Work Plan: Data Characterization Query Package

February 3, 2016
v3.00

TABLE OF CONTENTS

I. PURPOSE AND SCOPE 3
II. QUERY LIST 3
III. PROGRAM PACKAGE FILE STRUCTURE 5
IV. FILES INCLUDED IN QUERY REQUEST 6
V. OUTPUT FILES 6
VI. TABLE SHELLS: DEMOGRAPHIC QUERIES 7
VII. TABLE SHELLS: ENCOUNTER QUERIES 9
VIII. TABLE SHELLS: DIAGNOSIS QUERIES 15
IX. TABLE SHELLS: PROCEDURES QUERIES 19
X. TABLE SHELLS: ENROLLMENT QUERIES 22
XI. TABLE SHELLS: VITAL QUERIES 24
XII. TABLE SHELLS: HARVEST AND CROSS-TABLE QUERIES 31
XIII. RUNNING THE SAS QUERY 34

I. Purpose and Scope

The purpose of this document is to describe PCORnet's Data Characterization Query Package v3.00. This query package consists of 77 queries against seven PCORnet Common Data Model (CDM) v3.0 tables: DEMOGRAPHIC, ENROLLMENT, ENCOUNTER, DIAGNOSIS, PROCEDURES, VITAL, and HARVEST. Output tables will be produced by running SAS programs against static local DataMarts in PCORnet CDM v3.0 with SAS data types. An ETL Annotated Data Dictionary (ADD) must be submitted along with the query results.

Low Cell Count Threshold

DataMart Administrators may specify a low cell count threshold which establishes the minimum number of observations required to protect against possible identification of subject data. The default low cell count threshold value is zero (0). If the threshold is changed, query results greater than zero and less than the threshold will be changed to BT (below threshold). For example, if a DataMart sets a low cell count threshold of 11, cell counts between 1 and 10 will be changed to BT. The low cell count threshold applies to all query results except for (a) counts where the value of the specified field is null or missing, e.g. 10 records where RACE is null, (b) descriptive statistics about a field, e.g. minimum AGE=0, and (c) table-level counts, e.g. count of PATIDs in the DEMOGRAPHIC table. The low cell count threshold treatment for each query is shown in the Query List below. PCORnet's Distributed Research Network Operations Center (DRN OC) recommends setting low cell count thresholds no higher than 50.

Questions about this query package should be sent to Laura Qualls (laura.qualls@duke.edu).

II. Query List

PCORnet Table(s)	Query	Query Description	Low cell count threshold?
DEMOGRAPHIC	dem_l3_n	Counts PATID	No
DEMOGRAPHIC	dem_l3_ageyrsdist1	Descriptive statistics for age	No
DEMOGRAPHIC	dem_l3_ageyrsdist2	Age group frequency	Yes
DEMOGRAPHIC	dem_l3_hispdist	HISPANIC frequency	Yes
DEMOGRAPHIC	dem_l3_racedist	RACE frequency	Yes
DEMOGRAPHIC	dem_l3_sexdist	SEX frequency	Yes
ENCOUNTER	enc_l3_n	Counts PATID, ENCOUNTERID, and PROVIDERID	No
ENCOUNTER	enc_l3_n_visit	Counts visits (combination of PATID + PROVIDERID + ENC_TYPE)	No
ENCOUNTER	enc_l3_admsrc	ADMITTING_SOURCE frequency	Yes
ENCOUNTER	enc_l3_enctype_admsrc	ENC_TYPE by ADMITTING_SOURCE crosstab	Yes
ENCOUNTER	enc_l3_adate_y	ADMIT_DATE year frequency	Yes
ENCOUNTER	enc_l3_adate_ym	ADMIT_DATE year month frequency	Yes
ENCOUNTER	enc_l3_enctype_adate_ym	ENC_TYPE and ADMIT_DATE year month crosstab	Yes
ENCOUNTER	enc_l3_ddate_y	DISCHARGE_DATE year frequency	Yes
ENCOUNTER	enc_l3_ddate_ym	DISCHARGE_DATE year month frequency	Yes
ENCOUNTER	enc_l3_enctype_ddate_ym	ENC_TYPE and DISCHARGE_DATE year month crosstab	Yes
ENCOUNTER	enc_l3_disdisp	DISCHARGE_DISPOSITION frequency	Yes

1. PDF file limited to the 100 most frequent observations with counts above the low cell count threshold and sorted by descending record count.

PCORnet Table(s)	Query	Query Description	Low cell count threshold?
ENCOUNTER	enc_l3_enctype_disdisp	ENC_TYPE and DISCHARGE_DISPOSITION crosstab	Yes
ENCOUNTER	enc_l3_disstat	DISCHARGE_STATUS frequency	Yes
ENCOUNTER	enc_l3_enctype_disstat	ENC_TYPE and DISCHARGE_STATUS crosstab	Yes
ENCOUNTER	enc_l3_drg	DRG frequency	Yes
ENCOUNTER	enc_l3_drg_type	DRG_type frequency	Yes
ENCOUNTER	enc_l3_enctype_drg	ENC_TYPE and DRG_TYPE crosstab	Yes
ENCOUNTER	enc_l3_enctype	ENC_TYPE frequency	Yes
ENCOUNTER	enc_l3_dash1	Patients with encounters in the past 1, 2, 3, 4, or 5 years	Yes
ENCOUNTER	enc_l3_dash2	Patients with IP, ED, EI, or AV encounters in the past 1, 2, 3, 4, or 5 years	Yes
DIAGNOSIS	dia_l3_n	Counts PATID, ENCOUNTERID, and DIAGNOSISID	No
DIAGNOSIS	dia_l3_dx	DX frequency	Yes
DIAGNOSIS	dia_l3_dx_dxtype	DX and DX_TYPE crosstab	Yes
DIAGNOSIS	dia_l3_dxsource	DX_SOURCE frequency	Yes
DIAGNOSIS	dia_l3_dxtype_dxsource	DX_TYPE and DX_SOURCE crosstab	Yes
DIAGNOSIS	dia_l3_PDX	PDX frequency	Yes
DIAGNOSIS	dia_l3_PDX_enctype	PDX and ENC_TYPE crosstab	Yes
DIAGNOSIS	dia_l3_adate_y	ADMIT_DATE year frequency	Yes
DIAGNOSIS	dia_l3_adate_ym	ADMIT_DATE year month frequency	Yes
DIAGNOSIS	dia_l3_enctype	ENC_TYPE frequency	Yes
DIAGNOSIS	dia_l3_dxtype_enctype	DX_TYPE and ENC_TYPE crosstab	Yes
DIAGNOSIS	dia_l3_enctype_adate_ym	ENC_TYPE and ADMIT_DATE year month crosstab	Yes
DIAGNOSIS	dia_l3_dash1	Patients with diagnoses in the past 1, 2, 3, 4, or 5 years.	Yes
PROCEDURES	pro_l3_n	Counts PATID, ENCOUNTERID, and PROCEDURESID	No
PROCEDURES	pro_l3_px	PX frequency	Yes
PROCEDURES	pro_l3_adate_y	ADMIT_DATE year frequency	Yes
PROCEDURES	pro_l3_adate_ym	ADMIT_DATE year month frequency	Yes
PROCEDURES	pro_l3_pxdate_y	PX_DATE year frequency	Yes
PROCEDURES	pro_l3_px_enctype	PX by ENC_TYPE crosstab	Yes
PROCEDURES	pro_l3_enctype	ENC_TYPE frequency	Yes
PROCEDURES	pro_l3_pxtype_enctype	PX_TYPE and ENC_TYPE crosstab	Yes
PROCEDURES	pro_l3_enctype_adate_ym	ENC_TYPE and ADMIT_DATE year month crosstab	Yes
PROCEDURES	pro_l3_px_pxtype	PX and PX_TYPE crosstab	Yes
PROCEDURES	pro_l3_pxsource	PX_SOURCE frequency	Yes
ENROLLMENT	enr_l3_n	Counts PATID, ENR_START_DATE, and ENROLLID (combination of PATID, ENR_START_DATE, and ENR_BASIS)	No
ENROLLMENT	enr_l3_dist_start	Descriptive statistics for distinct ENR_START_DATE	No
ENROLLMENT	enr_l3_dist_end	Descriptive statistics for distinct ENR_END_DATE	No

1. PDF file limited to the 100 most frequent observations with counts above the low cell count threshold and sorted by descending record count.

PCORnet Table(s)	Query	Query Description	Low cell count threshold?
ENROLLMENT	enr_l3_dist_enrmonth	Distinct number of enrollment months frequency	Yes
ENROLLMENT	enr_l3_dist_enryear	Distinct number of enrollment years frequency	Yes
ENROLLMENT	enr_l3_enr_ym	ENR_START_DATE frequency	Yes
ENROLLMENT	enr_l3_basisdist	ENR_BASIS frequency	Yes
ENROLLMENT	enr_l3_per_patid	Descriptive statistics for number of enrollment periods per PATID	No
VITAL	vit_l3_n	Counts PATID, ENCOUNTERID, and VITALID	No
VITAL	vit_l3_mdate_y	MEASURE_DATE year frequency	Yes
VITAL	vit_l3_mdate_ym	MEASURE_DATE year month frequency	Yes
VITAL	vit_l3_vital_source	VITAL_SOURCE frequency	Yes
VITAL	vit_l3_ht	HT frequency	Yes
VITAL	vit_l3_ht_dist	Descriptive statistics for HT	No
VITAL	vit_l3_wt	WT frequency	Yes
VITAL	vit_l3_wt_dist	Descriptive statistics for WT	No
VITAL	vit_l3_diastolic	DIASTOLIC frequency	Yes
VITAL	vit_l3_systolic	SYSTOLIC frequency	Yes
VITAL	vit_l3_bmi	BMI frequency	Yes
VITAL	vit_l3_bp_position_type	BP POSITION_TYPE frequency	Yes
VITAL	vit_l3_smoking	SMOKING frequency	Yes
VITAL	vit_l3_tobacco	TOBACCO frequency	Yes
VITAL	vit_l3_tobacco_type	TOBACCO_TYPE frequency	Yes
VITAL	vit_l3_dash1	Patients with vital records in the past 1, 2, 3, 4 or 5 years.	Yes
ENCOUNTER, DIAGNOSIS, PROCEDURES, VITAL, ENROLLMENT	xtbl_l3_dates	Counts the minimum and maximum date and the number of records with non-missing dates, future dates, and dates prior to Jan 2010 for the following date fields: DEMOGRAPHIC.BIRTH_DATE, ENCOUNTER.ADMIT_DATE , ENCOUNTER.DISCHARGE_DATE, DIAGNOSIS.ADMIT_DATE, PROCEDURES.ADMIT_DATE, VITAL.MEASURE_DATE, ENROLLMENT. ENR_START_DATE, and ENROLLMENT.ENR_END_DATE	Yes for record counts; No for descriptive statistics
HARVEST	xtbl_l3_metadata	HARVEST table information, query package, response date, low cell count thresholds, operating system, and SAS version and packages	No
DIAGNOSIS, ENCOUNTER	xtbl_l3_dash1	Patients with at least 1 DX and 1 vital MEASURE_DATE in past 1, 2, 3, 4 or 5 years	Yes

