Virtual Data Warehouse
Data Specifications - Public

Version 5

7/21/2023

INTRODUCTION

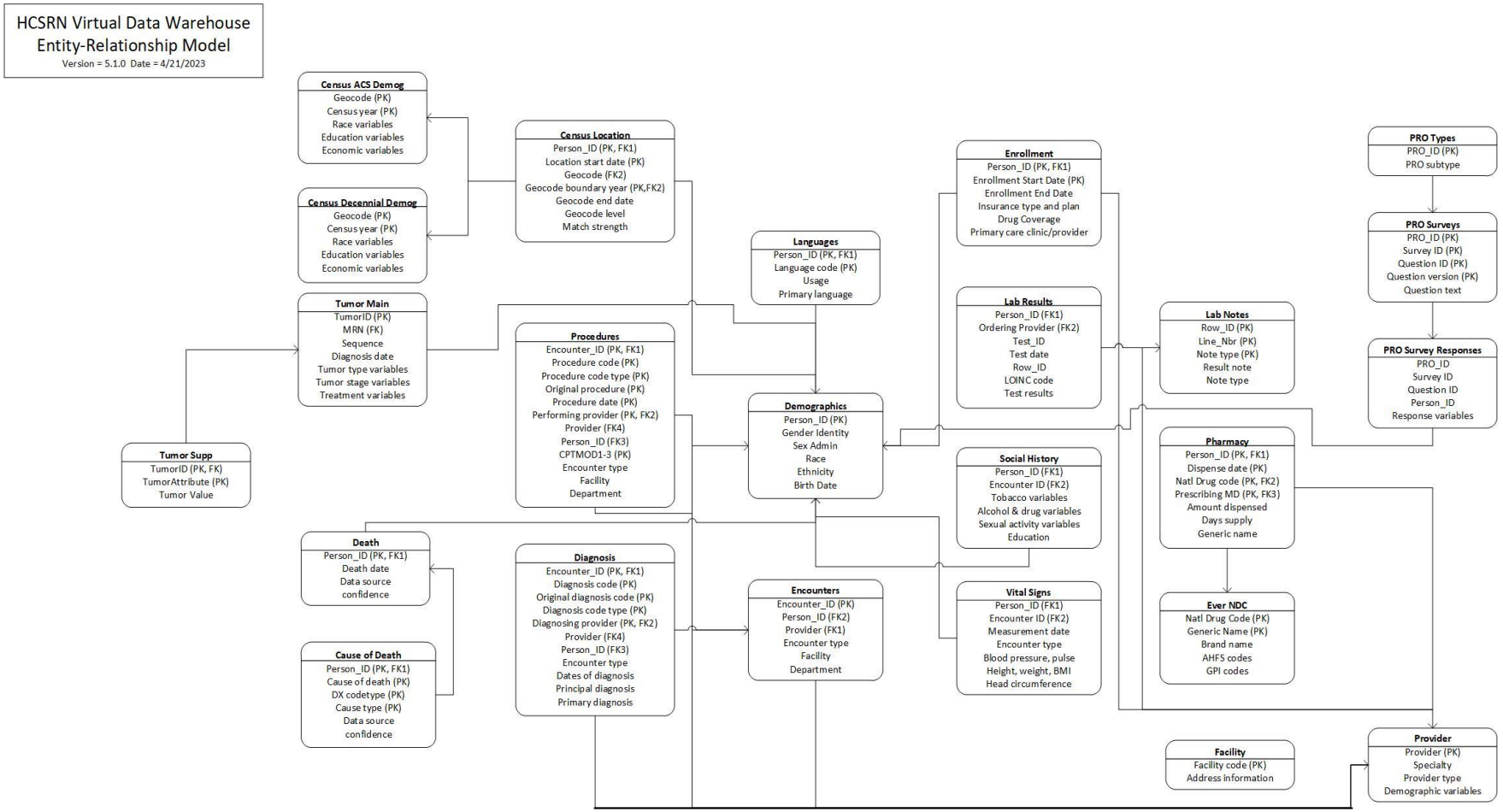
The HCSRN's Virtual Data Warehouse (VDW) is a pioneering common data model designed to support multi-site health system research. The VDW consists of clinical and claims data mapped to a common format to create tremendous efficiencies for data extraction, collection and management.

This document, the Virtual Data Warehouse Data Model Specifications, provides detailed data specifications for the VDW data model. These data areas are managed by site data managers and designated workgroups within the HCSRN sites. Some of the information contained in these specifications is only available to registered members of the HCSRN.

To submit a request for additional information regarding any of the data specs described in this document, please reach out to the VDW Operations Committee (VOC) at admin@hcsrn.org.

