

## PCORnet Data Checks, version 8

### Introduction

These data checks are included in PCORnet's foundational data curation package. All network partners must pass Data Checks 1.01-1.13, 3.04 and 3.05. Exceptions to other checks are permitted but must be explained in an online annotated data dictionary.

### Data Model Conformance Checks

Check	Description	Changes from version 7
DC 1.01	Required tables are not present	None
DC 1.02	Required tables are not populated	None
DC 1.03	Required fields are not present	None
DC 1.04	Required fields do not conform to data model specifications for data type, length, or name.	None
DC 1.05	Tables have primary key definition errors	None
DC 1.06	Required fields contain values outside of data model specifications	None
DC 1.07	Required fields have non-permissible missing values	None
DC 1.08	Tables contain orphan PATIDs	None
DC 1.09	Tables contain orphan ENCOUNTERIDs	None
DC 1.10	Replication errors between the ENCOUNTER, PROCEDURES and DIAGNOSIS tables	None
DC 1.11	More than 5% of encounters are assigned to more than one patient	None
DC 1.12	Tables contain orphan PROVIDERIDs	None
DC 1.13	More than 5% of ICD, CPT, LOINC, RXCUI, or NDC codes do not conform to the expected length or content	None
DC 1.14	Patients in the DEMOGRAPHIC table are not in the HASH_TOKEN table	New

## PCORnet Data Checks, version 8

### Data Plausibility Checks

Check	Description	Changes from version 7
DC 2.01	More than 5% of records have future dates	None
DC 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	None
DC 2.03	More than 5% of patients have illogical date relationships	Added DX_DATE
DC 2.04	The average number of encounters per visit is > 2.0 for inpatient (IP), emergency department (ED), or ED to inpatient (EI) encounters	None
DC 2.05	More than 5% of results for selected laboratory tests do not have the appropriate specimen source	None
DC 2.06	The median lab result value for selected tests is an outlier.	Increased the specificity of the outlier definition. Outliers are now defined as results outside of both the statistical and clinical normal range.
DC 2.07	The average number of principal diagnoses per encounter is above threshold [2.0 for inpatient (IP) and ED to inpatient (EI)]	None
DC 2.08	The monthly volume of encounter, diagnosis, procedure, vital, prescribing, or laboratory records is an outlier.	None

## PCORnet Data Checks, version 8

### Data Completeness Checks

Check	Description	Changes from version 7
DC 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]	None
DC 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]	None
DC 3.03	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	None
DC 3.04	Less than 50% of patients with encounters have DIAGNOSIS records	None
DC 3.05	Less than 50% of patients with encounters have PROCEDURES records	None
DC 3.06	More than 10% of IP (inpatient) or ED to inpatient (EI) encounters with any diagnosis don't have a principal diagnosis	None
DC 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	None
DC 3.08	Less than 80% of prescribing orders are mapped to a RXNORM_CUI which fully specifies the ingredient, strength and dose form	None
DC 3.09	Less than 80% of laboratory results are mapped to LAB_LOINC	None
DC 3.10	Less than 80% of quantitative results for tests mapped to LAB_LOINC fully specify the normal range	None
DC 3.11	Vital, prescribing, or laboratory records are less than 75% complete three months prior to the current month	None
DC 3.12	Less than 80% of quantitative results for tests mapped to LAB_LOINC fully specify the RESULT_UNIT	Removed SPECIMEN_SOURCE requirement
DC 3.13	The percentage of patients with selected lab tests is below threshold.	New

## PCORnet Data Checks, version 8

### Data Persistence Checks

Check	Description	Changes from version 7
DC 4.01	More than a 5% decrease in the number of patients or records in a CDM table	None
DC 4.02	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.	None
DC 4.03	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.	None