III. Program Package File Structure

Each request package distributed by PCORnet's DRN OC contains several sub-folders to organize program inputs and outputs. The subfolders must reside within an outer folder labeled with the query name

1. PDF file limited to the 100 most frequent observations with counts above the low cell count threshold and sorted by descending record count.

designated in the DRN Query Tool, e.g. PROD_P02_DQA_FDPRO_DCQ_NSD _r001_v01.
The subfolders are as follows:

- *dmlocal*: Contains output generated by the request that should be saved locally but not returned to DRN OC. Output may be used locally or to facilitate follow-up queries.
- *drnoc*: Contains output generated by the request that should be returned to the DRN OC via the PCORnet DRN Query Tool. These tables consist of aggregate data/output and transfer the minimum required to answer the analytic question.
- *sasprograms*: Contains the master SAS program that must be edited and then executed locally.
- *infolder:* Contains all input programs and files needed to execute the request. These are created for each request by the DRN OC Data Characterization team; the contents of this folder should not be edited.

IV. Files Included in Query Request

The following files will be included in the Zip file distributed with the query request.

1. Workplan Data Characterization v3.00.doc
2. run_queries.sas
3. data_characterization.sas
4. 2015-06-18 PCORnet CDM V3 0 ETL Annotated Data Dictionary Template

V. Output Files

Local files (*dmlocal* folder)

File name	File description
77 SAS datasets	The 77 tables created by the program (see Query List)
set.log	This contains the output results of the PROC SETINIT procedure. This information is used to populate parts of query table XTBL_L3_METADATA.

Files to be returned to the DRN OC (*drnoc* folder)

File name	File description
[DATAMARTID]_[DATE]_ data_characterization.cpt	This SAS transport file (similar to a Zip file) contains the 77 SAS tables produced by the query package.
[DATAMARTID]_[DATE]_ data_characterization.pdf	This PDF contains a partial print out of the 77 tables for the benefit of non-programmers. For ease of readibility, it excludes the first three columns of the table (DataMartID, Response Date, and Query Package), and large tables are limited to the 100 most frequent observations. Empty tables (this is unexpected) are not printed.
[DATAMARTID]_[DATE]_ data_characterization.log	A SAS log file. The log file is printed when the SAS program completes in order to exclude the DataMart's directory path information.

1. PDF file limited to the 100 most frequent observations with counts above the low cell count threshold and sorted by descending record count.

VI. Table Shells: DEMOGRAPHIC Queries

The table shells in this document are for illustrative purposes only, and individual DataMart output may differ. Column naming conventions are as follows: TAG is the field name; ALL_N and RECORD_N are counts of all records; RECORD_PCT is the percent of all records; DISTINCT_N is a count of records with unique values for the specified field; and NULL_N is the count of records with null or missing values for the specified field. RECORD_PCT will be blank for rows with values of 0 or BT (below threshold).

dem_l3_n (*Note*: **ALL_N** should equal **DISTINCT_N** and **NULL_N** should be 0)

DATAMARTID	RESPONSE_DATE	QUERY_PACKAGE	DATASET	TAG	ALL_N	DISTINCT_N	NULL_N
D1TEST	28OCT2015	DC V3.00	DEMOGRAPHIC	PATID			

dem_l3_ageyrsdist1

DATAMARTID	RESPONSE_DATE	QUERY_PACKAGE	STAT	RECORD_N
D1TEST	28OCT2015	DC V3.00	MIN	
D1TEST	28OCT2015	DC V3.00	MEAN	
D1TEST	28OCT2015	DC V3.00	MEDIAN	
D1TEST	28OCT2015	DC V3.00	MAX	
D1TEST	28OCT2015	DC V3.00	N	
D1TEST	28OCT2015	DC V3.00	NULL or missing	

dem_l3_ageyrsdist2

DATAMARTID	RESPONSE_DATE	QUERY_PACKAGE	AGE_GROUP	RECORD_N	RECORD_PCT
D1TEST	28OCT2015	DC V3.00	<0 yrs		
D1TEST	28OCT2015	DC V3.00	0-1 yrs		
D1TEST	28OCT2015	DC V3.00	2-4 yrs		
D1TEST	28OCT2015	DC V3.00	5-9 yrs		
D1TEST	28OCT2015	DC V3.00	10-14 yrs		
D1TEST	28OCT2015	DC V3.00	15-18 yrs		
D1TEST	28OCT2015	DC V3.00	19-21 yrs		
D1TEST	28OCT2015	DC V3.00	22-44 yrs		
D1TEST	28OCT2015	DC V3.00	45-64 yrs		
D1TEST	28OCT2015	DC V3.00	65-74 yrs		
D1TEST	28OCT2015	DC V3.00	75-110 yrs		
D1TEST	28OCT2015	DC V3.00	>110 yrs		
D1TEST	28OCT2015	DC V3.00	NULL or missing		

dem_l3_hispdist

DATAMARTID	RESPONSE_DATE	QUERY_PACKAGE	HISPANIC	RECORD_N	RECORD_PCT
D1TEST	28OCT2015	DC V3.00	N		
D1TEST	28OCT2015	DC V3.00	R		
D1TEST	28OCT2015	DC V3.00	Y		
D1TEST	28OCT2015	DC V3.00	NI		
D1TEST	28OCT2015	DC V3.00	UN		
D1TEST	28OCT2015	DC V3.00	OT		
D1TEST	28OCT2015	DC V3.00	NULL or missing		
D1TEST	28OCT2015	DC V3.00	Values outside of CDM specifications		

1. PDF file limited to the 100 most frequent observations with counts above the low cell count threshold and sorted by descending record count.

dem_l3_racedist

DATAMARTID	RESPONSE_DATE	QUERY_PACKAGE	RACE	RECORD_N	RECORD_PCT
D1TEST	28OCT2015	DC V3.00	01		
D1TEST	28OCT2015	DC V3.00	02		
D1TEST	28OCT2015	DC V3.00	03		
D1TEST	28OCT2015	DC V3.00	04		
D1TEST	28OCT2015	DC V3.00	05		
D1TEST	28OCT2015	DC V3.00	06		
D1TEST	28OCT2015	DC V3.00	07		
D1TEST	28OCT2015	DC V3.00	NI		
D1TEST	28OCT2015	DC V3.00	UN		
D1TEST	28OCT2015	DC V3.00	OT		
D1TEST	28OCT2015	DC V3.00	NULL or missing		
D1TEST	28OCT2015	DC V3.00	Values outside of CDM specifications		

dem_l3_sexdist

DATAMARTID	RESPONSE_DATE	QUERY_PACKAGE	SEX	RECORD_N	RECORD_PCT
D1TEST	28OCT2015	DC V3.00	A		
D1TEST	28OCT2015	DC V3.00	F		
D1TEST	28OCT2015	DC V3.00	M		
D1TEST	28OCT2015	DC V3.00	NI		
D1TEST	28OCT2015	DC V3.00	UN		
D1TEST	28OCT2015	DC V3.00	OT		
D1TEST	28OCT2015	DC V3.00	NULL or missing		
D1TEST	28OCT2015	DC V3.00	Values outside of CDM specifications		

1. PDF file limited to the 100 most frequent observations with counts above the low cell count threshold and sorted by descending record count.

VII. Table Shells: ENCOUNTER Queries

enc_l3_n (*Note*: For ENCOUNTERID and PATID, NULL_N should be 0. For EncounterID, ALL_N should equal DISTINCT_N)

DATAMARTID	RESPONSE_DATE	QUERY_ PACKAGE	DATASET	TAG	ALL_N	DISTINCT_N	NULL_N
D1TEST	28OCT2015	DC V3.00	ENCOUNTER	ENCOUNTERID			
D1TEST	28OCT2015	DC V3.00	ENCOUNTER	PATID			
D1TEST	28OCT2015	DC V3.00	ENCOUNTER	PROVIDERID			

enc_l3_n_visit (This query counts visits, which are a concatenation of PATID + PROVIDER_ID + ENC_TYPE + ADMIT_DT. ALL_N counts all records with non-missing values for all 4 fields).