Table of Contents

Version = 5.1.1 Date=4/21/2023


HCSRN Virtual Data Warehouse
Entity-Relationship Model
Version = 5.1.0 Date = 4/21/2023
Census ACS Demog
Geocode (PK)
Census year (PK)
Race variables
Education variables
Economic variables
Census Decennial Demog
Geocode (PK)
Census year (PK)
Race variables
Education variables
Economic variables
Census Location
Person_ID (PK, FK1)
Location start date (PK)
Geocode (FK2)
Geocode boundary year (PK,FK2)
Geocode end date
Geocode level
Match strength
Tumor Main
TumorID (PK)
MRN (FK)
Sequence
Diagnosis date
Tumor type variables
Tumor stage variables
Treatment variables
Tumor Supp
TumorID (PK, FK)
TumorAttribute (PK)
Tumor Value
Procedures
Encounter_ID (PK, FK1)
Procedure code (PK)
Procedure code type (PK)
Original procedure (PK)
Procedure date (PK)
Performing provider (PK, FK2)
Provider (FK4)
Person_ID (FK3)
CPTMOD1-3 (PK)
Encounter type
Facility
Department
Languages
Person_ID (PK, FK1)
Language code (PK)
Usage
Primary language
Demographics
Person_ID (PK)
Gender Identity
Sex Admin
Race
Ethnicity
Birth Date
Death
Person_ID (PK, FK1)
Death date
Data source
confidence
Cause of Death
Person_ID (PK, FK1)
Cause of death (PK)
DX codetype (PK)
Cause type (PK)
Data source
confidence
Diagnosis
Encounter_ID (PK, FK1)
Diagnosis code (PK)
Original diagnosis code (PK)
Diagnosis code type (PK)
Diagnosing provider (PK, FK2)
Provider (FK4)
Person_ID (FK3)
Encounter type
Dates of diagnosis
Principal diagnosis
Primary diagnosis
Encounters
Encounter_ID (PK)
Person_ID (FK2)
Provider (FK1)
Encounter type
Facility
Department
Enrollment
Person_ID (PK, FK1)
Enrollment Start Date (PK)
Enrollment End Date
Insurance type and plan
Drug Coverage
Primary care clinic/provider
Lab Results
Person_ID (FK1)
Ordering Provider (FK2)
Test_ID
Test date
Row_ID
LOINC code
Test results
Social History
Person_ID (FK1)
Encounter ID (FK2)
Tobacco variables
Alcohol & drug variables
Sexual activity variables
Education
Vital Signs
Person_ID (FK1)
Encounter ID (FK2)
Measurement date
Encounter type
Blood pressure, pulse
Height, weight, BMI
Head circumference
Lab Notes
Row_ID (PK)
Line_Nbr (PK)
Note type (PK)
Result note
Note type
Pharmacy
Person_ID (PK, FK1)
Dispense date (PK)
Natl Drug code (PK, FK2)
Prescribing MD (PK, FK3)
Amount dispensed
Days supply
Generic name
Ever NDC
Natl Drug Code (PK)
Generic Name (PK)
Brand name
AHFS codes
GPI codes
Facility
Facility code (PK)
Address information
PRO Types
PRO_ID (PK)
PRO subtype
PRO Surveys
PRO_ID (PK)
Survey ID (PK)
Question ID (PK)
Question version (PK)
Question text
PRO Survey Responses
PRO_ID
Survey ID
Question ID
Person_ID
Response variables
Provider
Provider (PK)
Specialty
Provider type
Demographic variables

DEMOGRAPHICS

Version = 5.1.1 Date = 5/10/2022 StdVar = &_vdw_demographic

Subject Area Description

The DEMOGRAPHICS table contains patient/enrollee level descriptives for the people found in VDW tables. It serves as a lookup dataset for MRNs. Every MRN appearing in any other VDW file should appear in the Demographics table, even if demographics information on the person is unknown.

Variable Name	Definition	Type(Len)	Values
MRN	Medical record number is the unique patient identifier within a site and should never leave the site	char(*)	Unique to each patient at each site
BIRTH_DATE	The person's date of birth	num(4)	SAS Date
SEX_ADMIN	The person's "administrative sex"–the value stored in official/legacy sources at last ascertainment	char(1)	F = Female M = Male X = Neither Male Nor Female O = Other U = Unknown /uncertain / missing
SEX_AT_BIRTH	The person's sex as assigned at birth	char(1)	F = Female M = Male I = Intersex O = Other U = Uncertain, Unknown or Not recorded on birth certificate C = Choose not to disclose

GENDER_IDENTITY	The person's gender identity as subjectively experienced, on last ascertainment.	char(2)	FF = Female MM = Male FM = Female to Male transgender MF = Male to Female transgender GQ = Genderqueer or non-conforming or non-binary or genderfluid OT = Other ND = Choose not to disclose UN = Unknown
RACE1 - RACE5	The person's race. Preference is for self-reported; please see comment 1 for recording multiple race values	char(2)	HP = Native Hawaiian / Pacific Islander IN = American Indian / Alaskan Native AS = Asian BA = Black or African American WH = White MU = Multiple races with particular unknown OT = Other, values that do not fit well in any other value UN = Unknown or Not Reported
HISPANIC	Whether the person is of Hispanic origin / ethnicity	char(1)	Y = Yes N = No U = Unknown
NEEDS_INTERPRETER	Whether the person needs an interpreter to communicate with an English-only speaker	char(1)	Y = Yes N = No U = Unknown
SEXUAL_ORIENTATION1-SEXUAL_ORIENTATION3	The person's response(s) to inquiry into their sexual orientation	char(1)	B = Bisexual T = Heterosexual M = Homosexual A = Asexual P = Pansexual Q = Queer O = Other D = Does not know N = Choose not to disclose U = Not asked/no information

Primary Key:

MRN

Foreign Key Relationship:

Source Variable (Foreign Key)	Target Table	Target Variable (Primary Key)	Orphans allowed?
[None Defined]			

LANGUAGE

Version = 5.0.0 Date = 11/16/2021 StdVar = &_vdw_language

Subject Area Description

The LANGUAGE table contains information on the languages that patients speak and write. There is one record per person per known language. People on whom you have no language information should not be included in the table.

