UTOPIAN Data Safe Haven - Data Dictionary

Version 1.4.0.3

General Information

Color-coding Legend of data completion rate:

	95 – 100%	HIGH
	75 – 94%	MEDIUM – HIGH
	50 – 74%	MEDIUM
	25 – 49%	LOW- MEDIUM
	0 – 24%	POOR
		The field is derived from an automatic algorithm

Revision Notes:

POINT TO VERSION 3.0 AND 4.0 TRANSITION DOCUMENTS (WHEN THEY EXIST).

Table: Network

	Field	Definition	Field Notes
	Network_ID	Unique identifier for each Network	Unique number assigned to each network from the central repository e.g. '9' for NS
	NetworkName	Name of the Network	Name of the Network in Acronym form (e.g., SAPCReN, MaRNet)
	GeographicArea	Geographical area of the Network	Examples. Southern Alberta; Northern Alberta; Maritimes, etc.

Table: Provider

	Field	Definition	Field Notes
	Provider_ID	Unique identifier for each provider	Unique number assigned by each network to the providers in its sites e.g. '15'
	BirthYear	Provider's year of birth	The field has a medium-high completion rate
	Sex	Provider's sex	Permitted values are 'Male' or 'Female' The field has a high completion rate

Table: Site

	Field	Definition	Field Notes
	Network_ID	Unique identifier for each Network	Only a Network_ID that exists in the Network table can be referenced here Unique number assigned to each network from the central repository e.g. '9' for NS Forms a Composite Primary Key with Site_ID
	Site_ID	Unique identifier for each Site in the Site_ID format	Site_ID: Unique number assigned by the network to identify each of its sites e.g. '1' for first site Site_ID remains constant across cycles Forms a Composite Primary Key with Network_ID
	LocationType	Type of location of the provider's practice	Permitted values enforced by check constraints e.g. 'Primary Care Clinic', 'Walk-in Clinic'
	PostalCode	Postal code of the Site location	e.g. 'X2X 2X2'
	Province	A unique 2-character province name	Permitted values enforced by check constraints e.g. 'AB', 'NS'

	EMRName	Name of the EMR used by the participating site	Permitted values enforced by check constraints Med Access; Wolf; Jonoke; Nightingale; etc.

Table: SiteProvider

	Field	Definition	Field Notes
	Network_ID	Foreign Key from the Network table	Only a Network_ID that exists in the Network table can be referenced here Forms a Composite Foreign Key with Site_ID Forms a Unique constraint with Site_ID and Provider_ID
	Site_ID	Foreign Key from the Site table	Only a Site_ID that exists in the Site table can be referenced here Forms a Composite Foreign Key with Network_ID Forms a Unique constraint with Network_ID and Provider_ID
	Provider_ID	Provider_ID from the Provider table	Only a Provider_ID that exists in the Provider table can be referenced here Forms a Unique constraint with Network_ID and Site_ID
	ProviderType	Role of this provider	Permitted values enforced by check constraints e.g. 'Family Physician', 'Nurse Practitioner' The field has a high completion rate
	ProviderRole	Role of this provider	For now, Role = ProviderType The field has a high completion rate
	StartDate	Date that the provider starts participating in CPCSSN at the site	Can either be the date the provider began with CPCSSN or the date that the provider moved to a new site

Table: GroupInfo

Table Notes:
Not all network use it.

	Field	Definition	Field Notes
	Network_ID	Foreign Key from the Network table	Only a Network_ID that exists in the Network table can be referenced here Forms a Composite Foreign Key with Site_ID
	Site_ID	Foreign Key from the Site table	Only a Site_ID that exists in the Site table can be referenced here Forms a Composite Foreign Key with Network_ID
	GroupName	Name of the group	e.g. Toronto Family Health Team, ASA Study
	GroupType	Type of group	e.g. Family Health Team, Study
	PaymentModel	Capitation model	
	CareModel	How care is provided	
	GovernanceModel	How the group is managed	
	Description	Other descriptions that may not fit in the above categories	

Table: ProviderGroup

Table Notes:
All of the Provider-group pairings.

	Field	Definition	Field Notes
	Provider_ID	Foreign Key from the Provider table	Only a Provider_ID that exists in the Provider table can be referenced here Forms a Unique constraint with GroupInfo_ID
	GroupInfo_ID	Foreign Key from the GroupInfo table	Only a GroupInfo_ID that exists in the GroupInfo table can be referenced here Forms a Unique constraint with Provider_ID

Table: Patient

Table Notes:
List of EMR patients whose primary provider is a consenting physician in the CPCSSN project (has an entry in the Provider table).

	Field	Definition	Field Notes
	Patient_ID	Unique and randomised patient ID	Assign each patient a meaningless number e.g. '123456'
	Sex	Patient's sex	Permitted values enforced by check constraints e.g. 'Male' or 'Female' The field has a high completion rate with a few missing data
	BirthYear	4-digit year of patient's birth date	Must be greater than or equal to 1850 The field has a high completion rate with a few missing data
	BirthMonth	Numerical value of patient's birth month	The field has a low completion rate due to REB restriction
	OptedOutDate	Date on which the patient opted out	

Table: PatientProvider

Table Notes:
This table identifies which patients are assigned to which providers. Notice that one patient may be assigned to multiple providers.

	Field	Definition	Field Notes
	Network_ID	Foreign Key from the Network table	Only a Network_ID that exists in the Network table can be referenced here Forms a Composite Foreign Key with Site_ID and Provider_ID
	Site_ID	Foreign Key from the Site table	Only a Site_ID that exists in the Site table can be referenced here Forms a Composite Foreign Key with Network_ID and Provider_ID
	Patient_ID	Foreign Key from the Patient table	Only a Patient_ID that exists in the Patient table can be referenced here
	Provider_ID	Foreign Key from the Provider table	Only a Provider_ID that exists in the Provider table can be referenced here Forms a Composite Foreign Key with Network_ID and Site_ID
	PatientType	How the patient is identified in the provider's roster	Suggested use as an indicator for rostered patients Not all patients have this information
	StartDate	Date that the provider starts providing care to the patient at this site	The field can be the date of the first encounter with the provider

Table: Allergy Intolerance

Table Notes:
All allergy and intolerance data for the patient.

	Field	Definition	Field Notes
	Network_ID	Foreign Key from the Network table	Only a Network_ID that exists in the Network table can be referenced here Forms a Composite Foreign Key with Site_ID
	Site_ID	Foreign Key from the Site table	Only a Site_ID that exists in the Site table can be referenced here Forms a Composite Foreign Key with Network_ID
	Patient_ID	Foreign Key from Patient table	Only a Patient_ID that exists in the Patient table can be referenced here
	Encounter_ID	Foreign Key from Encounter table	Only an Encounter_ID that exists in the Encounter table can be used May be null if the EMR does not have the data
	StartDate	Date that the allergy was first identified	The field can be the date of the encounter when the patient discusses the problem with the doctor, or the date the doctor put the information into the EMR. It is an approximation to the exact date when the allergy intolerance first occurs
	StopDate	Date that the allergy noted as inactive	If an allergy is inactive and StopDate cannot be found, Status field must still be set to 'Inactive'
	DIN	DIN for the medication	Extract this where available
	Name_orig	Name of the allergy as it appears in the EMR	
	CodeType_orig	Allergy CodeType of the allergy exactly as it appears in the EMR	e.g. 'ATC'
	Code_orig	Allergy code exactly as it appears in the EMR	
	Category	The category of the allergy	Permitted values enforced by check constraints e.g. 'Medication', 'Non-Medication', 'Medication Intolerance' The field has a medium completion rate. The null value in the field only means the information is not available in EMR.
	Severity	The severity of the reaction	Permitted values enforced by check constraints e.g. 'Mild', 'Moderate', 'Severe' The field has a medium completion rate. The null value in the field only means the information is not available in EMR
	AllergyStatus	Current status of the allergy	Permitted values enforced by check constraints (e.g. 'Active', 'Inactive') If allergy is inactive, this field is always filled, even if no StopDate exists The field has a low-medium completion rate. The null value in the field only means the information is not available in EMR.
	ReactionType	Type of reaction that occurs with allergy	e.g. 'Rash', 'Anaphylaxis', etc. The field has a medium completion rate and is in free text
	DateCreated	EMR date stamp of the record	The field indicates when the record is created in the EMR. It may not suitable as an approximation to StartDate
	Name_calc	Name_orig recoded into consistent text	The field is coded with an automatic algorithm. It may consist of errors. The error rate is low when Name_orig is indeed a medication
	CodeType_calc	Code set name for the Name_calc entry	e.g. 'ATC'
	Code_calc	Allergy code in the CodeType_calc code set	The field is coded with an automatic algorithm. It may consist of errors. The error rate is low when Name_orig is indeed a medication

Table: Billing

Table Notes:
All billing data submitted to the province for the patient. Notice that billing record may not be available for salaried doctors and the clinics not using EMR for billing.

	Field	Definition	Field Notes
	Network_ID	Foreign Key from the Network table	Only a Network_ID that exists in the Network table can be referenced here Forms a Composite Foreign Key with Site_ID
	Site_ID	Foreign Key from the Site table	Only a Site_ID that exists in the Site table can be referenced here Forms a Composite Foreign Key with Network_ID
	Patient_ID	Foreign Key from Patient table	Only a Patient_ID that exists in the Patient table can be referenced here
	Encounter_ID	Foreign Key from Encounter table	Only an Encounter_ID that exists in the Encounter table can be used May be null if the EMR does not have the data
	ServiceDate	Date the billing was performed/submitted	
	ServiceCode	Service code associated with the billing	e.g. '03.03A' in Alberta
	DiagnosisText_orig	Text exactly as it appears in the EMR	
	DiagnosisCodeType_orig	Type of code for any diagnosis codes attached to the billing as it appears in the EMR	e.g. 'ICD9'
	DiagnosisCode_orig	Diagnosis code associated with the billing as it appears in the EMR	e.g. '250.0'
	DateCreated	EMR date stamp of the record	The field indicates when the record is created in the EMR. It may not suitable as an approximation to ServiceDate
	DiagnosisText_calc	DiagnosisText_orig recoded into consistent text	The field is coded with an automatic algorithm and may consist errors, but the error rate is low
	DiagnosisCodeType_calc	DiagnosisCodeType_orig recoded into consistent text	e.g. 'ICD9'
	DiagnosisCode_calc	DiagnosisCode_orig recoded into consistent text	The field is coded with an automatic algorithm and may consist errors, but the error rate is low

Table: Encounter

Table Notes:
An encounter is an interaction of the patient with a provider in some fashion. The provider does not need to be a participating provider. It consists of all encounters of the patient.

	Field	Definition	Field Notes
	Encounter_ID	Unique identifier of an encounter	The field links to EncounterDiagnosis, Medication, Exam, and others, to provide a holistic picture of an encounter. However, the linkage can be missing.
	Network_ID	Foreign Key from the Network table	Only a Network_ID that exists in the Network table can be referenced here Forms a Composite Foreign Key with Site_ID and Provider_ID
	Site_ID	Foreign Key from the Site table	Only a Site_ID that exists in the Site table can be referenced here Forms a Composite Foreign Key with Network_ID and Provider_ID
	Patient_ID	Foreign Key from the Patient table	Only a Patient_ID that exists in the Patient table can be referenced here
	Provider_ID	Responsible provider for this encounter	The composite foreign key requires that the Network_ID, Site_ID, and Provider_ID MUST be in the SiteProvider table. If the provider for the encounter is not a CPCSSN sentinel, assign NULL Forms a Composite Foreign Key with Network_ID and Site_ID
	EncounterDate	Date the encounter occurred	
	Reason_orig	Text exactly as it appears in the EMR	If there is more than one source of this data, then precede the entry with 'Patient: ' and 'Provider: ' as appropriate The field has a low-medium completion rate and is in free text
	EncounterType	How or where the encounter was conducted	Permitted values enforced by check constraints e.g. ‘Phone’, ‘Walk-In’, ‘ER Visit’ Extract this data if available If recording is not straightforward, populate with original text and indicate this in the extraction notes The field has a low-medium completion rate. The null value only implies the information is not available in EMR
	DateCreated	Date the record was created in the EMR	The field indicates when the record is created in the EMR. It may not suitable as an approximation to EncounterDate
	Reason_calc	Re-coded or cleaned version of Reason_orig	We will recode this in a future cycle, do not fill in Permitted values NOT enforced by check constraints, they are the name of the coding set The field is not implemented yet

Table: EncounterDiagnosis

Table Notes:
All diagnoses resulting from an encounter with the patient. Notice not all encounters have encounter diagnoses.

	Field	Definition	Field Notes
	Network_ID	Foreign Key from Network table	Only a Network_ID that exists in the Network table can be referenced here Forms a Composite Foreign Key with Site_ID
	Site_ID	Foreign Key from Site table	Only a Site_ID that exists in the Site table can be referenced here Forms a Composite Foreign Key with Network_ID
	Patient_ID	Foreign Key from Patient table	Only a Patient_ID that exists in the Patient table can be referenced here
	Encounter_ID	Foreign Key from Encounter table	Only an Encounter_ID that exists in the Encounter table can be used May be null if the EMR does not have the data
	DiagnosisText_orig	Text exactly as it appears in the EMR	This is the final physician diagnosis, not the presenting complaint of the visit
	DiagnosisCodeType_orig	Name of the code set from which the original diagnosis code was taken	Populate this field ONLY if there is a value in the DiagnosisCode_orig field for this record e.g. 'ICD9', 'SNOMED'
	DiagnosisCode_orig	Original diagnosis code from EMR	Populate this only if it is already available in the EMR. Do not clean any data
	DateCreated	EMR date stamp of the record	The field indicates when the record is created in the EMR. It may not suitable as an approximation to EncounterDate
	DiagnosisText_calc	DiagnosisText_orig recoded into consistent text	Permitted values NOT enforced by check constraints, they are from the associated coding set The field is coded with an automatic algorithm and may consist errors, but the error rate is low
	DiagnosisCodeType_calc	DiagnosisCodeType_orig recoded into consistent text	Permitted values NOT enforced by check constraints, they are from the associated coding set e.g. 'ICD9'
	DiagnosisCode_calc	DiagnosisCode_orig recoded into consistent text	Permitted values NOT enforced by check constraints, they are from the associated coding set The field is coded with an automatic algorithm and may consist errors, but the error rate is low

Table: Exam

Table Notes:

Results of physical exams performed on the patient.
Extract only the designated physical exams includes blood pressure, weight, height, BMI, waist circumference, waist hip ratio, food exam, and PEFR.
A good number of BP, weight, height, BMI but waist hip ratio, food exam, and PEFR are scarce.

	Field	Definition	Field Notes
	Network_ID	Foreign Key from Network table	Only a Network_ID that exists in the Network table can be referenced here Forms a Composite Foreign Key with Site_ID
	Site_ID	Foreign Key from Site table	Only a Site_ID that exists in the Site table can be referenced here Forms a Composite Foreign Key with Network_ID
	Patient_ID	Foreign Key from Patient table	Only a Patient_ID that exists in the Patient table can be referenced here
	Encounter_ID	Foreign Key from Encounter table	Only Encounter_IDs existing in the Encounter table can be used May be null if the EMR does not have the data
	Exam1	Name of the physical exam; recoded into consistent text	Permitted values enforced by check constraints e.g. 'BMI (kg/m^2)', 'dBP (mmHg)', 'Foot Exam', 'sBP (mmHg)'
	Result1_orig	Result of the physical exam	Use consistent units. Units are contained in the 'Exam' name
	Exam2	Name of the paired physical exam	Currently the only paired exam is blood pressure 'Exam1' must be 'sBP (mmHg)' and 'Exam2' must be 'dBP (mmHg)'
	Result2_orig	Second result of paired physical exam	Currently the only paired exam is blood pressure 'Result1' must be a systolic value and 'Result2' must be a diastolic value
	UnitOfMeasure_orig	Units that accompany the ExamResult	Use consistent units. Units are contained in the 'Exam' name
	PairingMethod	PairingMethod for physical exam	Permitted values enforced by check constraints PairingMethod is only filled in for exams that have paired results, e.g. blood pressures, and **both** Exam1 and Exam2 fields are populated e.g. 'Algorithm', 'Pre-paired'
	DateCreated	EMR date stamp of the record	The field indicates when the record is created in the EMR. It may not suitable as an approximation to Exam date
	Result1_calc	Re-coding of Result1_orig	Populated by Exam Coding algorithm The field is calculated with an automatic algorithm that cleans the value and adjusts the unit of measurement. It may consist errors but has better data quality than Result_orig1
	Result2_calc	Re-coding of Result2_orig	Populated by Exam Coding algorithm The field is calculated with an automatic algorithm that cleans the value and adjusts the unit of measurement. It may consist errors but has better data quality than Result_orig2
	UnitOfMeasure_calc	Re-coding of UnitOfMeasure_orig	Populated by Exam Coding algorithm The field is calculated with an automatic algorithm that cleans the value and adjusts the unit of measurement. It may consist errors but has better data quality than UnitOfMeasure_orig

Table: FamilyHistory

Table Notes:
Family history of the patient.

	Field	Definition	Field Notes
	Network_ID	Foreign Key from Network table	Only a Network_ID that exists in the Network table can be referenced here Forms a Composite Foreign Key with Site_ID
	Site_ID	Foreign Key from Site table	Only a Site_ID that exists in the Site table can be referenced here Forms a Composite Foreign Key with Network_ID
	Patient_ID	Foreign Key from Patient table	Only a Patient_ID that exists in the Patient table can be referenced here
	Encounter_ID	Foreign Key from Encounter table	Only an Encounter_ID that exists in the Encounter table can be used May be null if EMR does not have the data
	DiagnosisText_orig	Text exactly as it appears in the EMR	
	DiagnosisCodeType_orig	Name of the code set from which the original diagnosis code was taken	Populate this field ONLY if there is a value in the DiagnosisCode_orig field for this record e.g. 'ICD9', 'SNOMED'
	DiagnosisCode_orig	Original diagnosis code from EMR	Populate this only if it is available in the EMR. Do not clean any data
	Relationship_orig	Original relationship from the EMR	The field has a low completion rate and is in free text
	AgeAtOnset	Age of onset of the condition	The field has a low completion rate
	VitalStatus	Whether relative is alive or deceased	The field has a low completion rate and is in free text
	WasCauseOfDeath	Was this condition the cause of death	
	AgeAtDeath	Relation's age at death	The field has a low completion rate
	DateCreated	EMR date stamp of the record	The field indicates when the record is created in the EMR. It may not suitable to approximate AgeAtOnset
	DiagnosisText_calc	DiagnosisText_orig recoded into consistent text	Filled by health condition coding algorithm Values are from the associated coding set, NOT enforced by check constraints The field is coded with an automatic algorithm. It may consist of errors but the error rate is low.
	DiagnosisCodeType_calc	DiagnosisCodeType_orig recoded into consistent text	Filled by health condition coding algorithm, e.g. 'ICD9'
	DiagnosisCode_calc	DiagnosisCode_orig recoded into consistent text	Filled by health condition coding algorithm Values are from the associated coding set, NOT enforced by check constraints The field is coded with an automatic algorithm. It may consist of errors but the error rate is low.
	RelationshipSide_calc	Coding of side of the relationship	e.g. 'Maternal' or 'Paternal' The field is not implemented yet
	RelationshipDegree_calc	Genetic degree of the relationship	Relationship degree defined here: http://www.cdc.gov/genomics/resources/diseases/breast_ovarian_cancer/risk_categories.htm The field is not implemented yet

Table: HealthCondition

Table Notes:
All health conditions of the patient.

	Field	Definition	Field Notes
	Network_ID	Foreign Key from Network table	Only a Network_ID that exists in the Network table can be referenced here Forms a Composite Foreign Key with Site_ID
	Site_ID	Foreign Key from Site table	Only Site_IDs that exist in the Site table can be referenced Forms a Composite Foreign Key with Network_ID
	Patient_ID	Foreign Key from Patient table	Only a Patient_ID that exists in the Patient table can be referenced here
	Encounter_ID	Foreign Key from Encounter table	Only an Encounter_ID that exists in the Encounter table can be used May be null if the EMR does not have the data
	DiagnosisText_orig	Text exactly as it appears in the EMR	
	DiagnosisCodeType_orig	Name of the code set from which the original diagnosis code was taken	Populate this field ONLY if there is a value in the DiagnosisCode_orig field for this record e.g. 'ICD9', 'SNOMED'
	DiagnosisCode_orig	Original diagnosis code from the EMR	Populate this only if it is already available in the EMR. Do not clean any data
	DateOfOnset	Date that the health condition began	The field has a low completion rate and is likely to consist of data entry errors
	SignificantNegativeFlag	An indicator that the Patient does NOT have this health condition	'True': does not have this condition 'False'/NULL: has this condition
	ActiveInactiveFlag	An indicator that the condition is active at the time of data extraction.	Permitted values enforced by check constraints If it is in the problem list, it is likely to be Active. If it is in the Past Medical History, it is likely to be Inactive The field has a high completion rate
	DateCreated	EMR date stamp of the record	The field indicates when the record is created in the EMR. It may not suitable as an approximation to DateOfOnset
	DiagnosisText_calc	DiagnosisText_orig recoded into consistent text	Populated by health condition coding algorithm Values are from the associated coding set The field is coded with an automatic algorithm. It may consist of errors but the error rate is low.
	DiagnosisCodeType_calc	DiagnosisCodeType_orig recoded into consistent text	Populated by health condition coding algorithm, e.g. 'ICD9'
	DiagnosisCode_calc	DiagnosisCode_orig recoded into consistent text	Populated by health condition coding algorithm Values are from the associated coding set The field is coded with an automatic algorithm. It may consist of errors but the error rate is low.

Table: Lab

Table Notes:
It consists results of lab tests relevant to Index Diseases and extracts only the designated lab tests.

	Field	Definition	Field Notes
	Network_ID	Foreign Key from Network table	Only a Network_ID that exists in the Network table can be referenced here Forms a Composite Foreign Key with Site_ID
	Site_ID	Foreign Key from Site table	Only a Site_ID that exists in the Site table can be referenced here Forms a Composite Foreign Key with Network_ID
	Patient_ID	Foreign Key from Patient table	Only a Patient_ID that exists in the Patient table can be referenced here
	Encounter_ID	Foreign Key from Encounter table	Only an Encounter_ID that exists in the Encounter table can be used May be null if EMR does not have the data
	PerformedDate	Date that the lab test was done	The field can be the date when the lab is conducted or the date when the lab is ordered
	Name_orig	Text exactly as it appears in the EMR	
	CodeType_orig	CodeType text exactly as it appears in the EMR	
	Code_orig	Code text exactly as it appears in the EMR	
	TestResult_orig	Result of the lab test	
	UpperNormal	Highest lab result value that is considered normal	May not be available
	LowerNormal	Lowest lab result value that is considered normal	May not be available
	NormalRange	Original text containing upper and lower lab ranges in one record from the EMR	If upper and lower ranges are already given in separate fields, this field can be left blank
	UnitOfMeasure_orig	Unit of measure for the value in Name_orig field	May not be available
	DateCreated	EMR date stamp of the record	The field indicates when the record is created in the EMR. It may not suitable a s an approximation to PerformedDate
	Name_calc	Name_orig recoded into consistent text	Permitted values enforced by check constraints Populated by lab result coding algorithm The field is coded with an automatic algorithm. It may consist of errors but the error rate is low
	CodeType_calc	CodeType_orig recoded into consistent text	Populated by lab result coding algorithm Values are the name of the coding set (NOT enforced by check constraints) e.g. ‘LOINC’, ‘Life Labs Proprietary’
	Code_calc	Code_orig orig recoded into consistent text	Populated by lab result coding algorithm Values are from the associated coding set The field is coded with an automatic algorithm. It may consist of errors but the error rate is low
	TestResult_calc	TestResult_orig recoded into consistent text	The field is calculated with an automatic algorithm that cleans the value and adjusts the unit of measurement. It may consist errors but has better data quality than TestResult_orig
	UnitOfMeasure_calc	UnitOfMeasure_orig recoded into consistent text	The field is calculated with an automatic algorithm that cleans the value and adjusts the unit of measurement. It may consist errors but has better data quality than UnitOfMeasure_orig

Table: MedicalProcedure

Table Notes:
All procedures performed on the patient.

	Field	Definition	Field Notes
	Network_ID	Foreign Key from Network table	Only a Network_ID that exists in the Network table can be referenced here Forms a Composite Foreign Key with Site_ID
	Site_ID	Foreign Key from Site table	Only a Site_ID that exists in the Site table can be referenced here Forms a Composite Foreign Key with Network_ID
	Patient_ID	Foreign Key from Patient table	Only a Patient_ID that exists in the Patient table can be referenced here
	Encounter_ID	Foreign Key from Encounter table	Only an Encounter_ID that exists in the Encounter table can be used May be null if EMR does not have the data
	PerformedDate	Date that the procedure was performed	The field has a medium-high completion rate
	Name_orig	Procedure text exactly as it appears in the EMR	The field is in free text
	DateCreated	EMR date stamp of the record	The field indicates when the record is created in the EMR. It may not suitable as an approximation to PerformedDate
	Name_calc	Name_orig recoded into consistent text	Re-coding will be done in a future cycle

Table: Medication

Table Notes:
All medications prescribed for the patient.

	Field	Definition	Field Notes
	Network_ID	Foreign Key from Network table	Only a Network_ID that exists in the Network table can be referenced here Forms a Composite Foreign Key with Site_ID
	Site_ID	Foreign Key from Site table	Only a Site_ID that exists in the Site table can be referenced here Forms a Composite Foreign Key with Network_ID
	Patient_ID	Foreign Key from Patient table	Only a Patient_ID that exists in the Patient table can be referenced here
	Encounter_ID	Foreign Key from Encounter table	Only an Encounter_ID that exists in the Encounter table can be used May be null if EMR does not have the data
	StartDate	Date that the Patient started taking the medication	The field is the date when the prescription is prescribed, or the date when the patient is supposed to start the medication. It is an approximation of the date when the patient starts the medication. It is possible that the patient may not fill the prescription or take the medication at all. The field has a high completion rate.
	StopDate	Date that the Patient stopped taking the medication	The field is the date when the patient stops the medication. It has a low-medium completion rate. The null value does not necessarily mean the patient is still on the medication
	Reason	Reason that Patient was prescribed the medication	Exact text as in the EMR, no cleaning for now The field has a low completion rate
	DIN	DIN number for the medication	Extract this where available
	Name_orig	Text exactly as it appears in the EMR	
	CodeType_orig	Original code set used in the EMR	e.g. 'ATC'
	Code_orig	Original code used in the EMR	
	Strength	Concentration of the medication	e.g. 40 The field has a medium completion rate and is in free text. DIN may provide more reliable information
	Dose	Number of units of the medication to be taken	e.g. 2 The field has a medium completion rate and is in free text
	UnitOfMeasure	Units of medication strength	e.g. mg, ml The field has a low-medium completion rate and is in free text. DIN may provide more reliable information
	Frequency	Frequency at which medication to be taken	e.g. bid, tid, q4hr The field has a low-medium completion rate and is in free text
	DurationCount	Length of time that the patient should take the medication	e.g. 10 The field has a low-medium completion rate and is in free text
	DurationUnit	The units of measure for the DurationCount	e.g. 'days' The field has a low-medium completion rate and is in free text
	DispensedCount	Number of units (as defined in DispensedForm) to be dispensed	e.g. 90 The field has a low-medium completion rate and is in free text
	DispensedForm	Form of dispensed medication	e.g. 'Vial', 'Tab', 'Capsule' The field has a low-medium completion rate and is in free text. DIN may provide

			more reliable information
	RefillCount	Number of refills	e.g. 3 The field is in the number format and consists of data entry errors
	DateCreated	EMR date stamp of the record	The field indicates when the record is created in the EMR. It may not suitable as an approximation to StartDate
	Name_calc	Name_orig recoded into consistent text	Populated by medication coding algorithm Values are from the associated coding set, NOT enforced by check constraints The field is coded with an automatic algorithm. It may consist of errors but the error rate is low
	CodeType_calc	CodeType_orig recoded into consistent text	Populated by medication coding algorithm. Values are from the associated coding set, NOT enforced by check constraints e.g. ‘ATC’
	Code_calc	Code_orig recoded into consistent text	Populated by medication coding algorithm Values are from the associated coding set, NOT enforced by check constraints The field is coded with an automatic algorithm. It may consist of errors but the error rate is low

Table: PatientDemographic

Table Notes:
Changeable characteristics/demographics of the patients are stored here.

	Field	Definition	Field Notes
	Network_ID	Foreign Key from Network table	Only a Network_ID that exists in the Network table can be referenced here Forms a Composite Foreign Key with Site_ID
	Site_ID	Foreign Key from Site table	Only a Site_ID that exists in the Site table can be referenced here Forms a Composite Foreign Key with Network_ID
	Patient_ID	Foreign Key from Patient table	Only a Patient_ID that exists in the Patient table can be referenced here
	Occupation	Patient's occupation	Extract this from the EMR if it is available The field has a very low completion rate and is in free text
	HighestEducation	Patient's highest education	Extract this from the EMR if it is available The field has a very low completion rate and is in free text
	HousingStatus	Patient's housing status	Extract this from the EMR if it is available The field has a very low completion rate and is in free text
	ResidencePostalCode	Postal code of the patient	e.g. 'X2X 2X2' The field has a high completion rate and it may be truncated due to privacy concern
	PatientStatus_orig	Status of the patient	e.g: deceased, expired, active, hospital, transient, senior clinic If both codes and description of codes are available(e.g. code=1, description=active), only enter the description The field has a high completion rate and is in free text
	PrimaryLanguage	Patient's primary language	The field has a very low completion rate and is in free text
	Ethnicity	Patient's ethnicity	Extract this from the EMR if it is available The field has a very low completion rate and is in free text
	DeceasedYear	The year in which the patient became deceased	The field is likely to consist of delayed and incomplete value
	DateCreated	EMR date stamp of the record	
	PatientStatus_calc	PatientStatus_orig recoded into consistent text	Permitted values are enforced by check constraints e.g. 'Active', 'Deceased', etc. The field is coded with an automatic algorithm. It may consist of errors but the error rate is low

Table: Referral

Table Notes:
All referrals made for the patient.
Include only referrals made by this provider/practice. Exclude referrals made by specialists to another provider.

	Field	Definition	Field Notes
	Network_ID	Foreign Key from Network table	Only a Network_ID that exists in the Network table can be referenced here Forms a Composite Foreign Key with Site_ID
	Site_ID	Foreign Key from Site table	Only a Site_ID that exists in the Site table can be referenced here Forms a Composite Foreign Key with Network_ID
	Patient_ID	Foreign Key from Patient table	Only a Patient_ID that exists in the Patient table can be referenced here
	Encounter_ID	Foreign Key from Encounter table	Only an Encounter_ID that exists in the Encounter table can be used May be null if EMR does not have the data
	CompletedDate	Date when the patient saw the provider to whom they were referred	Extract if available in your EMR The field has a low-medium completion rate
	Name_orig	Referral Text exactly as it appears in the EMR	We want the reason for referral, not the entire referral letter
	DateCreated	EMR date stamp of the record	
	Name_calc	Name_orig recoded into consistent text	Populated by the Referral Cleaning algorithm The field is coded with an automatic algorithm. It may consist of errors but the error rate is low.
	ConceptCode_calc	SNOMED concept code	Populated by the Referral Cleaning algorithm The field is coded with an automatic algorithm. It may consist of errors but the error rate is low.

Table: RiskFactor

Table Notes:
Risk factors recorded for the patient.
Extract only the designated risk factors.

	Field	Definition	Field Notes
	Network_ID	Foreign Key from Network table	Only a Network_ID that exists in the Network table can be referenced here Forms a Composite Foreign Key with Site_ID
	Site_ID	Foreign Key from Site table	Only a Site_ID that exists in the Site table can be referenced here Forms a Composite Foreign Key with Network_ID
	Patient_ID	Foreign Key from Patient table	Only a Patient_ID that exists in the Patient table can be referenced here
	Encounter_ID	Foreign Key from Encounter table	Only an Encounter_ID that exists in the Encounter table can be used May be null if EMR does not have the data
	StartDate	Date that the risk factor began	The field can be the date when the patient discusses the risk factor with the doctor or the date when the risk factor starts
	EndDate	Date that the risk factor ended	The null value most likely suggests the risk factor remains but not necessary
	Name_orig	Risk factor name exactly as it appears in the EMR	The field is in free text
	Name_calc	Name_orig recoded into consistent text	Permitted values are enforced by check constraints Re-coding is currently performed by the individual Data Managers The field is not necessarily implemented
	Value_orig	Measure of the risk factor	For example, if Risk Factor is 'smoker', value might be '3 pk/week' or '7/day' The field has a medium completion rate and is in free text
	Value_calc	Value_orig recoded into consistent text	The field is not necessarily implemented
	Status_orig	Original value denoting whether the risk is current, past, never, etc.	The field has a medium completion rate and is in free text
	Status_calc	Coded value of the status	Permitted values are enforced by check constraints The field is not necessarily implemented
	Frequency	How often the patient is currently affected by the specified risk factor	For example, if the person currently smokes less than 3 packs per day, value would be “3” The field is not necessarily implemented
	FrequencyType	For entries where a specific value is not provided, allows a comparative description of frequency length	Possible values: Greater than; Less than For example, if the person currently smokes less than 3 packs per day, value would be “Less than” The field is not necessarily implemented
	FrequencyUnit	Frequency Unit of Measure	For example, if the person currently smokes less than 3 packs per day, value would be “Packs per day” The field is not necessarily implemented
	Duration	Amount of time that the person has been affected by the specified risk factor	For example, if the person has been smoking for more than 10 years, value would be “10” The field is not necessarily implemented
	DurationType	For entries where a specific value is not provided, allows for the entry of a comparative description of Duration length	Possible values: Greater than; Less than For example, if the person has been smoking for more than 10 years, value would be “Greater than”

			The field is not necessarily implemented
	DurationUnit	Duration Unit of Measure	For example, if the person has been smoking for more than 10 years, value would be “Years” The field is not necessarily implemented
	EndDuration	Period of time since the person is no longer affected by the specified risk factor	For example, if the person has not had a drink for more than 5 years, value would be “5” The field is not necessarily implemented
	EndDurationType	For entries where a specific value is not provided, allows for the entry of a comparative description of EndDuration length	Possible values: Greater than; Less than For example, if the person has not had a drink for more than 5 years, value would be “Greater than” The field is not necessarily implemented
	EndDurationUnit	EndDuration Unit of Measure	For example, if the person has not had a drink for more than 5 years, value would be “Years” The field is not necessarily implemented
	RiskDetails	Any additional details regarding the risk factor	Possible entries: details regarding cessation attempts and type; relevant details to risk which currently may not fit structured fields, etc. The field is not necessarily implemented
	DateCreated	EMR date stamp of the record	The field indicates when the record is created in the EMR. It may not suitable as an approximation to StartDate

Table: Vaccine

Table Notes:
All vaccinations given to the patient.

	Field	Definition	Field Notes
	Network_ID	Foreign Key from Network table	Only a Network_ID that exists in the Network table can be referenced here Forms a Composite Foreign Key with Site_ID
	Site_ID	Foreign Key from Site table	Only a Site_ID that exists in the Site table can be referenced here Forms a Composite Foreign Key with Network_ID
	Patient_ID	Foreign Key from Patient table	Only a Patient_ID that exists in the Patient table can be referenced here
	Encounter_ID	Foreign Key from Encounter table	Only an Encounter_ID that exists in the Encounter table can be used May be null if EMR does not have the data
	GivenDate	Date of vaccine administration.	
	ExpiryDate	Vaccine expiry date.	The vaccine-batch expiry date, NOT the end-date of vaccine efficacy
	Name_orig	Text exactly as it appears in the EMR	
	CodeType_orig	Original code set used in the EMR	e.g. 'ATC'
	Code_orig	Original code used in the EMR	
	DIN	DIN number for the vaccine	Extract this where available
	Dose	Number of units/volumes of the administered vaccine	e.g. 0.1 The field has a high completion rate and is in free text
	UnitOfMeasure	Units used for the vaccine	e.g. mL The field has a high completion rate and is in free text. DIN may provide more reliable information
	NotGiven	Identifies if vaccination was prevented	e.g. 'Yes'=1, 'No'=0
	NotGivenReason	Represents the reason a vaccine was not administered to a patient	e.g. 'Patient Objection', 'Allergy', 'history of severe reaction' The field has a medium completion rate and is in free text
	Reaction	Adverse reaction related to immunization	e.g. 'Allergic', 'idiosyncratic', 'Intolerance', 'Overdose' The field has a low-medium completion rate and is in free text
	AdminSite	Site of vaccine administration	e.g. 'Left deltoid', 'Right gluteus' The field has a medium-high completion rate and is in free text
	Route	Route of vaccine administration	e.g. 'PO', 'IM', 'SC' The field has a medium-high completion rate and is in free text
	Lot	The vaccine lot number	e.g. 'C2274AA', 'M005060' The field has a medium-high completion rate and is in free text
	DateCreated	EMR date stamp of the record	The field indicates when the record is created in the EMR. It may not suitable as an approximation to GivenDate
	Name_calc	Name_orig recoded into consistent text	Populated by vaccine coding algorithm The field is coded with an automatic algorithm. It may consist of errors but the error rate is low
	CodeType_calc	CodeType_orig recoded into consistent text	Populated by vaccine coding algorithm, e.g. 'ATC'
	Code_calc	Code_orig recoded into consistent text	Populated by vaccine coding algorithm

			The field is coded with an automatic algorithm. It may consist of errors but the error rate is low

Table: DiseaseCase

Table Notes:
Patients in the Patient table who have one or more of the Index Diseases.
Populated by the case detection algorithm.

	Field	Definition	Field Notes
	Patient_ID	Foreign Key from Patient table	Only a Patient_ID that exists in the Patient table can be referenced here
	Disease	The Patient's chronic condition of interest to this database	Valid values are diseases with CPCSSN case definitions e.g. 'COPD', 'Depression', 'Diabetes Mellitus'
	DateOfOnset	Date that the health condition began	The field is the earliest date when the patient shows an indication of a disease. For example, a patient may first have an FBG over 7 and is diagnosed as Diabetes later. The field will use the date of the lab instead of the date of diagnosis

Table: DiseaseCaseIndicator

Table Notes:
Collects all of the reasons that a patient has been identified as having an index disease.
Populated by the case detection algorithm.

	Field	Definition	Field Notes
	Patient_ID	Foreign Key from Patient table	Only a Patient_ID that exists in the Patient table can be referenced here
	Disease	The Patient's chronic condition of interest to this database	Valid values are diseases with CPCSSN case definitions e.g. 'COPD', 'Depression', 'Diabetes Mellitus'
	IndicatorType	General category that the indicator falls under	Values are the names of the table where the original record can be found e.g. "HealthCondition", "Medication", etc.
	IndicatorValue	Data value from the original record	
	TableName	Name of the table storing the original record	
	TableKey	Primary key of the original record in the original table	
	DateCreated	EMR date stamp of the original record	