DATAMARTID	RESPONSE_DATE	QUERY_PACKAGE	DATASET	TAG	ALL_N	DISTINCT_N
D1TEST	28OCT2015	DC V3.00	ENCOUNTER	visit		

enc_l3_admsrc

DATAMARTID	RESPONSE_DATE	QUERY_PACKAGE	ADMITTING_SOURCE	RECORD_ N	RECORD_ PCT
D1TEST	28OCT2015	DC V3.00	AF		
D1TEST	28OCT2015	DC V3.00	AL		
D1TEST	28OCT2015	DC V3.00	AV		
D1TEST	28OCT2015	DC V3.00	ED		
D1TEST	28OCT2015	DC V3.00	HH		
D1TEST	28OCT2015	DC V3.00	HO		
D1TEST	28OCT2015	DC V3.00	HS		
D1TEST	28OCT2015	DC V3.00	IP		
D1TEST	28OCT2015	DC V3.00	NH		
D1TEST	28OCT2015	DC V3.00	RH		
D1TEST	28OCT2015	DC V3.00	RS		
D1TEST	28OCT2015	DC V3.00	SN		
D1TEST	28OCT2015	DC V3.00	NI		
D1TEST	28OCT2015	DC V3.00	UN		
D1TEST	28OCT2015	DC V3.00	OT		
D1TEST	28OCT2015	DC V3.00	NULL or missing		
D1TEST	28OCT2015	DC V3.00	Values outside of CDM specifications		

1. PDF file limited to the 100 most frequent observations with counts above the low cell count threshold and sorted by descending record count.

enc_l3_enctype_admsrc[1]

DATAMARTID	RESPONSE_DATE	QUERY_PACKAGE	ENCTYPE	ADMITTING_SOURCE	RECORD_N	RECORD_PCT
D1TEST	28OCT2015	DC V3.00	AV	AF		
D1TEST	28OCT2015	DC V3.00	AV	AL		
D1TEST	28OCT2015	DC V3.00	AV	AV		
D1TEST	28OCT2015	DC V3.00	AV	ED		
D1TEST	28OCT2015	DC V3.00	AV	HH		
D1TEST	28OCT2015	DC V3.00	AV	HO		
D1TEST	28OCT2015	DC V3.00	AV	HS		
D1TEST	28OCT2015	DC V3.00	AV	IP		
D1TEST	28OCT2015	DC V3.00	AV	NH		
D1TEST	28OCT2015	DC V3.00	AV	RH		
D1TEST	28OCT2015	DC V3.00	AV	RS		
D1TEST	28OCT2015	DC V3.00	AV	SN		
D1TEST	28OCT2015	DC V3.00	AV	NI		
D1TEST	28OCT2015	DC V3.00	AV	UN		
D1TEST	28OCT2015	DC V3.00	AV	OT		
D1TEST	28OCT2015	DC V3.00	AV	NULL or missing		
D1TEST	28OCT2015	DC V3.00	AV	Values outside of CDM specifications		

enc_l3_adate_y (*Note*: ADMIT_DATE should not be NULL or missing)

DATAMARTID	RESPONSE_DATE	QUERY_PACKAGE	ADMIT_DATE	RECORD_N	RECORD_PCT	DISTINCT_PATID_N
D1TEST	28OCT2015	DC V3.00	2015			
D1TEST	28OCT2015	DC V3.00	2016			

enc_l3_adate_ym [1]

DATAMARTID	RESPONSE_DATE	QUERY_PACKAGE	ADMIT_DATE	RECORD_N
D1TEST	28OCT2015	DC V3.00	2015_07	
D1TEST	28OCT2015	DC V3.00	2015_08	
D1TEST	28OCT2015	DC V3.00	NULL or missing	

enc_l3_enctype_adate_ym [1]

DATAMARTID	RESPONSE_DATE	QUERY_PACKAGE	ENCTYPE	ADMIT_DATE	RECORD_N
D1TEST	28OCT2015	DC V3.00	AV	2015_07	
D1TEST	28OCT2015	DC V3.00	AV	2015_08	
D1TEST	28OCT2015	DC V3.00	AV	NULL or missing	

enc_l3_ddate_y

DATAMARTID	RESPONSE_DATE	QUERY_PACKAGE	DISCHARGE_DATE	RECORD_N	RECORD_PCT
D1TEST	28OCT2015	DC V3.00	2015		
D1TEST	28OCT2015	DC V3.00	NULL or missing		

1. PDF file limited to the 100 most frequent observations with counts above the low cell count threshold and sorted by descending record count.

enc_l3_ddate_ym [1]

DATAMARTID	RESPONSE_DATE	QUERY_PACKAGE	DISCHARGE_DATE	RECORD_N
D1TEST	28OCT2015	DC V3.00	2015_07	
D1TEST	28OCT2015	DC V3.00	2015_08	
D1TEST	28OCT2015	DC V3.00	NULL or missing	

enc_l3_enctype_ddate_ym [1]

DATAMARTID	RESPONSE_DATE	QUERY_PACKAGE	ENCTYPE	DISCHARGE_DATE	RECORD_N
D1TEST	28OCT2015	DC V3.00	AV	2004_01	
D1TEST	28OCT2015	DC V3.00	AV	2004_02	
D1TEST	28OCT2015	DC V3.00	AV	NULL or missing	

enc_l3_disdisp

DATAMARTID	RESPONSE_DATE	QUERY_PACKAGE	DISCHARGE_DISPOSITION	RECORD_N	RECORD_PCT
D1TEST	28OCT2015	DC V3.00	A		
D1TEST	28OCT2015	DC V3.00	E		
D1TEST	28OCT2015	DC V3.00	NI		
D1TEST	28OCT2015	DC V3.00	UN		
D1TEST	28OCT2015	DC V3.00	OT		
D1TEST	28OCT2015	DC V3.00	NULL or missing		
D1TEST	28OCT2015	DC V3.00	Values outside of CDM specifications		

enc_l3_enctype_disdisp

DATAMARTID	RESPONSE_ DATE	QUERY_ PACKAGE	ENCTYPE	DISCHARGE_ DISPOSITION	RECORD_N	RECORD _PCT
D1TEST	28OCT2015	DC V3.00	AV	A		
D1TEST	28OCT2015	DC V3.00	AV	E		
D1TEST	28OCT2015	DC V3.00	AV	NI		
D1TEST	28OCT2015	DC V3.00	AV	UN		
D1TEST	28OCT2015	DC V3.00	AV	OT		
D1TEST	28OCT2015	DC V3.00	AV	NULL or missing		
D1TEST	28OCT2015	DC V3.00	AV	Values outside of CDM specifications		

1. PDF file limited to the 100 most frequent observations with counts above the low cell count threshold and sorted by descending record count.

enc_l3_disstat

DATAMARTID	RESPONSE_DATE	QUERY_PACKAGE	DISCHARGE_STATUS	RECORD_N	RECORD_PCT
D1TEST	28OCT2015	DC V3.00	AF		
D1TEST	28OCT2015	DC V3.00	AL		
D1TEST	28OCT2015	DC V3.00	AM		
D1TEST	28OCT2015	DC V3.00	AW		
D1TEST	28OCT2015	DC V3.00	EX		
D1TEST	28OCT2015	DC V3.00	HH		
D1TEST	28OCT2015	DC V3.00	HO		
D1TEST	28OCT2015	DC V3.00	HS		
D1TEST	28OCT2015	DC V3.00	IP		
D1TEST	28OCT2015	DC V3.00	NH		
D1TEST	28OCT2015	DC V3.00	RH		
D1TEST	28OCT2015	DC V3.00	RS		
D1TEST	28OCT2015	DC V3.00	SH		
D1TEST	28OCT2015	DC V3.00	SN		
D1TEST	28OCT2015	DC V3.00	NI		
D1TEST	28OCT2015	DC V3.00	UN		
D1TEST	28OCT2015	DC V3.00	OT		
D1TEST	28OCT2015	DC V3.00	NULL or missing		
D1TEST	28OCT2015	DC V3.00	Values outside of CDM specifications		

enc_l3_enctype_disstat [1]

DATAMARTID	RESPONSE_DATE	QUERY_PACKAGE	ENCTYPE	DISCHARGE_STATUS	RECORD_PCT
D1TEST	28OCT2015	DC V3.00	AV	AF	
D1TEST	28OCT2015	DC V3.00	AV	AL	
D1TEST	28OCT2015	DC V3.00	AV	AM	
D1TEST	28OCT2015	DC V3.00	AV	AW	
D1TEST	28OCT2015	DC V3.00	AV	EX	
D1TEST	28OCT2015	DC V3.00	AV	HH	
D1TEST	28OCT2015	DC V3.00	AV	HO	
D1TEST	28OCT2015	DC V3.00	AV	HS	
D1TEST	28OCT2015	DC V3.00	AV	IP	
D1TEST	28OCT2015	DC V3.00	AV	NH	
D1TEST	28OCT2015	DC V3.00	AV	RH	
D1TEST	28OCT2015	DC V3.00	AV	RS	
D1TEST	28OCT2015	DC V3.00	AV	SH	
D1TEST	28OCT2015	DC V3.00	AV	SN	
D1TEST	28OCT2015	DC V3.00	AV	NI	
D1TEST	28OCT2015	DC V3.00	AV	UN	
D1TEST	28OCT2015	DC V3.00	AV	OT	
D1TEST	28OCT2015	DC V3.00	AV	NULL or missing	

1. PDF file limited to the 100 most frequent observations with counts above the low cell count threshold and sorted by descending record count.

enc_l3_drg [1] **(***Note:* For DRG, values outside of CDM specifications signifies an incorrect data type (ie, DRG is not CHAR 3)

DATAMARTID	RESPONSE_DATE	QUERY_PACKAGE	DRG	RECORD_N	RECORD_PCT
D1TEST	28OCT2015	DC V3.00	001		
D1TEST	28OCT2015	DC V3.00	150		
D1TEST	28OCT2015	DC V3.00	NULL or missing		
D1TEST	28OCT2015	DC V3.00	Values outside of CDM specifications		

enc_l3_drg_type

DATAMARTID	RESPONSE_DATE	QUERY_PACKAGE	DRG_TYPE	RECORD_N	RECORD_PCT
D1TEST	28OCT2015	DC V3.00	01		
D1TEST	28OCT2015	DC V3.00	02		
D1TEST	28OCT2015	DC V3.00	NI		
D1TEST	28OCT2015	DC V3.00	UN		
D1TEST	28OCT2015	DC V3.00	OT		
D1TEST	28OCT2015	DC V3.00	NULL or missing		
D1TEST	28OCT2015	DC V3.00	Values outside of CDM specifications		

enc_l3_enctype_drg [1]

DATAMARTID	RESPONSE_DATE	QUERY_PACKAGE	ENCTYPE	DRG	RECORD_N	RECORD_PCT
D1TEST	28OCT2015	DC V3.00	AV	001		
D1TEST	28OCT2015	DC V3.00	AV	150		
D1TEST	28OCT2015	DC V3.00	AV	NULL or missing		
D1TEST	28OCT2015	DC V3.00	AV	Values outside of CDM specifications		

enc_l3_enctype (*Note:* ENCTYPE should not be NULL or Missing. Visits are a concatenation of PATID + PROVIDER_ID + ENC_TYPE + ADMIT_DT)

DATAMARTID	RESPONSE_DATE	QUERY_PACKAGE	ENCTYPE	RECORD_N	RECORD_PCT	DISTINCT_VISIT_N	DISTINCT_PATID_N
D1TEST	28OCT2015	DC V3.00	AV				
D1TEST	28OCT2015	DC V3.00	ED				
D1TEST	28OCT2015	DC V3.00	EI				
D1TEST	28OCT2015	DC V3.00	IP				
D1TEST	28OCT2015	DC V3.00	IS				
D1TEST	28OCT2015	DC V3.00	OA				
D1TEST	28OCT2015	DC V3.00	NI				
D1TEST	28OCT2015	DC V3.00	UN				
D1TEST	28OCT2015	DC V3.00	OT				
D1TEST	28OCT2015	DC V3.00	NULL or missing				
D1TEST	28OCT2015	DC V3.00	Values outside of CDM specifications				

1. PDF file limited to the 100 most frequent observations with counts above the low cell count threshold and sorted by descending record count.

enc_l3_dash1 (This query counts the number of patients with any encounter during the designated period of time prior to the maximum ADMIT_DATE. If the maximum ADMIT_DATE is in the future the current date is used instead).