Variable Name	Definition	Type(Len)	Values
MRN	Medical record number is the unique patient identifier within a site and should never leave the site	char(*)	Unique to each patient at each site
LANG_ISO	A code signifying the language.	char(3)	As defined by ISO-639-2 or 'unk' for unknown Note that value set is lowercase
LANG_USAGE	How the person uses this language.	char(1)	S = Spoken/signed W = Written B = Both spoken and written U = Unknown
LANG_PRIMARY	For spoken languages, whether this is the person's primary spoken language.	char(1)	Y = Yes N = No U = Unknown

Primary Key:

MRN + LANG_ISO

Foreign Key Relationship:

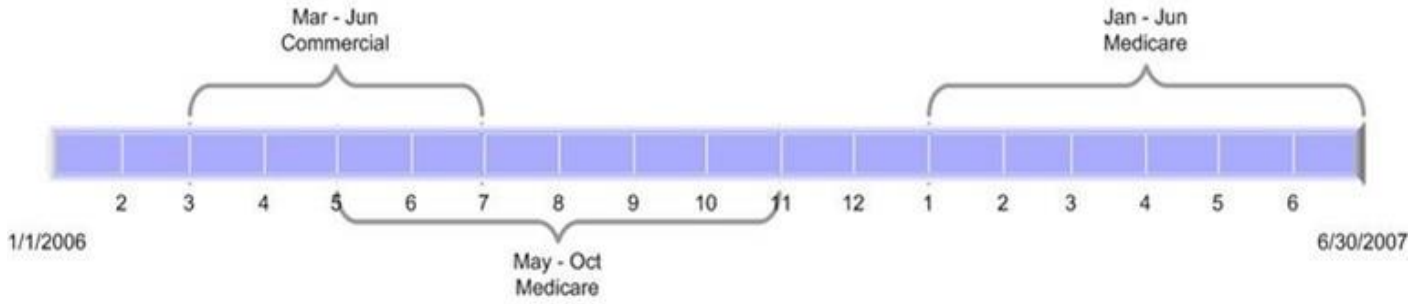
Source Variable (Foreign Key)	Target Table	Target Variable (Primary Key)	Orphans allowed?
MRN	DEMOGRAPHICS	MRN	No

ENROLLMENT

Version = 5.1.0 Date = 1/4/2022 StdVar = &_vdw_enroll

Subject Area Description

The ENROLLMENT table contains periods of time during which we believe to have at least the partial capture of medical care information either because of enrollment in a health insurance plan that we know about or as a patient of a care delivery system that we own or interact with. Insurance coverage is not the only basis of data capture (see ENROLLMENT_BASIS) nor do all records necessarily indicate completeness of data capture (see the capture warning variables, which all begin with 'INCOMPLETE'). Each record represents a period of time during which the information on the included variables was true. As many records as are necessary should be added to represent changes over time. There may be many contiguous records for a single period of enrollment, in order to account for these changes. However, periods may NOT overlap one another; there should be only one record covering any given day + medical record number. The following figure illustrates this concept:



Should result in 4 records:

MRN	enr_start	enr_end	ins_commercial	ins_medicare
111	03/01/2006	04/30/2006	Y	
111	05/01/2006	06/30/2006	Y	Y
111	07/01/2006	10/31/2006		Y
111	01/01/2007	06/30/2007		Y

Variable Name	Definition	Type(Len)	Values
MRN	Medical record number is the unique patient identifier within a site and should never leave the site	char(*)	Unique to each patient at each site
ENR_START	Beginning of the period at risk for medical care capture	num(4)	SAS Date
ENR_END	End of the period at risk for medical care capture	num(4)	SAS Date
INS_MEDICAID	Whether the person had any Medicaid insurance coverage during the period	char(1)	Y = Yes E = Yes, but from an External organization N = No U = Unknown
INS_COMMERCIAL	Whether the person had any commercial insurance coverage during the period	char(1)	Y = Yes E = Yes, but from an External organization N = No U = Unknown
INS_PRIVATEPAY	Whether the person had any insurance coverage in a private pay plan during the period	char(1)	Y = Yes E = Yes, but from an External organization N = No U = Unknown

INS_STATESUBSIDIZED	Whether the person had any state subsidized insurance coverage during the period	char(1)	Y = Yes E = Yes, but from an External organization N = No U = Unknown
INS_SELFFUNDED	Whether the person had any insurance coverage through an employer group that insures itself during the period	char(1)	Y = Yes E = Yes, but from an External organization N = No U = Unknown
INS_HIGHDEDUCTIBLE	Whether the person had any insurance coverage in a high deductible plan during the period as defined by the U.S. IRS (Pub 969) qualifying for a Health Savings Account	char(1)	Y = Yes E = Yes, but from an External organization N = No U = Unknown
INS_MEDICARE	Whether the person had any Medicare insurance coverage, including Medicare working aged, during the period	char(1)	Y = Yes E = Yes, but from an External organization N = No U = Unknown
INS_MEDICARE_A	Whether the person had Medicare Part A insurance coverage during the period	char(1)	Y = Yes E = Yes, but from an External organization N = No U = Unknown
INS_MEDICARE_B	Whether the person had Medicare Part B insurance coverage during the period	char(1)	Y = Yes E = Yes, but from an External organization N = No U = Unknown

