changes in Selected Code Types

This table shows changes in key DataMart attributes between the most recent approved DataMart refresh and the current DataMart refresh and supports Data Check 4.03 (more than a 5% decrease in the number of records or distinct codes for selected code types). Data check exceptions are highlighted in blue and should be investigated and explained in the ETL ADD.

	Records			Distinct Codes			Source table
	Previous Refresh	Current Refresh	Percent Change	Previous Refresh	Current Refresh	Percent Change	
DIAGNOSIS table							
09	68,094,275	62,750,199	-7.8	12,181	12,161	-0.2	DIA_L3_DX_DXTYPE
10	37,184,353	36,236,322	-2.5	29,455	29,972	1.8	DIA_L3_DX_DXTYPE
PROCEDURES table							
09	392,506	340,393	-13.3	2,581	2,591	0.4	PRO_L3_PX_PXTYPE
10	155,198	95,295	-38.6	7,622	6,100	-20.0	PRO_L3_PX_PXTYPE
CH	55,098,963	53,173,316	-3.5	7,054	6,971	-1.2	PRO_L3_PX_PXTYPE