DATAMARTID	RESPONSE_DATE	QUERY_PACKAGE	Period	DISTINCT_PATID_N
D1TEST	28OCT2015	DC V3.00	1 yr	
D1TEST	28OCT2015	DC V3.00	2 yrs	
D1TEST	28OCT2015	DC V3.00	3 yrs	
D1TEST	28OCT2015	DC V3.00	4 yrs	
D1TEST	28OCT2015	DC V3.00	5 yrs	
D1TEST	28OCT2015	DC V3.00	All yrs	

enc_l3_dash2 (This query counts the number of patients with any IP (Inpatient), EI (ED to Inpatient), ED (Emergency Department) or AV (Ambulatory Visit) encounter during the designated period of time prior to the maximum ADMIT_DATE. If the maximum ADMIT_DATE is in the future the current date is used instead).

DATAMARTID	RESPONSE_DATE	QUERY_PACKAGE	PERIOD	DISTINCT_PATID_N
D1TEST	28OCT2015	DC V3.00	1 yr	
D1TEST	28OCT2015	DC V3.00	2 yrs	
D1TEST	28OCT2015	DC V3.00	3 yrs	
D1TEST	28OCT2015	DC V3.00	4 yrs	
D1TEST	28OCT2015	DC V3.00	5 yrs	
D1TEST	28OCT2015	DC V3.00	All yrs	

1. PDF file limited to the 100 most frequent observations with counts above the low cell count threshold and sorted by descending record count.

VIII. Table Shells: DIAGNOSIS Queries

dia_l3_n (*Note*: For all fields, NULL_N should be 0. For DIAGNOSISID, ALL_N should equal DISTINCT_N)

DATAMARTID	RESPONSE_ DATE	QUERY_PACKAGE	DATASET	TAG	ALL_N	DISTINCT_N	NULL_N
D1TEST	28OCT2015	DC V3.00	DIAGNOSIS	ENCOUNTERID			
D1TEST	28OCT2015	DC V3.00	DIAGNOSIS	PATID			
D1TEST	28OCT2015	DC V3.00	DIAGNOSIS	DIAGNOSISID			

dia_l3_dx [1] (*Note*: DX should not be NULL or missing)

DATAMARTID	RESPONSE_DATE	QUERY_PACKAGE	DX	RECORD_N	RECORD_PCT
D1TEST	28OCT2015	DC V3.00	840		
D1TEST	28OCT2015	DC V3.00	953.5		

dia_l3_dx_dxtype [1] (*Note*: DX_TYPE should not be NULL or missing)

DATAMART ID	RESPONSE_ DATE	QUERY_ PACKAGE	DX	DX_TYPE	RECORD_N	DISTINCT_ PATID_N
D1TEST	28OCT2015	DC V3.00	008.45	09		
D1TEST	28OCT2015	DC V3.00	NULL or missing	10		
D1TEST	28OCT2015	DC V3.00	NULL or missing	11		
D1TEST	28OCT2015	DC V3.00	NULL or missing	SM		
D1TEST	28OCT2015	DC V3.00	NULL or missing	NI		
D1TEST	28OCT2015	DC V3.00	NULL or missing	UN		
D1TEST	28OCT2015	DC V3.00	NULL or missing	OT		
D1TEST	28OCT2015	DC V3.00	NULL or missing	NULL or missing		
D1TEST	28OCT2015	DC V3.00	NULL or missing	Values outside of CDM specifications		

1. PDF file limited to the 100 most frequent observations with counts above the low cell count threshold and sorted by descending record count.

dia_l3_dxsource (*Note*: DX_SOURCE should not be NULL or missing)

DATAMARTID	RESPONSE_DATE	QUERY_PACKAGE	DX_SOURCE	RECORD_N	RECORD_PCT
D1TEST	28OCT2015	DC V3.00	AD		
D1TEST	28OCT2015	DC V3.00	DI		
D1TEST	28OCT2015	DC V3.00	FI		
D1TEST	28OCT2015	DC V3.00	IN		
D1TEST	28OCT2015	DC V3.00	NI		
D1TEST	28OCT2015	DC V3.00	UN		
D1TEST	28OCT2015	DC V3.00	OT		
D1TEST	28OCT2015	DC V3.00	NULL or missing		
D1TEST	28OCT2015	DC V3.00	Values outside of CDM specifications		

dia_l3_dxtype_dxsource

DATAMARTID	RESPONSE_DATE	QUERY_PACKAGE	DX_TYPE	DX_SOURCE	RECORD_N
D1TEST	28OCT2015	DC V3.00	09	AD	
D1TEST	28OCT2015	DC V3.00	09	DI	
D1TEST	28OCT2015	DC V3.00	09	FI	
D1TEST	28OCT2015	DC V3.00	09	IN	
D1TEST	28OCT2015	DC V3.00	09	NI	
D1TEST	28OCT2015	DC V3.00	09	UN	
D1TEST	28OCT2015	DC V3.00	09	OT	
D1TEST	28OCT2015	DC V3.00	09	NULL or missing	
D1TEST	28OCT2015	DC V3.00	09	Values outside of CDM specifications	

dia_l3_PDX

DATAMARTID	RESPONSE_DATE	QUERY_PACKAGE	PDX	RECORD_N	RECORD_PCT
D1TEST	28OCT2015	DC V3.00	P		
D1TEST	28OCT2015	DC V3.00	S		
D1TEST	28OCT2015	DC V3.00	X		
D1TEST	28OCT2015	DC V3.00	NI		
D1TEST	28OCT2015	DC V3.00	UN		
D1TEST	28OCT2015	DC V3.00	OT		
D1TEST	28OCT2015	DC V3.00	NULL or missing		
D1TEST	28OCT2015	DC V3.00	Values outside of CDM specifications		

1. PDF file limited to the 100 most frequent observations with counts above the low cell count threshold and sorted by descending record count.

dia_l3_PDX_enctype

DATAMARTID	RESPONSE_DATE	QUERY_PACKAGE	PDX	ENCTYPE	RECORD_N	DISTINCT_ENCID_N	DISTINCT_PATID_N
D1TEST	28OCT2015	DC V3.00	P	AV			
D1TEST	28OCT2015	DC V3.00	P	ED			
D1TEST	28OCT2015	DC V3.00	P	EI			
D1TEST	28OCT2015	DC V3.00	P	IP			
D1TEST	28OCT2015	DC V3.00	P	IS			
D1TEST	28OCT2015	DC V3.00	P	OA			
D1TEST	28OCT2015	DC V3.00	P	NI			
D1TEST	28OCT2015	DC V3.00	P	UN			
D1TEST	28OCT2015	DC V3.00	P	OT			
D1TEST	28OCT2015	DC V3.00	P	NULL or missing			
D1TEST	28OCT2015	DC V3.00	P	Values outside CDM specifications			

dia_l3_adate_y

DATAMARTID	RESPONSE_DATE	QUERY_PACKAGE	ADMIT_DATE	RECORD_N	RECORD_PCT	DISTINCT_ENCID_N	DISTINCT_PATID_N
D1TEST	28OCT2015	DC V3.00	2015				
D1TEST	28OCT2015	DC V3.00	NULL or missing				

dia_l3_adate_ym [1]

DATAMARTID	RESPONSE_DATE	QUERY_PACKAGE	ADMIT_DATE	RECORD_N
D1TEST	28OCT2015	DC V3.00	2015_07	
D1TEST	28OCT2015	DC V3.00	2015_08	
D1TEST	28OCT2015	DC V3.00	NULL or missing	

dia_l3_enctype

DATAMARTID	RESPONSE_DATE	QUERY_PACKAGE	ENCTYPE	RECORD_N	RECORD_PCT
D1TEST	28OCT2015	DC V3.00	AV		
D1TEST	28OCT2015	DC V3.00	ED		
D1TEST	28OCT2015	DC V3.00	EI		
D1TEST	28OCT2015	DC V3.00	IP		
D1TEST	28OCT2015	DC V3.00	IS		
D1TEST	28OCT2015	DC V3.00	OA		
D1TEST	28OCT2015	DC V3.00	NI		
D1TEST	28OCT2015	DC V3.00	UN		
D1TEST	28OCT2015	DC V3.00	OT		
D1TEST	28OCT2015	DC V3.00	NULL or missing		
D1TEST	28OCT2015	DC V3.00	Values outside of CDM specifications		

1. PDF file limited to the 100 most frequent observations with counts above the low cell count threshold and sorted by descending record count.

dia_l3_dxtype_enctype

DATAMARTID	RESPONSE_DATE	QUERY_PACKAGE	DX_TYPE	ENCTYPE	RECORD_N
D1TEST	28OCT2015	DC V3.00	09	AV	
D1TEST	28OCT2015	DC V3.00	09	ED	
D1TEST	28OCT2015	DC V3.00	09	EI	
D1TEST	28OCT2015	DC V3.00	09	IP	
D1TEST	28OCT2015	DC V3.00	09	IS	
D1TEST	28OCT2015	DC V3.00	09	OA	
D1TEST	28OCT2015	DC V3.00	09	NI	
D1TEST	28OCT2015	DC V3.00	09	UN	
D1TEST	28OCT2015	DC V3.00	09	OT	
D1TEST	28OCT2015	DC V3.00	09	NULL or missing	
D1TEST	28OCT2015	DC V3.00	09	Values outside of CDM specifications	

dia_l3_enctype_adate_ym [1]

DATAMARTID	RESPONSE_DATE	QUERY_PACKAGE	ENCTYPE	ADMIT_DATE	DISTINCT_ENCID_N	RECORD_N
D1TEST	28OCT2015	DC V3.00	AV	2015_07		
D1TEST	28OCT2015	DC V3.00	ED	2015_07		
D1TEST	28OCT2015	DC V3.00	EI	2015_07		
D1TEST	28OCT2015	DC V3.00	IP	2015_07		
D1TEST	28OCT2015	DC V3.00	IS	2015_07		
D1TEST	28OCT2015	DC V3.00	OA	2015_07		
D1TEST	28OCT2015	DC V3.00	NI	2015_07		
D1TEST	28OCT2015	DC V3.00	UN	2015_07		
D1TEST	28OCT2015	DC V3.00	OT	2015_07		
D1TEST	28OCT2015	DC V3.00	NULL or missing	2015_07		
D1TEST	28OCT2015	DC V3.00	Values outside of CDM specifications	2015_07		

dia_l3_dash1 (This query counts the number of patients with any diagnosis record with a populated DX code during the designated period of time prior to the maximum ADMIT_DATE. If the maximum ADMIT_DATE is in the future the current date is substituted).