INS_MEDICARE_C	Whether the person had Medicare Part C insurance coverage during the period	char(1)	Y = Yes E = Yes, but from an External organization N = No U = Unknown
INS_MEDICARE_D	Whether the person had Medicare Part D insurance coverage during the period	char(1)	Y = Yes E = Yes, but from an External organization N = No U = Unknown
INS_OTHER	Whether the person had insurance coverage during the period that is not otherwise included in the INS_* variables during the period	char(1)	Y = Yes E = Yes, but from an External organization N = No U = Unknown
PLAN_HMO	Whether the person had insurance coverage under an HMO plan during the period	char(1)	Y = Yes E = Yes, but from an External organization N = No U = Unknown
PLAN_POS	Whether the person had insurance coverage in a point-of-service plan during the period	char(1)	Y = Yes E = Yes, but from an External organization N = No U = Unknown
PLAN_PPO	Whether the person had insurance coverage in a preferred provider organization plan during the period	char(1)	Y = Yes E = Yes, but from an External organization N = No U = Unknown

PLAN_INDEMNITY	Whether the person had insurance coverage in a traditional indemnity plan during the period	char(1)	Y = Yes E = Yes, but from an External organization N = No U = Unknown
DRUGCOV	Whether the person had *any* insurance coverage that included at least some coverage for prescription drugs	char(1)	Y = Yes E = Yes, but from an External organization N = No U = Unknown
INCOMPLETE_OUTPT_RX	Is outpatient pharmacy fill data suspected to be incomplete for this person/period?	char(1)	K = There are known reasons to suspect incomplete capture N = No, there is no known reason to suspect incomplete capture X = This variable not implemented at this site
INCOMPLETE_OUTPT_ENC	Is outpatient encounter data suspected to be incomplete for this person/period?	char(1)	K = There are known reasons to suspect incomplete capture N = No, there is no known reason to suspect incomplete capture X = This variable not implemented at this site
INCOMPLETE_INPT_ENC	Is inpatient encounter data suspected to be incomplete for this person/period?	char(1)	K = There are known reasons to suspect incomplete capture N = No, there is no known reason to suspect incomplete capture X = This variable not implemented at this site

INCOMPLETE_EMR	Is electronic medical record data (e.g. social history, vital signs) suspected to be incomplete for this person/period?	char(1)	K = There are known reasons to suspect incomplete capture N = No, there is no known reason to suspect incomplete capture X = This variable not implemented at this site
INCOMPLETE_TUMOR	Is tumor data suspected to be incomplete for this person/period?	char(1)	K = There are known reasons to suspect incomplete capture N = No, there is no known reason to suspect incomplete capture X = This variable not implemented at this site
INCOMPLETE_LAB	Is lab results data suspected to be incomplete for this person/period?	char(1)	K = There are known reasons to suspect incomplete capture N = No, there is no known reason to suspect incomplete capture X = This variable not implemented at this site
ENROLLMENT_BASIS	The basis for the claim that we have at least partial capture of the medical care information during the period. Historically 'I'nsurance was the only permissible basis	char(1)	G = Geography I = Insurance B = Both geography & insurance P = Non-enrolled patient
PCC	The primary care clinic to which the patient is paneled in administrative record.	char(*)	Same value set as UTILIZATION.FACILITY_CODE
PCP	The clinician to which the patient is paneled in administrative record	char(*)	Values link as a foreign key to UTILIZATION.PROVIDER

Primary Key:

MRN + ENR_START (no overlapping time periods are permitted)

Foreign Key Relationship:

Source Variable (Foreign Key)	Target Table	Target Variable (Primary Key)	Orphans allowed?
MRN	DEMOGRAPHICS	MRN	No
PCP	PROVIDER	PROVIDER	No

ENCOUNTER

Version = 5.0.0 Date = 11/16/2021 StdVar = &_vdw_utilization

Subject Area Description

The ENCOUNTER table contains one record for each significant medical-related interaction between a patient and a medical provider regardless of care setting or type of encounter. All known encounters should be included whether known through claims, through clinical systems, or a third source type. Inclusion, exclusions, classification, and uniqueness of an encounter is dependent on the type of service and care setting. For a precise description, please refer to the Utilization Implementation Guidelines and the Virtual Visits Proposal and Guidelines on the private HCSRN website,

Variable Name	Definition	Type (Len)	Values
MRN	Medical record number is the unique patient identifier within a site and should never leave the site	char(*)	Unique to each patient at each site. Nulls are not allowed
ADATE	Encounter or admit date for inpatient or institutional stay. If encounter/admit date is unknown, then use the first date of a claim.	num(4)	SAS date. Nulls are not allowed
ATIME	Start time of an encounter. - Use admission time for inpatient, emergency or institutional encounters. - For other encounters (such as ambulatory visits), use check-in time if this field is populated, otherwise use appointment time. If unknown, specify as null	num(4)	SAS time. # of seconds since midnight (0-86,400), If input data are default time values (placeholders and not actual time measurements) and there isn't another source for time, set these default time values to missing. Sites can restrict this rule to specific encounter type/subtype values and/or data sources. null values are allowed

DDATE	Discharge date for inpatient and overnight encounters. End date for other encounters if documented in the source data. If discharge date is unknown, then use the last date of a claim.	num(4)	SAS date null values are allowed although this field is expected to be populated for inpatient and institutional stays
DTIME	End time of an encounter. - Use discharge time for inpatient, emergency or institutional encounters. - For other encounters (such as ambulatory visits), use check-out time if populated. If unknown, specify as null	num(4)	SAS time. # of seconds since midnight (0-86,400), If input data are default time values (placeholders and not actual time measurements) and there isn't another source for time, set these default time values to missing. Sites can restrict this rule to specific encounter type/subtype values and/or data sources. Null values are allowed
PROVIDER	Identifies the provider most responsible for the encounter.	char(*)	Unique to each provider at each site. If unknown, then use value "UNK"
ENC_ID	Uniquely identifies the encounter. The value should not change at update. An encounter should be unique across MRN, ADATE, ENCTYPE, PROVIDER, ENCOUNTER_SUBTYPE, FACILITY_CODE and appointment time.	char(*)	Unique to each encounter at each site.

ENCTYPE	The type of encounter. Valid associated ENCOUNTER_SUBTYPE values are in brackets	char(2)	AV = Ambulatory visit including outpatient clinics, same-day surgeries, observation beds, urgent care visits, and same-day ambulatory hospital encounters. Excludes emergency department. [SUB = OC, OB, SD, HA, UC, RH, DI, OT] ED = Emergency department excluding urgent care [SUB = HA, OC] IP = Acute inpatient hospital stay including inpatient stays, same-day hospital discharges, hospital transfers when patient was admitted into hospital, acute inpatient pysch, and detox stays. [SUB = AI] IS = Non-acute institutional stays including hospice, SNF, rehab, nursing home, residential, overnight non-hospital dialysis and other non-hospital stays [SUB = HS, SN, NH, RH, DI, OT] LO = A lab only encounter that cannot be matched to another encounter [SUB = OC, OT] OE=Other Encounters: Non-overnight hospice visits, home health visits, SNF visits, or other visits that do not occur in a typical ambulatory clinic or hospital setting. This encounter type can include individual professional services or provider rounding visits for inpatient or institutional stays where no facility claim could be identified in which to merge these visits. [Encounter_subtype=HS, HH, SN, RH, DI, OT, AI, NH] RO = A radiology only encounter that cannot be matched to another encounter [SUB = OC, OT] VC= Virtual Care - an encounter that takes place by phone, video, chat, online portal or another type of non-face to face mechanism. Can include synchronous or asychronous contact. [SUB= CH, EM, OA, OI, OT, PP, TC, TN, TS, VV]

ENCOUNTER_SUBTYPE	Further specification of the type of encounter. Valid associated ENCTYPE values are in brackets	char(2)	AI = Acute inpatient stay excluding observation bed [TYPE=IP, OE] CH = Online Text Chat [TYPE = VC] DI = Dialysis [TYPE = IS, AV, OE] EM= Email/Patient Message [TYPE=VC] HA = Hospital ambulatory; outpatient care at hospital excluding same-day surgery and observation beds [TYPE = AV, ED] HH = Home health [TYPE = OE] HS = Hospice [TYPE = IS, OE] NH = Nursing home including intermediate care facilities [TYPE = IS, OE] OA = Other Asynchronous Virtual Care [TYPE=VC] OI = Online Intake/Triage Form [TYPE=VC] OB = Observation bed [TYPE = AV] OC = Outpatient clinic visit [TYPE = AV, LO, RO, ED] PP = Provider to Provider/Facility to Facility Virtual Care [TYPE=VC] RH = Rehab [TYPE = IS, AV, OE] SD = Same-day surgery [TYPE = AV] SN = Skilled nursing facility [TYPE = IS, OE] TC = Telephone - Call Center/Advice/Triage [TYPE=VC] TN = Telephone - Not Specified [TYPE=VC] TS = Telephone - Scheduled Visit [TYPE=VC] VV = Video Visit [TYPE=VC] UC = Urgent care [TYPE = AV] OT = Other non-hospital [TYPE = IS, OE, AV, LO, RO, VC]
DRG_VERSION	Identifies the version of the Diagnostic Related Group value (DRG_VALUE). Expected for hospital and some institutional stays but populate for all encounters where known.	char(1)	A = CMS-DRG (used prior to 10/1/07) B = MS-DRG (used post 10/1/07) C = APR-DRG (All Patients Refined DRG) Null values are allowed
DRG_VALUE	The Diagnostic Related Group value. Used for hospital encounters. Using leading zeros for codes less than 100.	char(3)	Values maintained by Centers for Medicare & Medicaid Services . Null values are allowed.

ENC_COUNT	The number of visits associated with this encounter. Value may be greater than one when a claim indicates a number of visits, but dates of visits are not specified as may occur in a bundled claim e.g. of dialysis services.	num(4)	Positive integer value
ADMITTING_SOURCE	The location from which the patient was admitted for events with admit.	char(2)	AV=Ambulatory Visit ED=Emergency Department AF=Adult Foster Home AL =Assisted Living Facility HH=Home Health HS=Hospice HO=Home / Self Care IP=Other Acute Inpatient Hospital NH=Nursing Home (Includes ICF) OT=Other RS=Residential Facility RH=Rehabilitation Facility SN=Skilled Nursing Facility UN=Unknown

DISCHARGE_STATUS	The status of the patient at discharge for events with discharge.	char(2)	AF=Adult Foster Home AL =Assisted Living Facility AM=Against Medical Advice AW=Absent without leave EX=Expired HH=Home Health HS=Hospice HO=Home / Self Care IP=Other Acute Inpatient Hospital NH=Nursing Home (Includes ICF) OT=Other RS=Residential Facility RH=Rehabilitation Facility SH=Still In Hospital SN=Skilled Nursing Facility UN=Unknown
DISCHARGE_DISPOSITION	The disposition of the patient at discharge for events with discharge.	char(1)	A = Alive E = Expired U = Unknown
FACILITY_CODE	A code indiciating the facility, hospital, or clinic in which the encounter occurred.	char(*)	Unique to each facility at each site. If unknown, then use value "UNK"
DEPT	The department where the encounter took place as documented in the source data. This is not necessarily the specialty of the clinician providing services.	char(6)	If department information is unavailable in the source data (as may happen in claims data), specify as “UNKNWN”. For valid DEPT values, please see "Appendix A1 Dept".
DEPARTMENT	DEPRECATED	char(4)	As of version 4.0, DEPT has replaced DEPARTMENT. See Appendix A2 - DEPARTMENT for legacy values if needed.

SOURCE_DATA	Classification of the database that was used to create this record.	char(1)	**E** = Your site's EHR (Electronic Health Record) operated by your health care organization. Excludes claims and billing data. This category includes the following: • Direct extract from your site's EHR (or associated reporting database). • A secondary database that houses your site's EHR data (no code changes for billing) • Data for services captured in other local systems (such as a separate lab or radiology system) and interfaced into your site's EHR. **B** = Billing data--for services performed by your organization and captured in your org's EHR, but formatted for billing an outside insurer. These are sometimes referred to as 'internal' or 'outgoing' claims, or transaction data. **C** = Claims data--for services performed outside of your health care organization, presented to your organization for payment as the patient's insurer. **L** = Local data source but unrelated to your site's EHR . Often Includes pre-EHR systems prior to EHR implementation such as appointment management systems. **M** = Multiple sources--the encounter is an amalgam of data from different sources (e.g., inpatient encounters with data from claims and EHR rounding) **O** = Other. Not from the sources listed above. **U** = Unknown Can not be null
ELECTRONIC_CHART_REVIEW	Can this encounter be chart reviewed electronically?	char(1)	Y=Yes N=No P=Partially (part of the record can be reviewed electronically). A couple of examples: - If only a discharge summary is available for inpatient stays - if only rounding data is available for inpatient stays U=Unknown Can not be null

Primary Key:

ENC_ID

Foreign Key Relationship:

Source Variable (Foreign Key)	Target Table	Target Variable (Primary Key)	Orphans allowed?

DIAGNOSIS

Version = 5.0.0 Date = 11/16/2021 StdVar = &_vdw_dx

Subject Area Description

The DIAGNOSIS table contains all recorded diagnoses associated with the encounters indicated in the ENCOUNTERS table with the exception of admitting diagnoses for inpatient stays.

Variable Name	Definition	Type(Len)	Values
MRN	Medical record number is the unique patient identifier within a site and should never leave the site	char(*)	Unique to each patient at each site
ADATE	Refer to the ADATE variable on the ENCOUNTER table for definition, type, length, and value set. This variable's redundancy is to improve querying performance.		
ENCTYPE	Refer to the ENCTYPE variable on the ENCOUNTER table for definition, type, length, and value set. This variable's redunancy is to improve querying performance.		
ENC_ID	Foreign key to the ENCOUNTER table uniquely identifying the encounter.	char(*)	Unique to each encounter at each site.
PROVIDER	Refer to the PROVIDER variable on the ENCOUNTER table for definition, type, length, and value set. This variable's redundancy is to improve querying performance.		
DIAGPROVIDER	Identifies the provider that made the diagnosis. If unknown, set the value equal to the PROVIDER varaible.	char(*)	

DX	The diagnosis made. For ICD diagnosis coding, include decimal points in the value.	char(*)	ICD-9-CM Format ###.##, V##.##, E###.# ICD-10-CM Format A#@.@@@@ (3-8 characters - includes a decimal point (except for 3 digit codes)) First digit is alpha (A); Digit 2 is numeric (#); Digits 3-7 are alpha or numeric (@) Centers for Medicare & Medicaid Services http://www.cms.gov/Medicare/Coding/ICD9ProviderDiagnosticCodes/index.html
DX_CODETYPE	The coding set used in the DX variable for this record.	char(2)	07=’ICD-7-CM’ (including 'ICD-7') 08=’ICD-8-CM’ (including 'ICD-8') 09=’ICD-9-CM’ (including 'ICD-9') 10=’ICD-10-CM’ (including 'ICD-10') 11=’ICD-11-CM’ (including 'ICD-11') OT=’Other’
ORIGDX	The diagnosis code as reported in source data without standardization or cleaning.	char(*)	
PRINCIPAL_DX	For hospital admissions, whether this diagnosis is the principal discharge diagnosis of the encounter. The principal diagnosis indicates the main reason why the patient was admitted to the hospital for care and the value on which a DRG is assigned.	char(1)	P = Principal diagnosis N = Not principal diagnosis X = Unknown or not classifiable
PRIMARY_DX	Whether this diagnosis is the primary diagnosis of the encounter. The primary diagnosis is the most serious or resource intensive diagnosis and is the primary reason for the procedures being rendered.	char(1)	P = Primary diagnosis S = Secondary diagnosis X = Unknown or not classifiable

SOURCE_DATA_DX	Classification of the database that was used to create this record.	char(1)	E = Your site's EHR (Electronic Health Record) operated by your health care organization. Excludes claims and billing data. This category includes the following: • Direct extract from your site's EHR (or associated reporting database). • A secondary database that houses your site's EHR data (no code changes for billing) • Data for services captured in other local systems (such as a separate lab or radiology system) and interfaced into your site's EHR. B = Billing data--for services performed by your organization and captured in your org's EHR, but formatted for billing an outside insurer. These are sometimes referred to as 'internal' or 'outgoing' claims, or transaction data. C = Claims data--for services performed outside of your health care organization, presented to your organization for payment as the patient's insurer. L = Local data source but unrelated to your site's EHR . Often Includes pre-EHR systems prior to EHR implementation such as appointment management systems. O = Other. Not from the sources listed above. U = Unknown

Primary Key:

ENC_ID + DX + ORIGDX + DIAGPROVIDER + DX_CODETYPE

Foreign Key Relationship:

Source Variable (Foreign Key)	Target Table	Target Variable (Primary Key)	Orphans allowed?
MRN	DEMOGRAPHICS	MRN	No
ENC_ID	ENCOUNTER	ENC_ID	No
PROVIDER	PROVIDER	PROVIDER	No
DIAG_PROVIDER	PROVIDER	PROVIDER	No

PROCEDURE

Version = 5.0.0 Date = 11/16/2021 StdVar = &_vdw_px

Subject Area Description

The PROCEDURE table contains all recorded procedures that were actually performed associated with the encounters indicated in the ENCOUNTERS table.

Variable Name	Definition	Type(Len)	Values
MRN	Medical record number is the unique patient identifier within a site and should never leave the site	char(*)	Unique to each patient at each site
ENCTYPE	Refer to the ENCTYPE variable on the ENCOUNTER table for definition, type, length, and value set. This variable's redunancy is to improve querying performance.		
ADATE	Refer to the ADATE variable on the ENCOUNTER table for definition, type, length, and value set. This variable's redundancy is to improve querying performance.		
PROCDATE	The actual date when the procedure was performed. Unknown values should be coded as missing/null.	num(4)	SAS Date
ENC_ID	Foreign key to the ENCOUNTER table uniquely identifying the encounter.	char(*)	Unique to each encounter at each site.
PROVIDER	Refer to the PROVIDER variable on the ENCOUNTER table for definition, type, length, and value set. This variable's redundancy is to improve querying performance.		

PERFORMINGPROVIDER	Identifies the provider that performed the procedure. If unknown, set the value equal to the PROVIDER varaible.	char(*)	
PX	Code identifying the procedure that was performed	char(*)	##.## or ##.# ICD-9, $$$$$$$ ICD-10, #####, ####A, ####T CPT-4, A#### for HCPCS, ### or #### for Revenue Codes (match source data) # = Numeric Digit, A=Alphabet Letter $=Alpha or Numeric Convert local codes to standard codes if possible. Value set is dependent on the PX_CODETYPE value and is set by external organizations. Revenue codes may appear as both a 3-digit value and an equivalent 4-digit value with a leading zero, for example, both '123' and '0123' may appear in the data. Users should include both variants in filter/query conditions.
ORIGPX	The procedure code as reported in source data without standardization or cleaning.	char(*)	

PX_CODETYPE	Identifies the coding scheme used in the PX variable to identify the procedure performed.	char(2)	09 = ICD-9-CM 10 = ICD-10-CM 11 = ICD-11-CM C4 = CPT-4 H4 = HCPCS-4 RV = Revenue code LO = Local homegrown OT = Other
PXCNT	Number of times that the procedure was performed.	num(4)	Positive integer value
CPTMOD1	First modifier to a CPT or HCPCS procedure code used to communicate special circumstances related to the performance of a procedure.	char(2)	Value set is set by the American Medical Association and proprietary
CPTMOD2	Second modifier to a CPT or HCPCS procedure code used to communicate special circumstances related to the performance of a procedure.	char(2)	Value set is set by the American Medical Association and proprietary
CPTMOD3	Third modifier to a CPT or HCPCS procedure code used to communicate special circumstances related to the performance of a procedure.	char(2)	Value set is set by the American Medical Association and proprietary
SOURCE_CATEGORY_PX	Source of the procedure information. Order and billing pertain to internal healthcare processes and data sources. Claim pertains to data from the bill fulfillment, generally data sources held by insurers and other health plans.	char(2)	OD=Order (such as from your EHR's order table) BI=Billing (billing data such as transaction tables and internal claims created from services inside the health care organization) CL=Claim (external claims for services outside the health care organization) NI=No information UN=Unknown OT=Other (including but not limited to your EHR's encounter table)

SOURCE_DATA_PX	Classification of the source database that was used to create this record.	char(1)	E = Your site's EHR (Electronic Health Record) operated by your health care organization. Excludes claims and billing data. This category includes the following: • Direct extract from your site's EHR (or associated reporting database). • A secondary database that houses your site's EHR data (no code changes for billing) • Data for services captured in other local systems (such as a separate lab or radiology system) and interfaced into your site's EHR. B = Billing data--for services performed by your organization and captured in your org's EHR, but formatted for billing an outside insurer. These are sometimes referred to as 'internal' or 'outgoing' claims, or transaction data. C = Claims data--for services performed outside of your health care organization, presented to your organization for payment as the patient's insurer. L = Local data source but unrelated to your site's EHR . Often Includes pre-EHR systems prior to EHR implementation such as appointment management systems. O = Other. Not from the sources listed above. U = Unknown

Primary Key:

ENC_ID + PROCDATE + PERFORMINGPROVIDER + ORIGPX + PX + PX_CODETYPE + CPTMOD1-3

Foreign Key Relationship:

Source Variable (Foreign Key)	Target Table	Target Variable (Primary Key)	Orphans allowed?
MRN	DEMOGRAPHICS	MRN	No
ENC_ID	ENCOUNTER	ENC_ID	No
PROVIDER	PROVIDER	PROVIDER	No
PERFORMINGPROVIDER	PROVIDER	PROVIDER	No

PROVIDER

Version = 5.0.0 Date = 11/16/2021 StdVar = &_vdw_provider_specialty

Subject Area Description

The PROVIDER table contains proivder level descriptives for the providers found in VDW tables. Every provider appearing in any other VDW file should appear in the PROVIDER table, even if demographics information is unknown.

Variable Name	Definition	Type(Len)	Values
PROVIDER	The unique provider identifier.	char(*)	Unique to each provider at each site
SPECIALTY	The provider's specialty.	char(3)	See Appendix C - SPECIALTY for value list
SPECIALTY2 - SPECIALTYN	Optional variables (as many as desired) to indicate when a provider has multiple specialties.	char(3)	See Appendix C - SPECIALTY for value list
PROVIDER_TYPE	The position, job, title, or role of the provider.	char(3)	See Appendix C - PROVIDER_TYPE for value list
PROVIDER_BIRTH_YEAR	The year that the provider was born.	num(4)	Leave missing if unknown
PROVIDER_GENDER	The provider's gender and/or sex; if both gender and sex are known, this variable holds gender.	char(1)	M = Male F = Female O = Other including transgendered U = Unknown
PROVIDER_RACE	Refer to the RACE1 variable on the DEMOGRAPHICS table for definition, type, length, and value set.		
PROVIDER_HISPANIC	Refer to the HISPANIC variable on the DEMOGRAPHICS table for definition, type, length, and value set.		
YEAR_GRADUATED	The year that the provider graduated from medical/nursing/technical school.	num(4)	Leave missing if unknown

Primary Key:

PROVIDER

Foreign Key Relationship:

Source Variable (Foreign Key)	Target Table	Target Variable (Primary Key)	Orphans allowed?
[None Defined]			

FACILITY

Version = 5.0.1 Date = 1/4/2022 StdVar = &_vdw_facility

Subject Area Description

Contains the most recent information about each health care facility found in the
Update frequency: Updated after every update of the encounter table.
Include in the VDW. This is a HMORN VDW table to be used in multi-site

Variable name	Definition	Type(Len)	Values
FACILITY_CODE	VDW Facility Code as defined by each site	char(*)	VDW facility_code values Can not be null
RELATIONSHIP	Relationship between the facility and your health care organization	char(1)	O=Owned and/or operated by your health care organization. E=External facility. A facility NOT owned nor operated by your health care organization. Includes contract facilities. U=Unknown Cannot be null
RELATIONSHIP_HISTORY	Describes the relationhship history between the facility and your health care organization	char(1)	O=Always owned and/or operated by your health care organization E=Always an external facility (includes contract facilities) 1=Was an owned facility, most recently an external facility 2=Was an external facility, most recently an owned facility U=Unknown
FULL_ADDRESS	Full address of the facility.	CHAR(*)	Street address, city, state, zip combined in to one variable.

STREET_ADDRESS	Street address of the facility.	CHAR(*)	This field is populated if the information is easily available in the source system. Null values are allowed.
CITY	City where the facility is located	CHAR(*)	This field is populated if the information is easily available in the source system. Null values are allowed.
STATE	State where the facility is located	CHAR(2)	This field is populated if the information is easily available in the source system. Null values are allowed.
ZIP	Zip code of the facility (5 or 9 digits).	CHAR(9)	5 or 9 digit zip code, with leading zeroes, and no dashes. For example: 01234 12345 123456789 Null values are allowed.
ADDRESS_FACILITY_TYPE	Specify whether the address information is for a clinical or billing facility	CHAR(1)	C=Clinical Facility B=Billing Facility. Typically, this is the zip code for a claims vendor. U=Unknown (specify if zip is not missing) