

Navigating CDW Data

Brian Robison MPH

VINCI Data Services



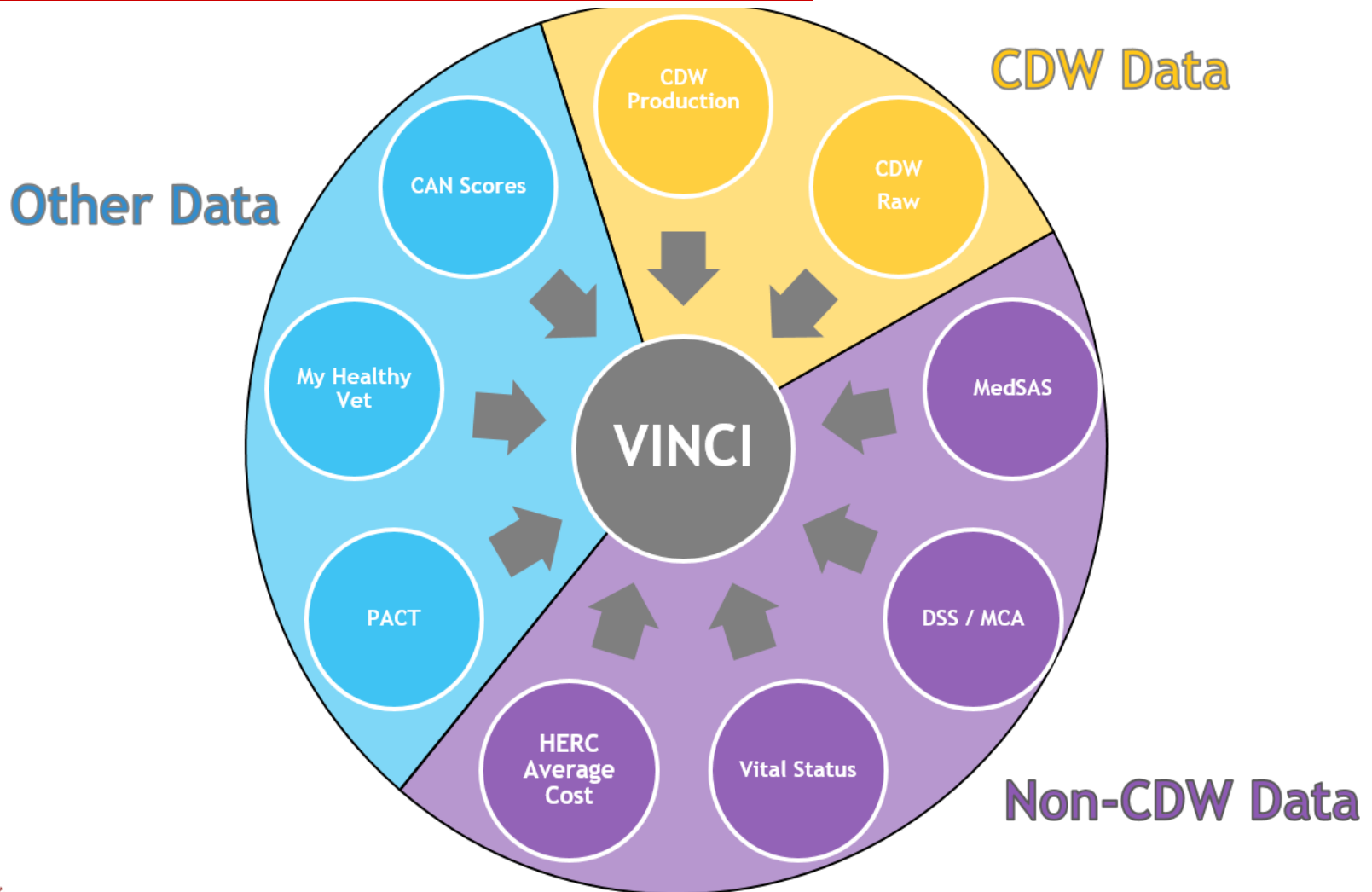
VA Informatics &
Computing
Infrastructure



Agenda

- VA Data Sources
- Finding Research Variables of Interest
- Understanding Patient Identifiers
- SQL Indexing

VA Data Sources

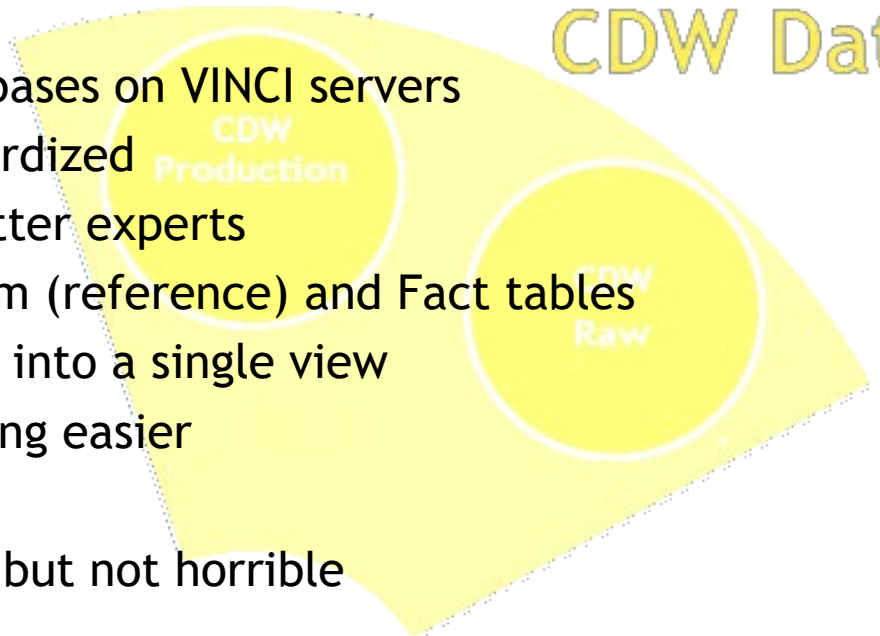


CDW Data

■ Consists of CDW-Production and CDW-Raw

- Data resides in [CDWork] databases on VINCI servers
- Production data has been standardized
- Reviewed by clinical subject matter experts
- Data objects categorized into Dim (reference) and Fact tables
- May combine multiple VistA files into a single view
- SIDs added by CDW - makes joining easier
- Indexes added by CDW
- Documentation exists; not great but not horrible
 - CDW Support Site Metadata Report
 - [Meta] schema in [CDWork]

CDW Data



Dimensions and NDim

- Dimension data does NOT contain any patient data.
 - All dimension tables are accessible to all studies.
 - Located in the CDWork database on all study database servers
- Contains reference/lookup information, including information about:
- ICD diagnosis & procedure code numbers/definitions
 - CPT procedure code numbers/definitions
 - Drugs and drug classes
 - Lab tests and specimen types
 - VistA sites / stations, VISNs, and regions
 - All other descriptive information about terms and items used in VA health care.
 - Used to efficiently identify patients for cohorts and to select data.
- NDim is a schema for any dimension table that is not sourced from a Vista system.