DATAMARTID	RESPONSE_DATE	QUERY_PACKAGE	PERIOD	DISTINCT_PATID_N
D1TEST	28OCT2015	DC V3.00	1 yr	
D1TEST	28OCT2015	DC V3.00	2 yrs	
D1TEST	28OCT2015	DC V3.00	3 yrs	
D1TEST	28OCT2015	DC V3.00	4 yrs	
D1TEST	28OCT2015	DC V3.00	5 yrs	
D1TEST	28OCT2015	DC V3.00	All yrs	

1. PDF file limited to the 100 most frequent observations with counts above the low cell count threshold and sorted by descending record count.

IX. Table Shells: PROCEDURES Queries

pro_l3_n (*Note:* For all fields, NULL_N should be 0. For PROCEDURESID, ALL_N should be equal to DISTINCT_N).

DATAMARTID	RESPONSE_ DATE	QUERY_ PACKAGE	DATASET	TAG	ALL_ N	DISTINCT_ N	NULL_ N
D1TEST	28OCT2015	DC V3.00	PROCEDURES	ENCOUNTERID			
D1TEST	28OCT2015	DC V3.00	PROCEDURES	PATID			
D1TEST	28OCT2015	DC V3.00	PROCEDURES	PROCEDURESID			

pro_l3_px [1] (*Note*: PX should not be NULL or missing)

DATAMARTID	RESPONSE_DATE	QUERY_PACKAGE	PX	RECORD_N	RECORD_PCT
D1TEST	28OCT2015	DC V3.00	00.11		
D1TEST	28OCT2015	DC V3.00	0067T		
D1TEST	28OCT2015	DC V3.00	NULL or missing		

pro_l3_adate_y

DATAMARTID	RESPONSE_ DATE	QUERY_ PACKAGE	ADMIT_ DATE	RECORD _N	RECORD _PCT	DISTINCT_ ENCID_N	DISTINCT_PATID _N
D1TEST	28OCT2015	DC V3.00	2015				
D1TEST	28OCT2015	DC V3.00	NULL or missing				

pro_l3_adate_ym [1]

DATAMARTID	RESPONSE_DATE	QUERY_PACKAGE	ADMIT_DATE	RECORD_N
D1TEST	28OCT2015	DC V3.00	2015_07	
D1TEST	28OCT2015	DC V3.00	2015_08	
D1TEST	28OCT2015	DC V3.00	NULL or missing	

pro_l3_pxdate_y

DATAMARTID	RESPONSE _DATE	QUERY_P ACKAGE	PX_ DATE	RECORD _N	RECORD_ PCT	DISTINCT_ENCID_ N	DISTINCT_PATID_ N
D1TEST	28OCT2015	DC V3.00	2015				
D1TEST	28OCT2015	DC V3.00	NULL or missing				

1. PDF file limited to the 100 most frequent observations with counts above the low cell count threshold and sorted by descending record count.

pro_l3_px_enctype [1]

DATAMARTID	RESPONSE_DATE	QUERY_PACKAGE	PX	ENCTYPE	RECORD_N
D1TEST	28OCT2015	DC V3.00	00.11	AV	
D1TEST	28OCT2015	DC V3.00	00.11	ED	
D1TEST	28OCT2015	DC V3.00	00.11	EI	
D1TEST	28OCT2015	DC V3.00	00.11	IP	
D1TEST	28OCT2015	DC V3.00	00.11	IS	
D1TEST	28OCT2015	DC V3.00	00.11	OA	
D1TEST	28OCT2015	DC V3.00	00.11	NI	
D1TEST	28OCT2015	DC V3.00	00.11	UN	
D1TEST	28OCT2015	DC V3.00	00.11	OT	
D1TEST	28OCT2015	DC V3.00	NULL or missing	NULL or missing	
D1TEST	28OCT2015	DC V3.00	NULL or missing	Values outside of CDM specifications	

pro_l3_enctype

DATAMARTID	RESPONSE_DATE	QUERY_PACKAGE	ENCTYPE	RECORD_N	RECORD_PCT
D1TEST	28OCT2015	DC V3.00	AV		
D1TEST	28OCT2015	DC V3.00	ED		
D1TEST	28OCT2015	DC V3.00	EI		
D1TEST	28OCT2015	DC V3.00	IP		
D1TEST	28OCT2015	DC V3.00	IS		
D1TEST	28OCT2015	DC V3.00	OA		
D1TEST	28OCT2015	DC V3.00	NI		
D1TEST	28OCT2015	DC V3.00	UN		
D1TEST	28OCT2015	DC V3.00	OT		
D1TEST	28OCT2015	DC V3.00	NULL or missing		
D1TEST	28OCT2015	DC V3.00	Values outside of CDM specifications		

pro_l3_pxtype_enctype [1] (*Note*: PX_TYPE should not be NULL or missing)

DATAMARTID	RESPONSE_DATE	QUERY_PACKAGE	PX_TYPE	ENCTYPE	RECORD_N
D1TEST	28OCT2015	DC V3.00	09	AV	
D1TEST	28OCT2015	DC V3.00	09	ED	
D1TEST	28OCT2015	DC V3.00	09	EI	
D1TEST	28OCT2015	DC V3.00	09	IP	
D1TEST	28OCT2015	DC V3.00	09	IS	
D1TEST	28OCT2015	DC V3.00	09	OA	
D1TEST	28OCT2015	DC V3.00	09	NI	
D1TEST	28OCT2015	DC V3.00	09	UN	
D1TEST	28OCT2015	DC V3.00	09	OT	
D1TEST	28OCT2015	DC V3.00	09	NULL or missing	
D1TEST	28OCT2015	DC V3.00	09	Values outside of CDM specifications	

1. PDF file limited to the 100 most frequent observations with counts above the low cell count threshold and sorted by descending record count.

pro_l3_enctype_adate_ym [1]

DATAMARTID	RESPONSE_ DATE	QUERY_PACKAGE	ENCTYPE	ADMIT_DATE	DIST_ENCID_N	RECORD_N
D1TEST	28OCT2015	DC V3.00	AV	2015_07		
D1TEST	28OCT2015	DC V3.00	ED	2015_07		
D1TEST	28OCT2015	DC V3.00	EI	2015_07		
D1TEST	28OCT2015	DC V3.00	IP	2015_07		
D1TEST	28OCT2015	DC V3.00	IS	2015_07		
D1TEST	28OCT2015	DC V3.00	OA	2015_07		
D1TEST	28OCT2015	DC V3.00	NI	2015_07		
D1TEST	28OCT2015	DC V3.00	UN	2015_07		
D1TEST	28OCT2015	DC V3.00	OT	2015_07		
D1TEST	28OCT2015	DC V3.00	NULL or missing	2015_07		

pro_l3_px_pxtype [1]

DATAMARTID	RESPONSE_DATE	QUERY_PACKAGE	PX	PX_TYPE	RECORD_N	DISTINCT_ PATID_N
D1TEST	28OCT2015	DC V3.00	00.11	09		

pro_l3_pxsource

DATAMARTID	RESPONSE_DATE	QUERY_PACKAGE	PX_SOURCE	RECORD_N	RECORD_PCT
D1TEST	28OCT2015	DC V3.00	BI		
D1TEST	28OCT2015	DC V3.00	CL		
D1TEST	28OCT2015	DC V3.00	OD		
D1TEST	28OCT2015	DC V3.00	NI		
D1TEST	28OCT2015	DC V3.00	UN		
D1TEST	28OCT2015	DC V3.00	OT		
D1TEST	28OCT2015	DC V3.00	NULL or missing		
D1TEST	28OCT2015	DC V3.00	Values outside of CDM specifications		

1. PDF file limited to the 100 most frequent observations with counts above the low cell count threshold and sorted by descending record count.

X. Table Shells: ENROLLMENT Queries

enr_l3_n (ENROLLID is a concatenation of PATID + ENR_START_DATE + ENR_BASIS. For ENROLLID, ALL_N should equal DISTINCT_N. For all fields, NULL_N should be 0)

DATAMARTID	RESPONSE_ DATE	QUERY_ PACKAGE	DATASET	TAG	ALL_ N	DISTINCT_ N	NULL _N
D1TEST	28OCT2015	DC V3.00	ENROLLMENT	PATID			
D1TEST	28OCT2015	DC V3.00	ENROLLMENT	ENR_START_DATE			
D1TEST	28OCT2015	DC V3.00	ENROLLMENT	ENROLLID			

enr_l3_dist_start (This query measures the distribution of ENROLL_START_DATE).

DATAMARTID	RESPONSE_ DATE	QUERY_ PACKAGE	STAT	RECORD_N
D1TEST	28OCT2015	DC V3.00	MIN	
D1TEST	28OCT2015	DC V3.00	P1	
D1TEST	28OCT2015	DC V3.00	P5	
D1TEST	28OCT2015	DC V3.00	P25	
D1TEST	28OCT2015	DC V3.00	MEDIAN	
D1TEST	28OCT2015	DC V3.00	P75	
D1TEST	28OCT2015	DC V3.00	P95	
D1TEST	28OCT2015	DC V3.00	P99	
D1TEST	28OCT2015	DC V3.00	MAX	
D1TEST	28OCT2015	DC V3.00	N	
D1TEST	28OCT2015	DC V3.00	NULL or missing	

enr_l3_dist_end (This query measures the distribution of ENROLL_END_DATE).

DATAMARTID	RESPONSE_ DATE	QUERY_ PACKAGE	STAT	RECORD_N
D1TEST	28OCT2015	DC V3.00	MIN	
D1TEST	28OCT2015	DC V3.00	P1	
D1TEST	28OCT2015	DC V3.00	P5	
D1TEST	28OCT2015	DC V3.00	P25	
D1TEST	28OCT2015	DC V3.00	MEDIAN	
D1TEST	28OCT2015	DC V3.00	P75	
D1TEST	28OCT2015	DC V3.00	P95	
D1TEST	28OCT2015	DC V3.00	P99	
D1TEST	28OCT2015	DC V3.00	MAX	
D1TEST	28OCT2015	DC V3.00	N	
D1TEST	28OCT2015	DC V3.00	NULL or missing	

enr_l3_dist_enrmonth [1] (This query counts the number of records in the ENROLLMENT table by number of enrollment months. Enrollment months are calculated as the difference between the ENR_END_DATE and ENR_START_DATE in months. Records with null or missing ENR_END_DATE or ENR_START_DATE are excluded from the calculation)

DATAMARTID	RESPONSE_ DATE	QUERY_ PACKAGE	ENROLL_M	RECORD_N	RECORD_PCT
D1TEST	28OCT2015	DC V3.00	<0		
D1TEST	28OCT2015	DC V3.00	0		
D1TEST	28OCT2015	DC V3.00	1		
D1TEST	28OCT2015	DC V3.00	2		
D1TEST	28OCT2015	DC V3.00	3		
D1TEST	28OCT2015	DC V3.00	4		
D1TEST	28OCT2015	DC V3.00	NULL or missing		

1. PDF file limited to the 100 most frequent observations with counts above the low cell count threshold and sorted by descending record count.

enr_l3_dist_enryear [1] (This query counts the number of records in the ENROLLMENT table by enrollment years. Enrollment years are calculated as the difference between the ENR_END_DATE and ENR_START_DATE in years. Records with null or missing ENR_END_DATE or ENR_START_DATE are excluded from the calculation)

DATAMARTID	RESPONSE_ DATE	QUERY_ PACKAGE	ENROLL_Y	RECORD_N	RECORD_PCT
D1TEST	28OCT2015	DC V3.00	1		
D1TEST	28OCT2015	DC V3.00	2		
D1TEST	28OCT2015	DC V3.00	NULL or missing		

enr_l3_enr_ym [1] **(**This query counts the number of records in the ENROLMENT table by ENR_START_DATE)

DATAMARTID	RESPONSE_ DATE	QUERY_ PACKAGE	MONTH	RECORD_N
D1TEST	28OCT2015	DC V3.00	2015_07	
D1TEST	28OCT2015	DC V3.00	2015_08	
D1TEST	28OCT2015	DC V3.00	NULL or missing	

enr_l3_basisdist

DATAMARTID	RESPONSE_ DATE	QUERY_ PACKAGE	ENR_BASIS	RECORD_N	RECORD_PCT
D1TEST	28OCT2015	DC V3.00	A		
D1TEST	28OCT2015	DC V3.00	E		
D1TEST	28OCT2015	DC V3.00	G		
D1TEST	28OCT2015	DC V3.00	I		
D1TEST	28OCT2015	DC V3.00	NULL or missing		
D1TEST	28OCT2015	DC V3.00	Values outside of CDM specifications		

enr_l3_per_patid (This query calculates the distribution of the number of enrollment records per PATID)

DATAMARTID	RESPONSE_ DATE	QUERY_ PACKAGE	STAT	RECORD_N
D1TEST	28OCT2015	DC V3.00	MIN	
D1TEST	28OCT2015	DC V3.00	P1	
D1TEST	28OCT2015	DC V3.00	P5	
D1TEST	28OCT2015	DC V3.00	P25	
D1TEST	28OCT2015	DC V3.00	MEDIAN	
D1TEST	28OCT2015	DC V3.00	P75	
D1TEST	28OCT2015	DC V3.00	P95	
D1TEST	28OCT2015	DC V3.00	P99	
D1TEST	28OCT2015	DC V3.00	MAX	
D1TEST	28OCT2015	DC V3.00	N	
D1TEST	28OCT2015	DC V3.00	NULL or missing	

1. PDF file limited to the 100 most frequent observations with counts above the low cell count threshold and sorted by descending record count.

XI. Table Shells: VITAL Queries

vit_l3_n (*Note:* For PATID and VITALID, NULL_N should be 0. For VITALID, ALL_N should be equal to DISTINCT_N).

DATAMARTID	RESPONSE_ DATE	QUERY_ PACKAGE	DATASET	TAG	ALL_N	DISTINCT_N	NULL_N
D1TEST	28OCT2015	DC V3.00	VITAL	ENCOUNTERID			
D1TEST	28OCT2015	DC V3.00	VITAL	PATID			
D1TEST	28OCT2015	DC V3.00	VITAL	VITALID			

vit_l3_mdate_y (*Note:* MEASURE_DATE should not be NULL or missing)

DATAMARTID	RESPONSE_ DATE	QUERY_ PACKAGE	MEASURE_DATE	RECORD_N	RECORD_PCT	DISTINCT_PATID_N
D1TEST	28OCT2015	DC V3.00	2015			
D1TEST	28OCT2015	DC V3.00	2016			

vit_l3_mdate_ym [1]

DATAMARTID	RESPONSE_ DATE	QUERY_ PACKAGE	MEASURE_DATE	RECORD_N
D1TEST	28OCT2015	DC V3.00	2015_07	
D1TEST	28OCT2015	DC V3.00	2015_08	
D1TEST	28OCT2015	DC V3.00	NULL or missing	

vit_l3_vital_source (*Note:* VITAL_SOURCE should not be NULL or missing)

DATAMARTID	RESPONSE_ DATE	QUERY_ PACKAGE	VITAL_SOURCE	RECORD_N	RECORD_PCT
D1TEST	28OCT2015	DC V3.00	HC		
D1TEST	28OCT2015	DC V3.00	HD		
D1TEST	28OCT2015	DC V3.00	PD		
D1TEST	28OCT2015	DC V3.00	PR		
D1TEST	28OCT2015	DC V3.00	NI		
D1TEST	28OCT2015	DC V3.00	UN		
D1TEST	28OCT2015	DC V3.00	OT		
D1TEST	28OCT2015	DC V3.00	NULL or missing		
D1TEST	28OCT2015	DC V3.00	Values outside of CDM specifications		

1. PDF file limited to the 100 most frequent observations with counts above the low cell count threshold and sorted by descending record count.
2. NULL or missing rates are expected to be high due to the VITAL table structure

vit_l3_ht [2]

DATAMARTID	RESPONSE_ DATE	QUERY_ PACKAGE	HT_GROUP	RECORD_N	RECORD_PCT	DISTINCT_PATID_N
D1TEST	28OCT2015	DC V3.00	<0			
D1TEST	28OCT2015	DC V3.00	0-10			
D1TEST	28OCT2015	DC V3.00	11-20			
D1TEST	28OCT2015	DC V3.00	21-45			
D1TEST	28OCT2015	DC V3.00	46-52			
D1TEST	28OCT2015	DC V3.00	53-58			
D1TEST	28OCT2015	DC V3.00	59-64			
D1TEST	28OCT2015	DC V3.00	65-70			
D1TEST	28OCT2015	DC V3.00	71-76			
D1TEST	28OCT2015	DC V3.00	77-82			
D1TEST	28OCT2015	DC V3.00	83-88			
D1TEST	28OCT2015	DC V3.00	89-94			
D1TEST	28OCT2015	DC V3.00	>=95			
D1TEST	28OCT2015	DC V3.00	NULL or missing			

vit_l3_ht_dist [2]

DATAMARTID	RESPONSE_DATE	QUERY_PACKAGE	STAT	RECORD_N
D1TEST	28OCT2015	DC V3.00	MIN	
D1TEST	28OCT2015	DC V3.00	MEAN	
D1TEST	28OCT2015	DC V3.00	MEDIAN	
D1TEST	28OCT2015	DC V3.00	MAX	
D1TEST	28OCT2015	DC V3.00	N	
D1TEST	28OCT2015	DC V3.00	NULL or missing	

1. PDF file limited to the 100 most frequent observations with counts above the low cell count threshold and sorted by descending record count.
2. NULL or missing rates are expected to be high due to the VITAL table structure

vit_l3_wt [2]

DATAMARTID	RESPONSE_DATE	QUERY_PACKAGE	WT_GROUP	RECORD_N	RECORD_PCT	DISTINCT_PATD_N
D1TEST	28OCT2015	DC V3.00	<0			
D1TEST	28OCT2015	DC V3.00	0-1			
D1TEST	28OCT2015	DC V3.00	2-6			
D1TEST	28OCT2015	DC V3.00	7-12			
D1TEST	28OCT2015	DC V3.00	13-20			
D1TEST	28OCT2015	DC V3.00	21-35			
D1TEST	28OCT2015	DC V3.00	36-50			
D1TEST	28OCT2015	DC V3.00	51-75			
D1TEST	28OCT2015	DC V3.00	76-100			
D1TEST	28OCT2015	DC V3.00	101-125			
D1TEST	28OCT2015	DC V3.00	126-150			
D1TEST	28OCT2015	DC V3.00	151-175			
D1TEST	28OCT2015	DC V3.00	176-200			
D1TEST	28OCT2015	DC V3.00	201-225			
D1TEST	28OCT2015	DC V3.00	226-250			
D1TEST	28OCT2015	DC V3.00	251-275			
D1TEST	28OCT2015	DC V3.00	276-300			
D1TEST	28OCT2015	DC V3.00	301-350			
D1TEST	28OCT2015	DC V3.00	>350			
D1TEST	28OCT2015	DC V3.00	NULL or missing			

vit_l3_wt_dist [2]

DATAMARTID	RESPONSE_DATE	QUERY_PACKAGE	STAT	RECORD_N
D1TEST	28OCT2015	DC V3.00	MIN	
D1TEST	28OCT2015	DC V3.00	MEAN	
D1TEST	28OCT2015	DC V3.00	MEDIAN	
D1TEST	28OCT2015	DC V3.00	MAX	
D1TEST	28OCT2015	DC V3.00	N	
D1TEST	28OCT2015	DC V3.00	NULL or missing	

1. PDF file limited to the 100 most frequent observations with counts above the low cell count threshold and sorted by descending record count.
2. NULL or missing rates are expected to be high due to the VITAL table structure

vit_l3_diastolic [2]

DATAMARTID	RESPONSE_DATE	QUERY_PACKAGE	DIASTOLIC_GROUP	RECORD_N	RECORD_PCT
D1TEST	28OCT2015	DC V3.00	<40		
D1TEST	28OCT2015	DC V3.00	40-60		
D1TEST	28OCT2015	DC V3.00	61-75		
D1TEST	28OCT2015	DC V3.00	76-80		
D1TEST	28OCT2015	DC V3.00	81-90		
D1TEST	28OCT2015	DC V3.00	91-100		
D1TEST	28OCT2015	DC V3.00	101-110		
D1TEST	28OCT2015	DC V3.00	111-120		
D1TEST	28OCT2015	DC V3.00	>120		
D1TEST	28OCT2015	DC V3.00	NULL or missing		

vit_l3_systolic [2]

DATAMARTID	RESPONSE_DATE	QUERY_PACKAGE	SYSTOLIC_GROUP	RECORD_N	RECORD_PCT
D1TEST	28OCT2015	DC V3.00	<40		
D1TEST	28OCT2015	DC V3.00	41-50		
D1TEST	28OCT2015	DC V3.00	51-60		
D1TEST	28OCT2015	DC V3.00	61-70		
D1TEST	28OCT2015	DC V3.00	71-80		
D1TEST	28OCT2015	DC V3.00	81-90		
D1TEST	28OCT2015	DC V3.00	91-100		
D1TEST	28OCT2015	DC V3.00	101-110		
D1TEST	28OCT2015	DC V3.00	111-120		
D1TEST	28OCT2015	DC V3.00	121-130		
D1TEST	28OCT2015	DC V3.00	131-140		
D1TEST	28OCT2015	DC V3.00	141-150		
D1TEST	28OCT2015	DC V3.00	151-160		
D1TEST	28OCT2015	DC V3.00	161-170		
D1TEST	28OCT2015	DC V3.00	171-180		
D1TEST	28OCT2015	DC V3.00	181-190		
D1TEST	28OCT2015	DC V3.00	191-200		
D1TEST	28OCT2015	DC V3.00	201-210		
D1TEST	28OCT2015	DC V3.00	>210		
D1TEST	28OCT2015	DC V3.00	NULL or missing		

1. PDF file limited to the 100 most frequent observations with counts above the low cell count threshold and sorted by descending record count.
2. NULL or missing rates are expected to be high due to the VITAL table structure

vit_l3_BMI [2]

DATAMARTID	RESPONSE_DATE	QUERY_PACKAGE	BMI_GROUP	RECORD_N	RECORD_PCT
D1TEST	28OCT2015	DC V3.00	<0		
D1TEST	28OCT2015	DC V3.00	0-1		
D1TEST	28OCT2015	DC V3.00	2-5		
D1TEST	28OCT2015	DC V3.00	6-10		
D1TEST	28OCT2015	DC V3.00	11-15		
D1TEST	28OCT2015	DC V3.00	16-20		
D1TEST	28OCT2015	DC V3.00	21-25		
D1TEST	28OCT2015	DC V3.00	26-30		
D1TEST	28OCT2015	DC V3.00	31-35		
D1TEST	28OCT2015	DC V3.00	36-40		
D1TEST	28OCT2015	DC V3.00	41-45		
D1TEST	28OCT2015	DC V3.00	46-50		
D1TEST	28OCT2015	DC V3.00	>50		
D1TEST	28OCT2015	DC V3.00	NULL or missing		

vit_l3_BP_position_type [2]

DATAMARTID	RESPONSE_DATE	QUERY_PACKAGE	BP_POSITION	RECORD_N	RECORD_PCT
D1TEST	28OCT2015	DC V3.00	01		
D1TEST	28OCT2015	DC V3.00	02		
D1TEST	28OCT2015	DC V3.00	03		
D1TEST	28OCT2015	DC V3.00	NI		
D1TEST	28OCT2015	DC V3.00	UN		
D1TEST	28OCT2015	DC V3.00	OT		
D1TEST	28OCT2015	DC V3.00	NULL or missing		
D1TEST	28OCT2015	DC V3.00	Values outside of CDM specifications		

1. PDF file limited to the 100 most frequent observations with counts above the low cell count threshold and sorted by descending record count.
2. NULL or missing rates are expected to be high due to the VITAL table structure

vit_l3_smoking [2]

DATAMARTID	RESPONSE_DATE	QUERY_PACKAGE	SMOKING	RECORD_N	RECORD_PCT
D1TEST	28OCT2015	DC V3.00	01		
D1TEST	28OCT2015	DC V3.00	02		
D1TEST	28OCT2015	DC V3.00	03		
D1TEST	28OCT2015	DC V3.00	04		
D1TEST	28OCT2015	DC V3.00	05		
D1TEST	28OCT2015	DC V3.00	06		
D1TEST	28OCT2015	DC V3.00	07		
D1TEST	28OCT2015	DC V3.00	08		
D1TEST	28OCT2015	DC V3.00	NI		
D1TEST	28OCT2015	DC V3.00	UN		
D1TEST	28OCT2015	DC V3.00	OT		
D1TEST	28OCT2015	DC V3.00	NULL or missing		
D1TEST	28OCT2015	DC V3.00	Values outside of CDM specifications		

vit_l3_tobacco [2]

DATAMARTID	RESPONSE_DATE	QUERY_PACKAGE	TOBACCO	RECORD_N	RECORD_PCT
D1TEST	28OCT2015	DC V3.00	01		
D1TEST	28OCT2015	DC V3.00	02		
D1TEST	28OCT2015	DC V3.00	03		
D1TEST	28OCT2015	DC V3.00	04		
D1TEST	28OCT2015	DC V3.00	06		
D1TEST	28OCT2015	DC V3.00	NI		
D1TEST	28OCT2015	DC V3.00	OT		
D1TEST	28OCT2015	DC V3.00	UN		
D1TEST	28OCT2015	DC V3.00	NULL or missing		
D1TEST	28OCT2015	DC V3.00	Values outside of CDM specifications		

1. PDF file limited to the 100 most frequent observations with counts above the low cell count threshold and sorted by descending record count.
2. NULL or missing rates are expected to be high due to the VITAL table structure

vit_l3_tobacco_type [2]

DATAMARTID	RESPONSE_DATE	QUERY_PACKAGE	TOBACCO_TYPE	RECORD_N	RECORD_PCT
D1TEST	28OCT2015	DC V3.00	01		
D1TEST	28OCT2015	DC V3.00	02		
D1TEST	28OCT2015	DC V3.00	03		
D1TEST	28OCT2015	DC V3.00	04		
D1TEST	28OCT2015	DC V3.00	05		
D1TEST	28OCT2015	DC V3.00	NI		
D1TEST	28OCT2015	DC V3.00	UN		
D1TEST	28OCT2015	DC V3.00	OT		
D1TEST	28OCT2015	DC V3.00	NULL or missing		
D1TEST	28OCT2015	DC V3.00	Values outside of CDM specifications		

vit_l3_dash1 (This query counts the number of patients with any vital record with a populated MEASURE_DATE during the designated period of time prior to the maximum MEASURE_DATE. If the maximum MEASURE_DATE is in the future the current date is substituted).

DATAMARTID	RESPONSE_DATE	QUERY_PACKAGE	PERIOD	DISTINCT_PATID_N
D1TEST	28OCT2015	DC V3.00	1 yr	
D1TEST	28OCT2015	DC V3.00	2 yrs	
D1TEST	28OCT2015	DC V3.00	3 yrs	
D1TEST	28OCT2015	DC V3.00	4 yrs	
D1TEST	28OCT2015	DC V3.00	5 yrs	
D1TEST	28OCT2015	DC V3.00	All yrs	

1. PDF file limited to the 100 most frequent observations with counts above the low cell count threshold and sorted by descending record count.
2. NULL or missing rates are expected to be high due to the VITAL table structure

XII. Table Shells: HARVEST and Cross-Table Queries

xtbl_l3_dates (this query counts the minimum and maximum date and the number of records with non-missing dates, future dates, and dates prior to Jan 2010)

DATAMARTID	RESPONSE_ DATE	QUERY_ PACKAGE	DATASET	TAG	MIN	MAX	N	FUTURE _DT_N	PRE-2010_ N
D1TEST	28OCT2015	DC V3.00	DEMOGRAPHIC	BIRTH_DATE					
D1TEST	28OCT2015	DC V3.00	ENCOUNTER	ADMIT_DATE					
D1TEST	28OCT2015	DC V3.00	ENCOUNTER	DISCHARGE_ DATE					
D1TEST	28OCT2015	DC V3.00	DIAGNOSIS	ADMIT_DATE					
D1TEST	28OCT2015	DC V3.00	PROCEDURES	ADMIT_DATE					
D1TEST	28OCT2015	DC V3.00	PROCEDURES	PX_DATE					
D1TEST	28OCT2015	DC V3.00	VITAL	MEASURE_ DATE					
D1TEST	28OCT2015	DC V3.00	ENROLLMENT	ENR_START_ DATE					
D1TEST	28OCT2015	DC V3.00	ENROLLMENT	ENR_END_ DATE					

xtbl_l3_metadata (this query captures metadata about the DATAMART from the HARVEST table, information about this query package, and SAS configuration information. There should only be 1 record in the table. Please verify your NETWORKID, NETWORK_NAME, DATAMARTID, and DATAMART_NAME against the HARVEST Reference Table (https://pcornet.imeetcentral.com/p/aQAAAAACl2-K).

DATAMARTID	RESPONSE_DATE	QUERY_PACKAGE	TAG	VALUE
D1TEST	28OCT2015	DC V3.00	NETWORKID	
D1TEST	28OCT2015	DC V3.00	NETWORK_NAME	
D1TEST	28OCT2015	DC V3.00	DATAMARTID	
D1TEST	28OCT2015	DC V3.00	DATAMART_NAME	
D1TEST	28OCT2015	DC V3.00	DATAMART_PLATFORM	
D1TEST	28OCT2015	DC V3.00	CDM_VERSION	
D1TEST	28OCT2015	DC V3.00	DATAMART_CLAIMS	
D1TEST	28OCT2015	DC V3.00	DATAMART_EHR	
D1TEST	28OCT2015	DC V3.00	BIRTH_DATE_MGMT	
D1TEST	28OCT2015	DC V3.00	ENR_START_DATE_MGMT	
D1TEST	28OCT2015	DC V3.00	ENR_END_DATE_MGMT	
D1TEST	28OCT2015	DC V3.00	ADMIT_DATE_MGMT	
D1TEST	28OCT2015	DC V3.00	DISCHARGE_DATE_MGMT	
D1TEST	28OCT2015	DC V3.00	PX_DATE_MGMT	
D1TEST	28OCT2015	DC V3.00	RX_ORDER_DATE_MGMT	
D1TEST	28OCT2015	DC V3.00	RX_START_DATE_MGMT	
D1TEST	28OCT2015	DC V3.00	RX_END_DATE_MGMT	
D1TEST	28OCT2015	DC V3.00	DISPENSE_DATE_MGMT	

1. PDF file limited to the 100 most frequent observations with counts above the low cell count threshold and sorted by descending record count.
2. NULL or missing rates are expected to be high due to the VITAL table structure

DATAMARTID	RESPONSE_DATE	QUERY_PACKAGE	TAG	VALUE
D1TEST	28OCT2015	DC V3.00	LAB_ORDER_DATE_MGMT	
D1TEST	28OCT2015	DC V3.00	SPECIMEN_DATE_MGMT	
D1TEST	28OCT2015	DC V3.00	RESULT_DATE_MGMT	
D1TEST	28OCT2015	DC V3.00	MEASURE_DATE_MGMT	
D1TEST	28OCT2015	DC V3.00	ONSET_DATE_MGMT	
D1TEST	28OCT2015	DC V3.00	REPORT_DATE_MGMT	
D1TEST	28OCT2015	DC V3.00	RESOLVE_DATE_MGMT	
D1TEST	28OCT2015	DC V3.00	PRO_DATE_MGMT	
D1TEST	28OCT2015	DC V3.00	REFRESH_DEMOGRAPHIC_DATE	
D1TEST	28OCT2015	DC V3.00	REFRESH_ENROLLMENT_DATE	
D1TEST	28OCT2015	DC V3.00	REFRESH_ENCOUNTER_DATE	
D1TEST	28OCT2015	DC V3.00	REFRESH_DIAGNOSIS_DATE	
D1TEST	28OCT2015	DC V3.00	REFRESH_PROCEDURES_DATE	
D1TEST	28OCT2015	DC V3.00	REFRESH_VITAL_DATE	
D1TEST	28OCT2015	DC V3.00	REFRESH_DISPENSING_DATE	
D1TEST	28OCT2015	DC V3.00	REFRESH_LAB_RESULT_CM_DATE	
D1TEST	28OCT2015	DC V3.00	REFRESH_CONDITION_DATE	
D1TEST	28OCT2015	DC V3.00	REFRESH_PRO_CM_DATE	
D1TEST	28OCT2015	DC V3.00	REFRESH_PRESCRIBING_DATE	
D1TEST	28OCT2015	DC V3.00	REFRESH_PCORNET_TRIAL_DATE	
D1TEST	28OCT2015	DC V3.00	REFRESH_DEATH_DATE	
D1TEST	28OCT2015	DC V3.00	REFRESH_DEATH_CAUSE_DATE	
D1TEST	28OCT2015	DC V3.00	REFRESH_MAX	
D1TEST	28OCT2015	DC V3.00	LOW_CELL_CNT	
D1TEST	28OCT2015	DC V3.00	OPERATING_SYSTEM	
D1TEST	28OCT2015	DC V3.00	QUERY_PACKAGE	
D1TEST	28OCT2015	DC V3.00	RESPONSE_DATE	
D1TEST	28OCT2015	DC V3.00	SAS_VERSION	
D1TEST	28OCT2015	DC V3.00	SAS_BASE	
D1TEST	28OCT2015	DC V3.00	SAS_GRAPH	
D1TEST	28OCT2015	DC V3.00	SAS_STAT	
D1TEST	28OCT2015	DC V3.00	SAS_ETS	
D1TEST	28OCT2015	DC V3.00	SAS_AF	
D1TEST	28OCT2015	DC V3.00	SAS_IML	
D1TEST	28OCT2015	DC V3.00	SAS_CONNECT	
D1TEST	28OCT2015	DC V3.00	SAS_MYSQL	
D1TEST	28OCT2015	DC V3.00	SAS_ODBC	
D1TEST	28OCT2015	DC V3.00	SAS_ORACLE	

1. PDF file limited to the 100 most frequent observations with counts above the low cell count threshold and sorted by descending record count.
2. NULL or missing rates are expected to be high due to the VITAL table structure

DATAMARTID	RESPONSE_DATE	QUERY_PACKAGE	TAG	VALUE
D1TEST	28OCT2015	DC V3.00	SAS_POSTGRES	
D1TEST	28OCT2015	DC V3.00	SAS_SQL	
D1TEST	28OCT2015	DC V3.00	SAS_TERADATA	

xtbl_l3_dash1 (This query counts the number of patients with any Vital record with a populated MEASURE_DATE and a Diagnosis record with a populated DX and ADMIT_DATE during the designated period of time prior to the maximum ADMIT_DATE. If the maximum ADMIT_DATE is in the future the current date is substituted).

DATAMARTID	RESPONSE_DATE	QUERY_PACKAGE	PERIOD	DISTINCT_PATID_N
D1TEST	28OCT2015	DC V3.00	1 yr	
D1TEST	28OCT2015	DC V3.00	2 yrs	
D1TEST	28OCT2015	DC V3.00	3 yrs	
D1TEST	28OCT2015	DC V3.00	4 yrs	
D1TEST	28OCT2015	DC V3.00	5 yrs	
D1TEST	28OCT2015	DC V3.00	All yrs	

1. PDF file limited to the 100 most frequent observations with counts above the low cell count threshold and sorted by descending record count.
2. NULL or missing rates are expected to be high due to the VITAL table structure

XIII. Running the SAS query

1) Confirm that your SAS data files are locked/static (ie, will not be changed until the next DataMart refresh) and that your HARVEST table contains the correct DATAMARTID and NETWORKID (https://pcornet.imeetcentral.com/p/aQAAAAACl2-K).
2) Open the zip file distributed with the query and extract the contents. Save the run_queries.sas program in the *sasprograms* folder, and the data_characterization.sas program in the *infolder* folder.
3) Open run_queries.sas. You will need to modify the directory paths. For reasons of compatibility and standardization, directory paths must meet the following criteria:
    i) DO use forward slashes (e.g. /) which are always compatible on both UNIX and WINDOWS.
    ii) DO use end of path separators (e.g. /xyz/ and not /xyz) which are assumed by many programs.
    iii) DO use beginning of path separators (e.g. /xyz) on UNIX.
    iv) DO NOT use beginning of path separators on WINDOWS (e.g. P:/xyz not /P:/xyz).
    v) DO NOT surround directory paths with quotes (e.g. /xyz/ not "/xyz/").
4) If your CDM data is stored in SAS datasets, modify the user inputs in run_queries.sas as instructed below. Otherwise, go to Step 5. DO NOT modify any other part of the program.
    a) After `%let dpath=,` provide the directory path where your PCORnet CDM v3.0 SAS data is located.
    b) After `%let qpath=,` provide the outer folder where the required folders were created.
    c) After `%let threshold=`, provide the low cell count threshold value for the DataMart.
5) If your CDM data is stored in database tables (e.g. SAS data step views of RDBMS tables), open data_characterization.sas and run_queries.sas and modify the user inputs as follows. DO NOT modify any other part of the program.
    a) Use the appropriate SAS/ACCESS options on a LIBNAME statement so that the program knows where to find the database tables. The examples below show connection information for an Oracle database; connecting to other database systems may require different connection information.
        i) In the run.queries.sas program, edit the `dpath` variable on Line 60 to include the appropriate database connection information. Be sure to use the `%str()` function to mask the embedded equal signs. For example: `%let dpath = %str(oracle user="myuserid" orapw=mypasswd path=mydbname schema=myschema);`
        ii) In the data_characterization.sas program, edit the libname pcordata statement on Line 28 to remove the quotation marks, as: `libname pcordata &dpath;`
    b) Modify the user inputs in the run_queries.sas program following steps 4b and 4c.
6) Save and run the run.queries.sas program. The program will call data_characterization.sas and process the queries. As it processes each query, the program will print results to a PDF file, create permanent SAS datasets for each output table, create log files, and import all of the permanent SAS datasets into the SAS transport file (.cpt).
7) Review the log files for errors and warnings, the data output for values outside of CDM specifications or other unexpected results, and the names of the files in the *drnoc* folder to ensure that they conform to the naming conventions listed in Section V. If desired, you may verify the contents of the cpt file by using use a `proc cimport` statement, as shown in the example below:

```
libname outlib 'F:/pcornet/myproject/';
%let infile= 'F:/pcornet/myproject/T1D3_20151101_Data_Characterization.cpt';
proc cimport infile=&infile library=outlib;
run;
```

8) Zip the three files in the *drnoc* folder and your completed ETL ADD and save them as DC_Query_Pkg_V300_Results.zip. Return this file to the PCORnet DRN OC via the DRN Query Tool.

1. PDF file limited to the 100 most frequent observations with counts above the low cell count threshold and sorted by descending record count.
2. NULL or missing rates are expected to be high due to the VITAL table structure