ADR Enrollment

- ❑ Contains data from the Enrollment System Administrative Data Repository
 - ❑ Includes information about Veteran Status, Enrollment status, and Priority Group
- ❑ Fact tables:
 - ADR.ADREnrollHistory
 - Veteran.ADRPerson
 - Veteran.ADRPersonMerge
- ❑ Dim tables:
 - NDim.ADRApplicationCloseReason
 - NDim.ADRDeathNotificationSource
 - NDim.ADREnrollCategory
 - NDim.ADREnrollSource
 - NDim.ADREnrollStatus
 - NDim.ADRPriorityGroup
 - NDim.ADRPrioritySubGroup
 - NDim.ADRReferenceType

Allergy

- This domain contains allergy and adverse drug reaction information generated by the VA.
 - Contains information about hypersensitive reactions, including:
 - Medications
 - Foods
 - Chemicals
 - Noxious stimuli
 - Contains information about patient reactions
- Fact tables:
 - Allergy.AllergicReaction
 - Allergy.AllergicReactionAssessment
 - Allergy.Allergy
 - Allergy.AllergyComment
 - Allergy.AllergyDrugClass
 - Allergy.AllergyDrugIngredient
 - Dim tables:
 - Dim.Reactant
 - Dim.ReactantDrugClass
 - Dim.ReactantDrugIngredient
 - Dim.ReactantSynonym
 - Dim.Reaction

AntiCoag

- Contains information from a combination of VistA File#130 and the ORAM Flowsheet through the Anticoagulation Management Tool
- Not all 130 facilities use this tool so data will be missing from those facilities
- Includes data on International Normalized Ratio, and indication for anticoagulation treatment
- Fact tables: AntiCoag.Flowsheet and AntiCoag.FlowsheetActivity

Appointment

- ❑ Contains information about appointments, including appointment date, purpose, status, reason cancelled, staff member who cancelled.
- ❑ Fact tables:
 - Appt.Appointment
Scheduled and historical appointments
Station
 - Appt.AppointmentMultiple
Sub-appointments for appointments in the Appointment table.
- ❑ Dim tables
 - Dim.AppointmentType
 - Dim.CancellationReason
 - Dim.Eligibility
 - Dim.Location
 - Dim.LocationProvider

BeneTravel

- Contains information on records of travel payments to patients and patient care givers
- Fact tables:
 - BeneTravel.Claim
 - BeneTravel.Certification
- Dim table:
 - Dim.BeneTravelAccount

Consult

- The consult domain contains information about requests for services, including clinical specialty care, procedures, and non- VA care.
- Fact tables:
 - Con.Consult
 - Con.ConsultActivity
- Dim tables:
 - Dim.ActivityType
 - Dim.AssociatedStopCode
 - Dim.ClinicalProcedure
 - Dim.GMRCProcedure

COVID-19 Shared Data Resource

- Contains information related to COVID-19 both in and outside of VA
- Collaborative effort utilizing multiple data sources
- Includes case definitions, concept definitions, VA data mappings, and other information
- 33 Fact tables and 7 Dim tables

CPRS Orders

- ❑ Contains Orders and Requisitions made through the Order Entry Option (OR) using any CPRS package.
 - Labs, Radiology (x-ray, CT, MRI, DEXA scan), Medications, etc.
- ❑ Fact tables:
 - CPRSOrder.CPRSOrder
 - CPRSOrder.OrderAction
 - CPRSOrder.OrderCheckInstance
 - CPRSOrder.OrderCheckMessage
 - CPRSOrder.OrderCheckOverrideReason
 - CPRSOrder.OrderedItem
 - CPRSOrder.OrderIntervention

Dental

- Dental Data
 - Contains records of dental exams, diagnoses, history, transactions, and procedures, including:
 - Basic hygiene
 - Restorative surgery
 - Orthodontics
 - Periodontics
 - Endodontics
 - 29 Fact tables and 17 Dim tables

Emergency Dept. Int. Software (EDIS)

- Data of patients as they were tracked through the Emergency Department through the Emergency Department Integration Software
 - Includes Patient Arrivals, Dispositions, Departures, and Reasons for visit
- Fact tables:
 - EDIS.EDISLog
 - EDIS.EDISLogDischargeDiagnosis
 - EDIS.EDISLogHistory
 - EDIS.EDISOrder
- Dim tables
 - Dim.EDISRole
 - Dim.EDISRoomBed
 - Dim.EDISStaff
 - Dim.EDISTrackingArea
 - Dim.EDISTrackingCode

Health Factors

- Contains information about over 3,000 different types of factors which affect the health of specific patients. Common health factors include:
 - Tobacco use
 - Alcohol use
 - Drug use
- Usage of health factors varies substantially by station.
- Health factors are often used with clinical reminders.
- Health Factor Data uses mostly free text
- Fact table:
 - HF.HealthFactor
- Dim table:
 - Dim.HealthFactorType

Immunizations

- Contains records of individual patient immunizations
- There are over 52 million immunization records
- Records can contain links to outpatient visit records via the VisitSID
- Fact table:
 - Immun.Immunization
- Dim tables:
 - Dim.ImmunizationAdministrationRoute
 - Dim.ImmunizationAdministrationSite
 - Dim.ImmunizationInformationSource
 - Dim.ImmunizationName

Inpatient

- ❑ Generated from Inpatient VistA : Patient Treatment File(PTF), 45 which contains all PTF information generated from admissions, treating specialty transfers, and PTF screen edits.
- ❑ Includes records of inpatient:
 - Admissions
 - Discharges
 - Surgical Procedures
 - ICD/CPT Procedures
 - Primary, secondary...etc. diagnoses for stay
 - Ward location
- ❑ Primary domain tables
 - Inpat.Inpatient
 - Inpat.InpatientDiagnosis
 - Inpat.InpatientCPTProcedure
 - Inpat.InpatientICDProcedure

Integrated Billing

- Contains data from bills generated on behalf of patients
- Generated from VistA files: Bill/Claims (399) and Accounts Receivable (430)
- 36 Fact tables and 31 Dim tables

Lab Accession

- Contains information on Lab Pathology Data from VistA File #68
- Fact tables:
 - LabAccession.AccessionAreaNumber
 - LabAccession.AccessionAreaPeriod
 - LabAccession.AccessionTest
 - LabAccession.AccessionTestWorkload

Lab Chemistry

- Use of the PatientLabChem table is CDW's recommended path to access the LabChem data for most patient analysis projects. This restricts the laboratory results to only those records that refer to human patients and filters out the administrative test, referral test, and research test records.
- Primary Domain tables:
 - Chem_PatientLabChem
 - Chem_LabPanel
 - Chem_OrderedLabPanel
- Primary Dim tables:
 - Dim.LabChemTest
 - Dim.LOINC
 - Dim.Topography

Mental Health Assessment

- Contains the data collected during the administration of a specified instrument from the MH TESTS AND SURVEYS file (#601.71) given to a patient at a specific date and time
- NOTE: Some of the Survey text fields include hidden characters (tabs, carriage returns, etc.)
 - Can use the REPLACE function along with ascii codes in SQL to identify and remove them
- Fact tables:
 - MH.SurveyAdministration
 - MH.SurveyAnswer
 - MH.SurveyAnswerText
 - MH.SurveyResult
- Dim tables:
 - Dim.Survey
 - Dim.SurveyChoice
 - Dim.SurveyChoiceType
 - Dim.SurveyContent
 - Dim.SurveyQuestion
 - Dim.SurveyScale
 - Dim.SurveyScaleGroup
 - Dim.SurveyScoringKey

Microbiology

- Contains data of report-level and organism/antibiotic susceptibility-level records
- Includes etiologies involved in pathogenic Microbiology
- Much of the data is text-based
- 15 Fact tables and 2 Dim tables

Master Veteran Index (MVI)

- A single master index of all Veterans Health Administration patients
- Acts to maintain and synchronize identities and demographic data for VA clients, Veterans, and beneficiaries
- Includes PatientICN status to assist with identifying deactivated PatientICN's
- 4 Fact tables and 11 Dim tables

OMOP

International Common Data Model

- Standard Vocabularies
- Standardized Data allowing data interchange
 - Condition
 - Procedure
 - Device Exposure
 - Drug - Exposure, Dose, Era
 - Measurement
 - Observation
- OMOP CDM defines table structures for each of the data domains in a Person and Provider-centric model

Outpatient

The Outpatient Visits domain replaced the CDW Encounter domain. Unlike the Enc tables, no records are filtered out of Output (aside from the normal excluding records before 10/1/1999).

- Contains data from the VistA ENCOUNTER (409.68), VISIT (9000010) and related files.
- Primary domain tables
 - Output_Visit
 - Output_VDiagnosis
 - Output_VProcedure
 - Output_ProblemList

Pathology

- Provides data from several pathology VistA files.
- Includes information on autopsies, cytopathology, and surgical pathology
- 69 Fact tables and 2 Dim tables

Patient Enrollment

- The data in the Patient Enrollment domain originates in the VistA version of the VA Enrollment data. When a Veteran goes to a VA Medical Center (VAMC) to enroll in VA care, information pertaining to his or her enrollment is stored in the local VistA system and is then sent to the Health Eligibility Center (HEC) for verification. HEC's enrollment data are considered the authoritative source. Therefore, this VistA enrollment data are synchronized with HEC data once a determination is made. A CDW field for SourceOfEnrollment indicates whether or not a particular record has been synchronized.
- Contains information such as: enrollment status, dates of enrollment periods, severity of disability, service-connected disability, military discharge status, type and number of benefits being paid, priority group, etc.
- Fact table: Patient.Enrollment

Pharmacy

Bar Code Medication Administration (BCMA)

- Medication administration process for the inpatient setting
 - Dates and times medication was ordered, delivered, and administered; Details of the medication; Variances or deviations of the administration process
- Primary domain tables
 - BCMAMedicationLog (Parent table)
 - BCMADispensedDrug
 - BCMAAdditive
 - BCMASolution

Pharmacy Outpatient

- All outpatient pharmacy data: administrative pharmacy services and info on patients receiving prescriptions
 - Fill dates, process dates, number of days' supply, refills, renewals
 - Includes specific drug information:
 - Drug Name
 - VA Drug Class
- Primary domain tables
 - RxOutpat
 - RxOutpatFill
 - RxOutpatMedInstructions
 - RxOutpatSig

Program Integrity Tool (PIT)

- Data on VA claims that are processed through the PIT including, but not limited to, Choice Claims
- 40 Fact tables
- 27 Dim tables

Purchased Care Authorized (Fee)

- Data on reimbursement of non-VA medical care from an external entity
 - Vendor, reason for visit, dates of service, status of the reimbursement; Amount billed, amount paid, type of payment
 - Patient information such as initial diagnosis, insurance, admission/discharge dates and dispositions
- 12 Fact tables and 9 Dim tables

Prosthetics

- Contains data on prosthetics and prosthetic purchasing transactions
- Generated from VistA file 660
- 27 Fact tables and 22 Dim tables

Radiology

- Administrative and Clinical information for radiology exams, orders, reports
- Schemas:
 - Rad: fact tables
 - SPatientText: report text, notes, written clinical histories
- 20 Fact tables and 21 Dim tables

Rx IV

- Contains information on VA IV activity, renewals, label tracking, and comments from providers
- Generated from VistA file 55
- 7 Fact tables and 2 Dim tables

SPatient & Patient

- All patients followed by medical centers and/or outpatient clinics
 - Demographics such as age, race, ethnicity
 - Sensitive patient data such as names, addresses, Real SSN
 - Military service information such as service-connected disabilities, period of service
- Primary domain tables
 - Patient and SPatient
 - SPatientAddress and SPatientPhone
 - MilitaryServiceEpisode

SStaff & Staff

- Information on staff affiliated with the VA
 - SStaff: Secured Staff PII data including staff SSN
 - Staff: Basic non-PII staff information
- Primary domain tables
 - Staff and SStaff
 - StaffSub.ProviderTypeAssignment
 - StaffSub.VistaPermission

Surgery PRE, INTRA and POST

- Data extracted from the Surgery package
- Split vertically into three conceptual groupings:
 - Surgery_PRE- surgical case scheduling and preparation
 - Surgery_INTRA - data during a surgical case progression
 - Surgery_POST - post-case data, provider notes, indications for operations, surgically implanted prosthetics
- Each view has surgery case identifier, patient identifier, date & time of the surgery, location, etc.
- Primary domain tables
 - SurgeryPRE, SurgeryINTRA, SurgeryPOST
 - Tables for associated Diagnosis, Procedure, CPT codes

Text Integration Utility (TIU)

- ❑ TIU contains free text documents such as progress notes and consult reports. TIU notes on VINCI are full text indexed
- ❑ Utilize the table value function tvf_TIU_FullTextSearch when searching the full text for keyword(s) for best results
- ❑ This data is extracted from the VistA file: TIU Document (8925)
- ❑ Domain tables in CDWork:
 - TIUDocument
 - TIUDocument_Recent
 - TIUIncompleteRecord

Vital Signs

- ❑ Contains Vital sign data from all sites. Extracted from VistA file: GMRV Vital Measurement 120.5
- ❑ There is no unit of measure with each type of vital. However, it is assumed that Height is measured in inches, weight in pounds, respiration in breaths per minute, pain from 1 to 10, pulse in beats per minute, temperature in degrees Fahrenheit, etc. or vital qualifiers which record things like the arm from which the blood pressure was taken.
- ❑ Domain table: Vital.VitalSign
- ❑ Dim table: Dim.VitalType

Women's Health

- The Women's Health domain provides information from the Women's Health VistA files in the 790 range. The information in the domain contains administrative and clinical information pertaining to care for female veterans. Domain Tables:
 - WH.WHLabTest
 - WH.WHNotification
 - WH.WHPapRegimenLog
 - WH.PregnancyLog
 - WH.WHRefusal
 - Patient.WHPatient

CDW-Raw Data

- Minimal processing: the Raw team's goal is to simply get it out of VistA and into SQL
- Swim at your own risk - they're "Raw" for a reason
- Must use [Sta3n]+[IEN] for joins (with some exceptions)
- Not indexed (with some exceptions)
- Documented... but you may be sorry you asked

CDW Data

Production

CDW
Raw

Echocardiogram

- ❑ Echo domain contains echocardiogram procedures done on a patient.
- ❑ This does not have Echo Reports. However, some Echocardiogram results are located in the TIU notes
- ❑ The [Echo] domain is a limited extract of the following VistA files (among others):

691	ECHO	691.09	RISK FACTOR
691.01	MED (CARDIOLOGY	691.15	DIAGNOSIS
691.02	PRE-DX	691.19	ICD DIAGNOSIS
691.04	REGIONAL	691.5	ELECTROCARDIOGRAM (EKG)
691.05	DOPPLER	691.6	HOLTER
691.06	FINDINGS	692.02	EKG
691.08	SYMPTOM	693	MEDICAL DESCRIPTION

Oncology

- [Oncology] is a CDW-Raw domain and is a limited extract of the Vista Oncology module.
- The Oncology module automates the tumor registry and supports tumor registrars in abstracting cancer cases, following up on cancer patients and producing the Hospital Annual Report.

Non-CDW Data

- Highly refined data sets based upon VistA and other data sources
 - Usually very clean
 - Tend to be de-normalized - “flat” not relational
 - Main patient identifier is typically [ScrSSN] ... sometimes [PatientICN]
- Typically predate the concept of CDW
- Intended for very specific business needs, i.e., very focused
- Tend to have *some* demographic information but usually not complete
- Reasonably well documented

Non-CDW Data

PSSG

- ❑ The PSSG (Planning Systems Support Group) Geocoded Enrollee files are a group of quarterly datasets comprising detailed and extensive geographic information. These data contain one record per enrollee.
- ❑ VINCI receives SAS files and the data is loaded into SQL server for research access
- ❑ Includes data on patient Addresses including drive times

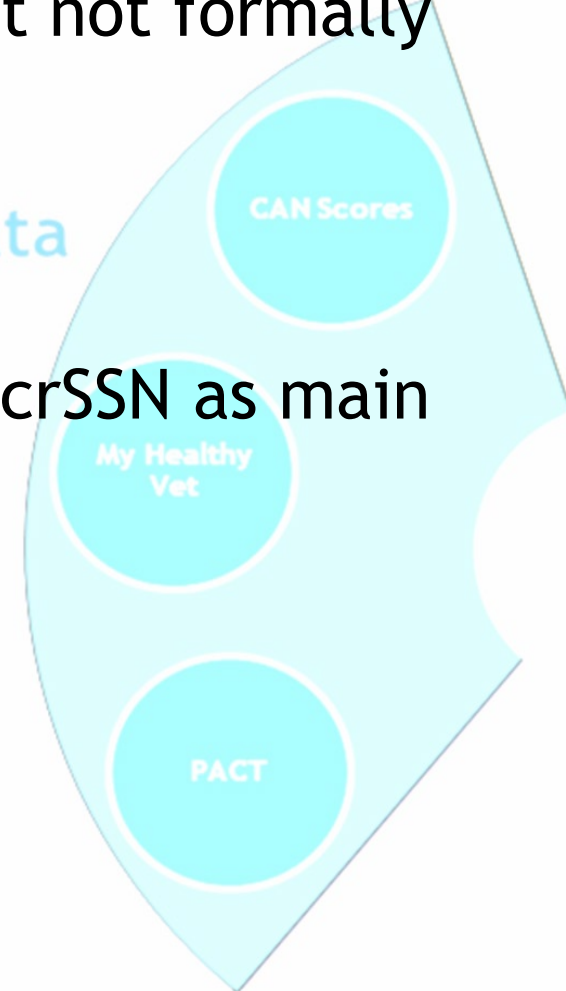
VASQIP

- VA Surgical Quality Improvement Program (VASQIP)
- Surgical Outcome Data
 - Morbidity and Mortality
- Data on the quality of processes and structures of surgical care

Other Data

- Data sets based upon CDW data but not formally produced or managed by CDW
- Documentation is hit and miss
- Frequently contain PatientICN or ScrSSN as main patient identifier
- Indexing is hit and miss

Other Data



Care Assessment Need (CAN) Score

- The Care Assessment Need (CAN) score reflects the estimated probability of hospital admission or death within a specified time frame of 90 days or 1 year.
- The score is expressed as a percentile, ranging from 0 (lowest risk) to 99 (highest risk) compared to other VA patients
- Data used to generate the score is obtained from VHA administrative databases
- Very accurate for large groups of patients



DaVINCI

- Department of Defense and VA-VINCI data integration
- Interagency collaboration and resource sharing
- Military Health System Data and service-related information
- Includes:
 - Health care events
 - Historical beneficiary data
 - Service-related information
 - Demographics



Left Ventricular Ejection Fraction (LVEF)

- This domain contains information on the left ventricular ejection fraction measurement extracted using Natural Language Processing(NLP)
 - The total amount of blood in the left ventricle that is pumped out with each heartbeat.
- The measurement is given in a range value. The LVEF data set separates the lower and upper values of each range that a patient received.
- Data available from VINCI upon request

Patient Aligned Care Team (PACT)

- VHA's implementation of the patient-centered medical home model, aims to improve health outcomes through team-based care, improved access and care management.
- Composite scores were constructed for eight core domains of PACT: access, continuity, care coordination, comprehensiveness, self-management support, patient-centered care and communication, shared decision making and team-based care.
- The main outcomes and measures are patient satisfaction, rates of hospitalization and emergency department use, quality of care and staff burnout.

Finding Research Variables of Interest



VA Informatics &
Computing
Infrastructure



Commonly used data sources

- CDW Production
- CDW Raw
- Vital Status
- OMOP

CDW Production documentation sources:

- VINCI Central data sources page
- Meta data report (via CDW share point site)
- Meta data viewer application (accessed via VINCI central applications)
- Dim Data Viewer application (accessed via VINCI central applications)
- VIREC Documentation and Factbooks
- Meta view (via CDWork database on production servers)
- VHA Data Portal

BISL Meta data Report

CDW Metadata

Contains a grouped list of available CDW ER Diagrams and members.
Click ERDiagramTitle to show associated views, URL link to see diagram.

Right click View Name to drill through.

ERDiagramTitle	ERDiagramDate	URL
Pyramid GIS 1.0	06 Aug 2015 00:00:00	↗
Radiology 1.0 Dimensions	31 Oct 2017 08:45:01	↗
Radiology 1.0 Fact	31 Oct 2017 08:45:52	↗
Radiology 1.0 Fact 2	31 Oct 2017 08:46:52	↗
Radiology 1.0 Patient	31 Oct 2017 08:46:31	↗
Recall Reminders 1.0	23 Feb 2016 00:00:00	↗
RPCMM Dimensions	01 Oct 2018 07:37:49	↗
RPCMM Fact	17 May 2021 15:57:48	↗
RPCMM Hidden	29 May 2018 10:21:16	↗
RxIV (Dimension)	09 Dec 2020 08:57:53	↗
RxIV (Fact)	09 Dec 2020 08:58:04	↗
RxUD (Fact)	09 Dec 2020 07:22:53	↗
SHRMVI - Authorized Access Only	11 Oct 2018 11:13:01	↗
Staff Release 1.1	11 Oct 2018 11:46:42	↗
Surgery Supplemental 1.0	10 Sep 2021 14:35:06	↗
Surgery Supplemental 1.0	29 Sep 2021 11:32:36	↗
SurgeryINTRA 1.1	26 Sep 2019 09:19:08	↗
SurgeryPOST 1.1	15 Dec 2016 00:00:00	↗
SurgeryPRE 1.1	30 Sep 2019 07:46:14	↗
Text Document Titles 1.0	10 Aug 2018 07:34:54	↗
Text Integration (TIU) Templates	06 Oct 2021 07:22:19	↗
Vista Compensation and Pension 1.6	22 Dec 2020 10:58:44	↗
Vista Compensation and Pension 3.2	03 Aug 2021 07:26:50	↗
Vista Data Dictionary 1.0	20 Mar 2015 00:00:00	↗
Vista Permissions 1.0	11 Aug 2014 00:00:00	↗
Vista WaitList 1.0	17 Jul 2014 00:00:00	↗
Vital Sign 1.1	26 Oct 2017 08:13:01	↗
Womens Health	15 Jun 2021 13:47:05	↗
Women's Health v1.0	16 Jun 2016 00:00:00	↗

View Name	Field...	FileMan File Data Source	Deployed View Version	Partition Key	Cutoff Field
Dim.RadiologyDiagnosticCode	9	DIAGNOSTIC CODES (78.3)	xDWWork: 2		
Dim.WHAgeRangeDefault	6	WV AGE RANGE DEFAULT (790.72)	xDWWork: 2		
Dim.WHBreastTreatmentNeed	4	WV BREAST TX NEED (790.51)	xDWWork: 2		
Dim.WHCaseManager	7	WV CASE MANAGER (790.01)	xDWWork: 3		
Dim.WHCaseParameter	79	WV SITE PARAMETER (790.02)	xDWWork: 3		
Dim.WHCervicalTreatmentNeed	4	WV CERVICAL TX NEED (790.5)	xDWWork: 2		
Dim.WHNotificationOutcome	5	WV NOTIFICATION OUTCOME (790.405)	xDWWork: 2		
Dim.WHNotificationPurpose	13	WV NOTIFICATION PURPOSE (790.404)	xDWWork: 2		
Dim.WHNotificationType	6	WV NOTIFICATION TYPE (790.403)	xDWWork: 4		
Dim.WHPapRegimen	5	WV PAP REGIMEN (790.03)	xDWWork: 2		
Dim.WHProcedureType	15	WV PROCEDURE TYPE (790.2)	xDWWork: 4		
Dim.WHRadiologyDiagnosisCode	5	WV DIAGNOSTIC CODE TRANSLATION (7...	xDWWork: 2		
Dim.WHReferralSource	5	WV REFERRAL SOURCE (790.07)	xDWWork: 2		
Dim.WHResult	13	WV RESULTS/DIAGNOSIS (790.31)	xDWWork: 7		
Dim.WHSnapshotReport	53	WV SNAPSHOT REPORTS (790.71)	xDWWork: 4		
Patient.WHPatient	33	WV PATIENT (790)	xDWWork: 8		
WH.WHPapRegimenLog	8	WV PAP REGIMEN LOG (790.04)	xDWWork: 2	BeginDateTime	BeginDateTime
WH.WHLactationStatus	32	LACTATION STATUSES (790.16)	ViewValidationReady: 1	EnteredDateTime	EnteredDateTime
WH.WHLactationStatusErrorReason	20	REASONS ENTERED IN ERROR (790.18)	ViewValidationReady: 1	EnteredDateTime	EnteredDateTime
WH.WHPregnancyLog	41	PREGNANCY STATUSES (790.05)	ViewValidationReady: 8	EnteredDateTime	EnteredDateTime
WH.WHPregnancyLogContraception	18	CONTRACEPTIVE METHODS USED (790.17)	ViewValidationReady: 1	EnteredDateTime	EnteredDateTime
WH.WHPregnancyLogErrorReason	17	REASONS ENTERED IN ERROR (790.15)	ViewValidationReady: 1	EnteredDateTime	EnteredDateTime
WH.WHNotification	40	WV NOTIFICATION (790.4)	ViewValidationReady: 8	NotificationOpenedDat...	NotificationOper
WH.WHLabTest	19	WV LAB TESTS (790.08)	xDWWork: 6	ProcedureDateTime	ProcedureDateTi
WH.WHProcedure	49	WV PROCEDURE (790.1)	xDWWork: 7	ProcedureDateTime	ProcedureDateTi
WH.WHProcedureAction	24	PROCEDURE ACTIONS (790.23)	ViewValidationReady: 1	ProcedureDateTime	ProcedureDateTi
WH.WHRefusal	10	WV REFUSALS (790.3)	xDWWork: 5	RefusalDateTime	RefusalDateTime



VINCI Meta data viewer

Meta Data Viewer

Conduct a Search to Begin



Matching Schemas

- Dental (CDW Production)
- Dim (CDW Production)
- DSS (DSS)
- Fee (CDW Production)
- Inpat (CDW Production)**
- MedSAS (MedSAS)
- Outpat (CDW Production)

Inpat Tables/Views

InpatientCPTProcedure #46 INPATIENT CPT CODE

InpatientCPTProcedure

No Description Available

InpatientCPTProcedure Fields

Showing 1 to 10 of 13 entries

Filter:

Name	Data Type	PK	FK	Field Codes	ERDiagrams	VISTA File	VISTA Field	VISTA Field #
<u>CPTProcedureDateTime</u>	datetime2(6)	N	N		Inpatient 3.0; Inpatient 3.0	INPATIENT CPT CODE	CPT RECORD DATE/TIME	.14
<u>CPTProcedureDateTimeTransformSID</u>	bigint(8)	N	N		Inpatient 3.0; Inpatient 3.0	INPATIENT CPT CODE	CPT RECORD DATE/TIME	.14
<u>CPTProcedureVistaErrorDate</u>	varchar(50)	N	N		Inpatient 3.0; Inpatient 3.0	INPATIENT CPT CODE	CPT RECORD DATE/TIME	.14
<u>AdmitDateTime</u>	datetime2(6)	N	N	DGPT(Inpatient 3.0; Inpatient 3.0	INPATIENT CPT CODE	PTF	.16
<u>AdmitDateTimeTransformSID</u>	bigint(8)	N	N	DGPT(Inpatient 3.0; Inpatient 3.0	INPATIENT CPT CODE	PTF	.16
<u>AdmitVistaErrorDate</u>	varchar(50)	N	N	DGPT(Inpatient 3.0; Inpatient 3.0	INPATIENT CPT CODE	PTF	.16
<u>DeleteDateTime</u>	datetime2(6)	N	N		Inpatient 3.0; Inpatient 3.0	INPATIENT CPT CODE	DELETE DATE	1



VINCI Dim data viewer

Dim Data Viewer

User Guide

Welcome to the Dim Data Viewer. This application allows you to browse/search most of the Dim Tables within the CDW database for the purpose of constructing research cohorts. Please see the [User Guide](#) for additional information on how to use this application.

Smoking



Select a Category

View All Dims
Administrative (8219)
Environment (1997)
Event (4)
Finding (967)
Organism (6)
Procedure (2840)

Select a Finding Dim

ClinicalTerm (285)
DentalCannedComments (12)
HealthFactorType (666)
ICDDescription (1)
Reactant (1)
SurgeryCancelReason (2)

HealthFactorType Description

This file contains a descriptive name of health factors, and their health factor category.

Select HealthFactorType Search Terms

The following are the distinct search terms contained in your selected Dim Table. Select the terms relevant to your search to view the associated Dim Rows.

- 02 NONCOMPLIANCE SMOKING
- 02 NONCOMPLIANCE SMOKING SUPPLIES
- 1 MONTH SMOKING NO
- 1 MONTH SMOKING YES
- 1 YEAR SMOKING NO
- 1 YEAR SMOKING YES
- 2 MONTHS SMOKING NO
- 2 MONTHS SMOKING YES
- 2 WK SMOKING NO
- 2 WK SMOKING YES
- 3 MONTHS SMOKING NO
- 3 MONTHS SMOKING YES
- 520 HISTORY OF SMOKING
- 6 MONTHS SMOKING NO
- 6 MONTHS SMOKING YES
- 688 ID SMOKING CESSATION SESSION 1
- 688 ID SMOKING CESSATION SESSION 2
- 688 ID SMOKING CESSATION SESSION 3
- 688 ID SMOKING CESSATION SESSION 4
- 688 ID SMOKING CESSATION SESSION 5

First Previous 1 2 3 4 5 6 7 Next Last

View Dim Rows



VIReC Documentation

CDW Documentation

Overview

VIReC provides documentation on select Corporate Data Warehouse (CDW) production domains. CDW documentation is designed to help new and seasoned users better understand the structure and contents of the CDW. It might be necessary to review a combination of resources to fully understand the complexities of the CDW and data use environment.

Sign-up for VIReC product news and updates. Email the VIReC [HelpDesk](#) to receive notification of VIReC's new CDW products and product updates.

How are we doing? Take our [short survey](#) to share your thoughts about the CDW resources below. We value your feedback.

Data Documentation

Expand each type of documentation below to view these resources.

+ Expand All

- + Getting Started with CDW
- + Factbooks
- + The Researcher's Notebook
- + Statistical Snapshot
- + Domain Layout & Descriptions
- + Data Contents, Discrete Frequencies, Record & Null Counts

Learn About CDW

- [Overview](#)
- [Data Transition to CDW](#)
- [Documentation](#)
- [FAQs](#)

General Resources

- + [Data Access](#)
- + [Data Sources](#)
- + [Data Tools](#)
- + [Data Topics](#)
- + [Products & Services](#)
- + [Special Projects](#)



Meta View

- On VINCI production servers VHACDWxxx, Under CDWork database, Meta views

The screenshot displays the SQL Server Enterprise Manager interface. On the left, the Object Explorer shows the database structure for 'CDWork', with 'Meta.DWViewField' selected. The main window shows a SQL query and its results.

```
select top 100 * from [CDWork].Meta.[DWViewField]
where [DWViewFieldName] like '%ICD9%'
```

The results are displayed in a table with the following columns: DWViewSchema, DWViewName, ViewVersion, and DWViewFieldName.

	DWViewSchema	DWViewName	ViewVersion	DWViewFieldName
1	Inpat	SpecialtyTransferDiagnosis	6	ICD9SID
2	Outpat	VProcedureDiagnosis	15	ICD9SID
3	Inpat	InpatientSurgicalProcedure	11	ICD9ProcedureSID
4	Fee	FeeServiceProvided	18	ICD9SID
5	Inpat	CensusSurgicalProcedure	3	ICD9ProcedureSID
6	Outpat	VDiagnosis	24	ICD9SID
7	Inpat	Inpatient	37	PrincipalDiagnosisICD9SID
8	Inpat	Inpatient501TransactionDiagnosis	2	ICD9SID

CDW Raw documentation source

! This section is read-only.

CDW-Raw Domain Documents

- IV Meds
- Audiology
- Data Std
- PCM
- HR Smart
- Immunizations
- MCCF
- MDR
- MHV
- Oncology
- PAID
- PCMM
- PGD
- PFT
- Prosthetics
- Reminders

VistA Files & Fields (Oncology)

Tuesday, December 10, 2019 6:06 AM

This section contains the specification for VistA Oncology, from FileManFields.

4/5/2021:

The attached file is only 1MB but it's slow to open in Word. It's got all of the FileManFields definitions for the Oncology tables in CDW.

filenumber in ('160','160.075','160.19','165.51','165.52','164.0111','165.5','165.5','165.5','164.18')

Report - VistA Files (04/05/21 12:32:44)

File	File Name	Parent File	Description
160	ONCOLOGY PATIENT		Demographic and followup data concerning Oncology patients is stored in this file. (Tumor-related data is stored in the Primary File). Data is NOT exported with this file, but it is populated on site.
160.075	SUSPENSE	160	This multiple field stores the SUSPENSE data for each division where the patient is on SUSPENSE.
160.19	FACILITY		The American College of Surgeons identification number. This file contains 5000+ medical locations nationwide.
164.0111	5TH ED-SURGERY PRIMARY SITE	164	
164.18	CHEMOTHERAPEUTIC DRUGS		Chemotherapeutic drugs as listed in SEER SELF-INSTRUCTIONAL MANUAL FOR TUMOR REGISTRARS Book 8 - ANTINEOPLASTIC DRUGS Third Edition
165.5	ONCOLOGY PRIMARY		Tumor-related data for Oncology patients is stored in this file. (Demographic and follow-up data is in the Oncology Patient File). File is populated in the field by using the abstracting options.
165.51	SUBSEQUENT COURSE OF TREATMENT	165.5	Enter subsequent therapy (therapy provided after completion of the first course of therapy).
165.52	RADIATION TREATMENT	165.5	Record the type of radiation therapy.

8 rows selected (from CDW-Prod in 16 milliseconds)



Vital status documentation source

- <http://vaww.vhadataportal.med.va.gov/DataSources/VitalStatusFile.aspx>

Quick View

The Vital Status File contains dates of death and demographic data for Veterans.

Vital Status File

- Overview
- Content
- Features
- Access
- Resources

Vital Status File

Overview

The VHA Vital Status File (VSF) is a mortality file that contains data from multiple VA and non-VA data sources including the BIRLS Death file, Medicare Vital Status File, and the Social Security Administration (SSA) Death Master File. It comprises two cumulative files: the Master and Mini Files. VIREC's VA-NDI Data Merge Project was the impetus for the development of the VHA VSF. The project evaluated the completeness and accuracy of vital status information in VA data sources to that in the National Death Index (NDI). Visit the ["Accuracy and Completeness of Death Dates"](#) page for an overview of findings from this project.

Content

Data in the VSF includes dates of death and demographic data for Veterans who received VHA care since 1992, are VHA enrolled, or received VBA compensation or pension benefits since 2002. The VHA Vital Status File data is updated quarterly. The table below provides names, descriptions, and the availability of each VHA VSF.

Name	Description
Master	Contains one record for each SSN-date of birth (DOB)-gender combination found in VA source data. Data includes dates of birth, gender, dates of death, dates of last activity, Veteran status flags, flags indicating whether SSN was included in the yearly finder file used to request Medicare data, and variables to indicate whether or not the person was enrolled in Medicare.
Mini	Contains one record for each SSN. An algorithm developed by VHA National Data Systems (NDS) is used to select the "best" DOB, gender, and Date of Death (DOD) for each SSN. Non-Veteran SSNs are excluded from the Mini file.

OMOP

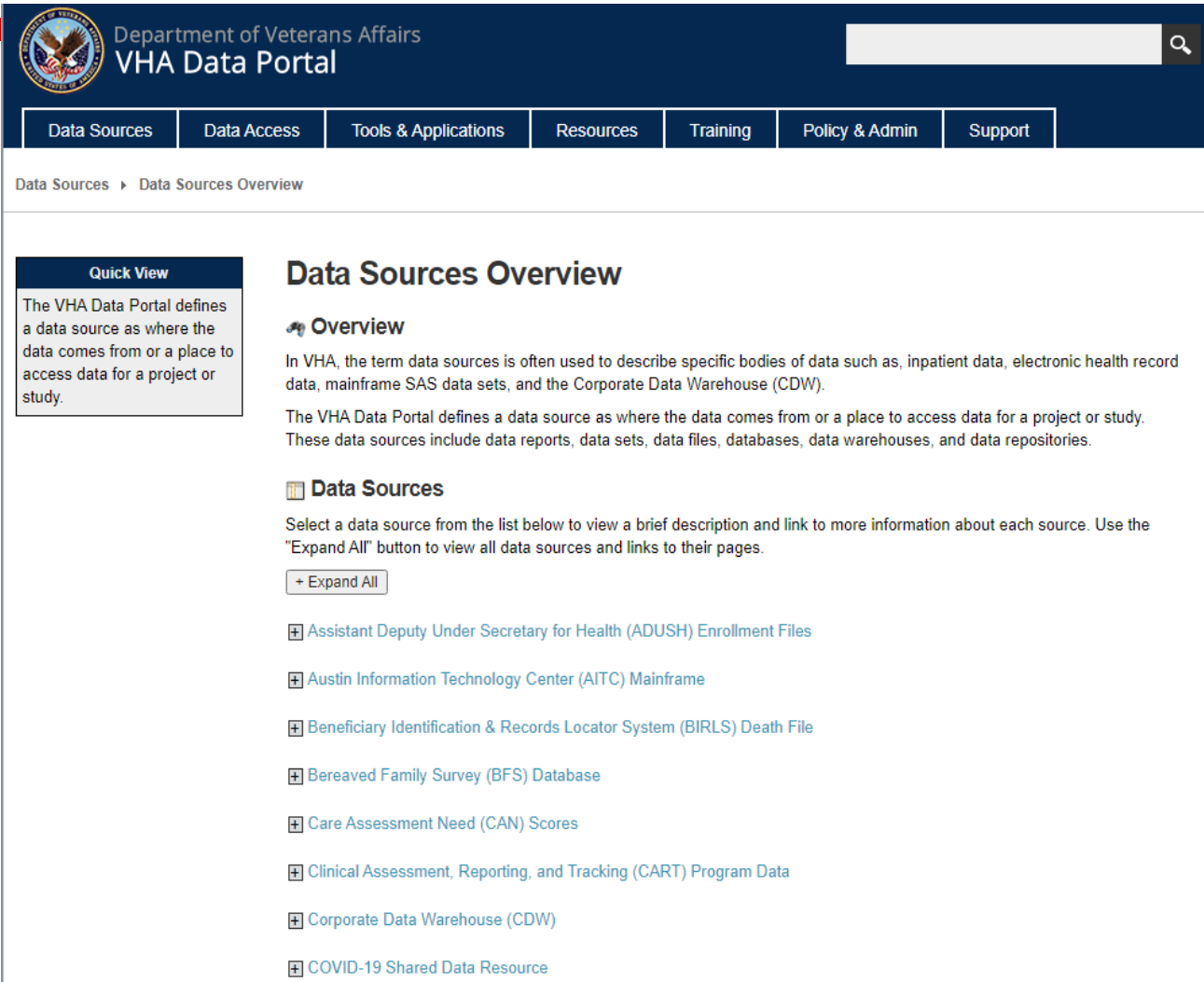
■ OMOP Concepts using the OMOP CONCEPT Table

■ Query Result from OMOP CONCEPT Table:

```
WHERE CONCEPT_NAME LIKE '%Diabetes%' AND DOMAIN_ID = 'CONDITION'
```

CONCEPT_ID	CONCEPT_NAME	DOMAIN_ID	VOCABULARY_ID	CONCEPT_CLASS_ID	STANDARD_CONCEPT	CONCEPT_CODE
45542732	Drug or chemical induced diabetes mellitus with diabetic cataract	Condition	ICD10CM	5-char billing code		E09.36
45542733	Drug or chemical induced diabetes mellitus with neurological complications with diabetic polyneuropathy	Condition	ICD10CM	5-char billing code		E09.42
45542737	Type 1 diabetes mellitus with diabetic peripheral angiopathy with gangrene	Condition	ICD10CM	5-char billing code		E10.52
45542738	Type 2 diabetes mellitus with hyperosmolarity without nonketotic hyperglycemic-hyperosmolar coma (NKHHC)	Condition	ICD10CM	5-char billing code		E11.00
45542741	Other specified diabetes mellitus with diabetic nephropathy	Condition	ICD10CM	5-char billing code		E13.21
45547617	Diabetes mellitus due to underlying condition with diabetic polyneuropathy	Condition	ICD10CM	5-char billing code		E08.42
45547620	Drug or chemical induced diabetes mellitus with neurological complications with diabetic neuropathy, unspecified	Condition	ICD10CM	5-char billing code		E09.40
45547621	Type 1 diabetes mellitus with diabetic chronic kidney disease	Condition	ICD10CM	5-char billing code		E10.22
45547622	Type 1 diabetes mellitus with diabetic peripheral angiopathy without gangrene	Condition	ICD10CM	5-char billing code		E10.51
45547624	Type 1 diabetes mellitus with other specified complication	Condition	ICD10CM	5-char billing code		E10.69
45547625	Type 2 diabetes mellitus with diabetic mononeuropathy	Condition	ICD10CM	5-char billing code		E11.41
45547635	Other specified diabetes mellitus with other diabetic ophthalmic complication	Condition	ICD10CM	5-char billing code		E13.39

VHA Data Portal



The screenshot shows the VHA Data Portal interface. At the top, there is a dark blue header with the Department of Veterans Affairs logo and the text "Department of Veterans Affairs VHA Data Portal". A search bar is located on the right side of the header. Below the header is a navigation menu with the following items: Data Sources, Data Access, Tools & Applications, Resources, Training, Policy & Admin, and Support. The "Data Sources" item is selected, and a breadcrumb trail shows "Data Sources > Data Sources Overview".

Quick View

The VHA Data Portal defines a data source as where the data comes from or a place to access data for a project or study.

Data Sources Overview

Overview

In VHA, the term data sources is often used to describe specific bodies of data such as, inpatient data, electronic health record data, mainframe SAS data sets, and the Corporate Data Warehouse (CDW).

The VHA Data Portal defines a data source as where the data comes from or a place to access data for a project or study. These data sources include data reports, data sets, data files, databases, data warehouses, and data repositories.

Data Sources

Select a data source from the list below to view a brief description and link to more information about each source. Use the "Expand All" button to view all data sources and links to their pages.

[+ Expand All](#)

- [Assistant Deputy Under Secretary for Health \(ADUSH\) Enrollment Files](#)
- [Austin Information Technology Center \(AITC\) Mainframe](#)
- [Beneficiary Identification & Records Locator System \(BIRLS\) Death File](#)
- [Bereaved Family Survey \(BFS\) Database](#)
- [Care Assessment Need \(CAN\) Scores](#)
- [Clinical Assessment, Reporting, and Tracking \(CART\) Program Data](#)
- [Corporate Data Warehouse \(CDW\)](#)
- [COVID-19 Shared Data Resource](#)



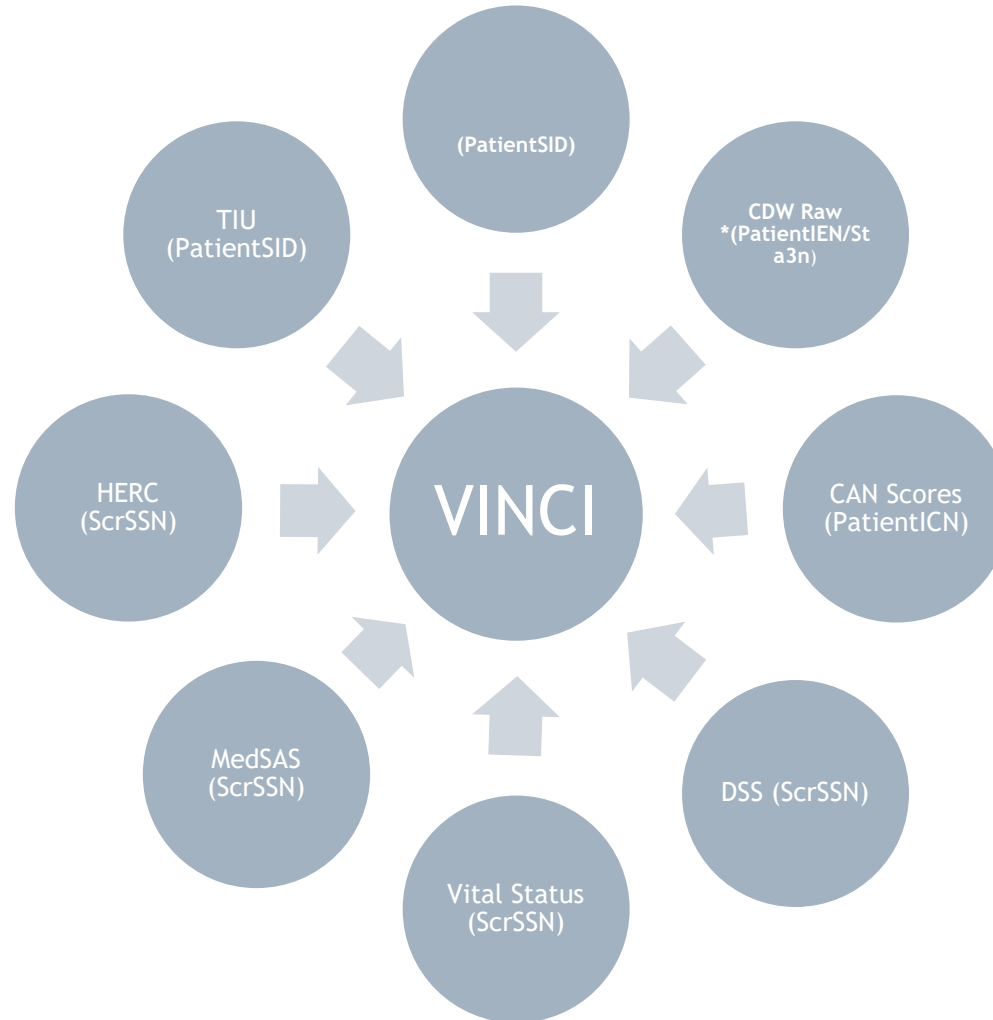
Understanding Patient Identifiers



VA Informatics &
Computing
Infrastructure



VINCI - Patient Identifiers By Data Provider



VINCI - Patient Identifiers

Description - Real SSN

- Generally 9 digits.
- Can have leading zeros.
- If an SSN is not provided by the patient then a pseudo SSN is assigned by the VistA system which serves as a place holder until the real SSN is obtained(a nine-digit number generated by the VistA system followed by the character 'P').
- Pseudo SSN generated is not a unique number, and the same pseudo SSN may be assigned to different patients.
- May rarely contain other formats (e.g., 10-digit numbers) if the SSN was entered as such in VistA.

VINCI - Patient Identifiers

Description - ScrSSN

- Used as a replacement of the SSN in many VHA data stores for privacy and security.
- Generated by the CDW as well as other VHA systems using the same algorithm.
- Used to identify patients in some non-CDW domains (MedSAS, DSS, Vital Status, HERC).
- The ScrSSN will be the same for a patient with a pseudo SSN as a patient that does not have a pseudo SSN if the first nine digits in their PatientSSN columns are the same. The patient with the pseudo SSN will have a 'P' following the first nine digits of their PatientSSN, but there is no 'P' on the end of their scrambled SSN. Thus, the ScrSSN may not be a unique identifier for a small number of patients.

VINCI - Patient Identifiers

Description - Sta3n

- 3-digit number representing the Vista system (station) where the data was entered.
- When using RAW data that has not had any processing and using any IEN, **MUST** use the sta3n as well to ensure uniqueness.
 - Example: “P.PatientIEN = X.PatientIEN AND P.Sta3n = X.Sta3n”

VINCI - Patient Identifiers

Description - PatientIEN

- Found in both the Patient.Patient and SPatient.SPatient tables.
- Internal Entry Number in a VistA system for a patient.
- Unique to a patient only in one VistA system, but not across all VistA systems.
- When using RAW data that has not had any processing, joins must use the PatientIEN and the Sta3n or duplicate rows will occur.

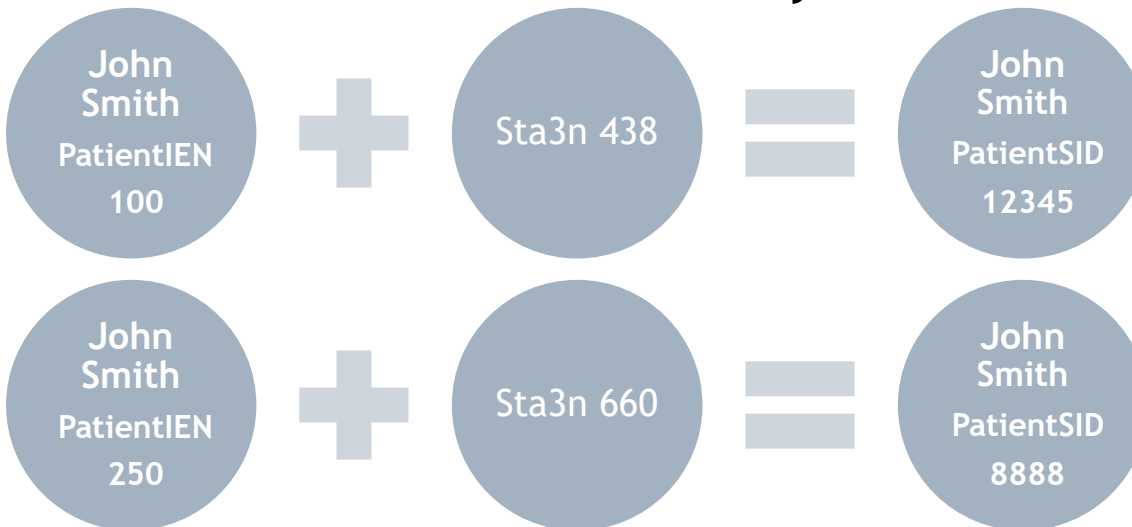
John Smith
PatientIEN 100
Sta3n 438

Beth Adams
PatientIEN 100
Sta3n 660

VINCI - Patient Identifiers

Description - PatientSID

- Primary patient identifier for CDW Production, TIU and some processed RAW data.
- Primary key of the Patient.Patient and SPatient.SPatient tables.
- Generated by the CDW for each unique combination of patient and VistA system (Sta3n).
- A patient may have more than one PatientSID if they have information in more than one VistA system.

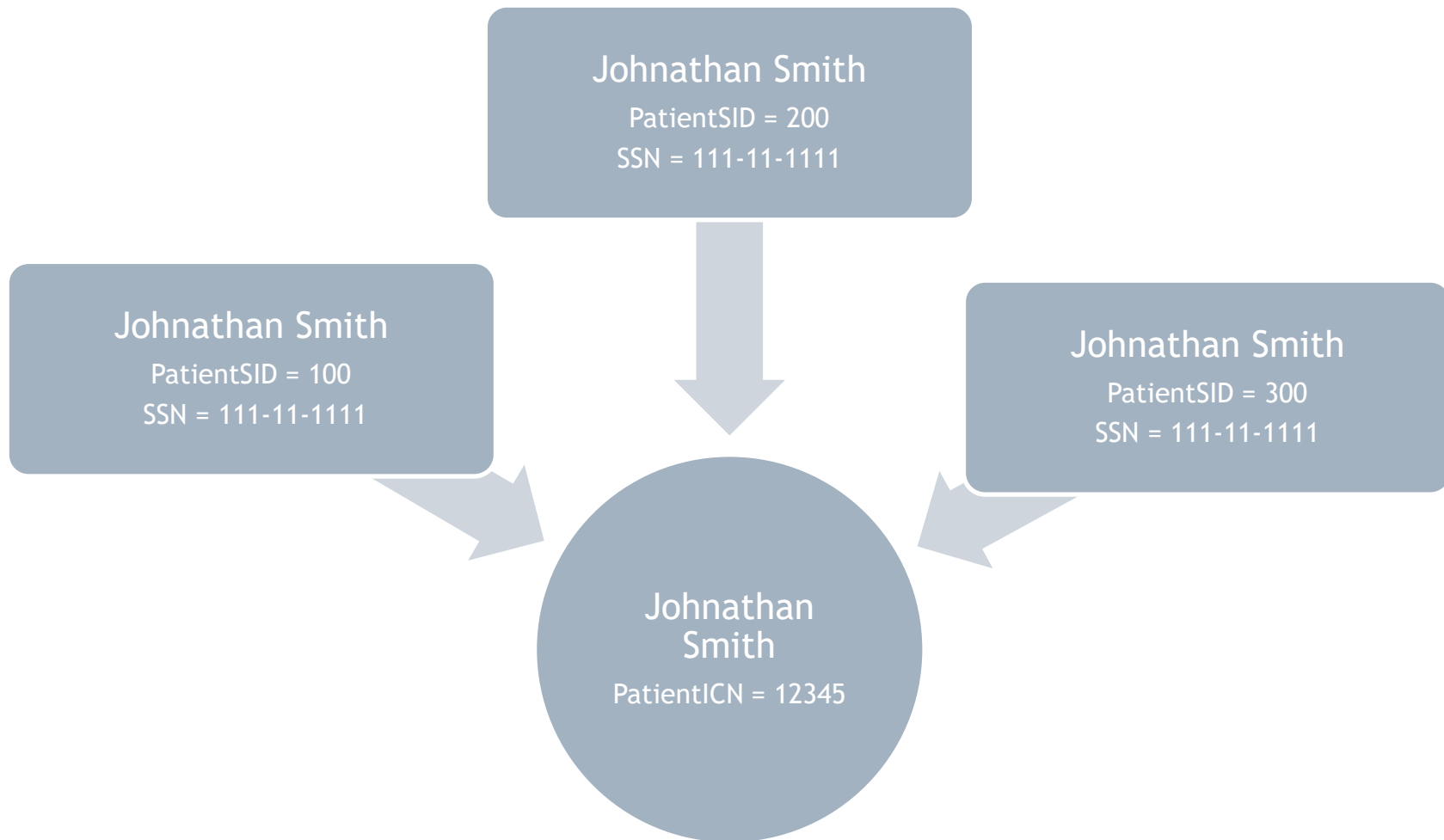


VINCI - Patient Identifiers

Description - PatientICN

- PatientICN is found in both the Patient.Patient and SPatient.SPatient tables.
- Used as the patient identifier for the CAN Scores domain.
- Created by Master Veteran Index (MVI).
- Not reused and are unique across all patients and all VHA systems including VistA systems.
- A patient should only have one ICN.
- If a study requests data from a domain that uses ScrSSN as a primary identifier, but **DOES NOT** want ScrSSN access, PatientICN can be substituted as an identifier.

VINCI - Patient Identifiers Description - PatientICN



Word of Caution

- Similarly named identifiers are not always matches
- PatientSID from the Patient domain is not the same as PITPatientSID from the PIT domain
- Check the Primary/Foreign Key relationship before you match

Indexing



VA Informatics &
Computing
Infrastructure



SQL Server Index

- An on-disk structure associated with a table or view that speeds retrieval of rows from that table or view.
- Contains keys built from one or more columns in the table or view

Clustered and Non-clustered

■ Clustered Index

- Sort and store the data rows in the table or view

■ Non-clustered Index

- Structure separate from the data rows
- Contains the non-clustered index key values and each has a pointer to the data row that contains the key value

Indexes and Constraints

- Indexes are automatically created when PRIMARY KEY and UNIQUE constraints are used on Table Columns
 - A PRIMARY KEY constraint will automatically create a clustered index
 - A UNIQUE constraint will automatically create a non-clustered index.

Clustered Columnstore

- Combines query efficiency and data compression
- Recommended for use on large tables like fact tables that tend to require full table scans

References

- VHA Data Portal - vaww.vhadataportal.med.va.gov
 - Data Sources
 - Data Access

- VIREC - vaww.virec.research.va.gov
 - HSRDataL-ListServ -
vaww.virec.research.va.gov/Support/HSRData-L.htm
 - Research User Guides(RUGs)
 - Researcher's Notebook

- CDW Support Site - [CDW Support \(sharepoint.com\)](http://CDW Support (sharepoint.com))

- BISL- [Business Intelligence Service Line \(BISL\) \(sharepoint.com\)](http://Business Intelligence Service Line (BISL) (sharepoint.com))

- CDW-Raw site - [CDW-Raw \(sharepoint.com\)](http://CDW-Raw (sharepoint.com))

References

- VINCI Central - [VINCI Central - Home \(va.gov\)](#)
 - Guides
 - Videos
 - VINCI environment documentation
 - VINCI Data Sources: [Data Sources \(va.gov\)](#)

- VHA Data Quality -
<http://vaww.vhadatportal.med.va.gov/AboutUs/DataQualityProgram.aspx>

- MCA/DSS - vaww.dss.med.va.gov/nationalrptg/nr_extracts.asp
✂ <http://vaww.virec.research.va.gov/NDEs/Documentation.htm>

- HERC
 - Dataset Guides: vaww.herc.research.va.gov/include/page.asp?id=guidebooks
 - Average Cost data
 - Non-VA Medical Care (Fee Basis)
 - MCA (DSS) Cost Extracts

References

- MedSAS: [VIReC/MedSAS/Documentation](#)
- COVID-19 Shared Data Resource: [COVID-19:Shared Data Resource - VA Phenomics Library](#)
- [VA Phenomics Library](#)
- VINCI OMOP Academy: [OMOP Academy \(va.gov\)](#)

References

- Indexing in SQL Server: [Indexes - SQL Server | Microsoft Docs](#)
- Clustered Columnstore Index: [Microsoft Columnstore Index](#)

Questions?

VINCI@va.gov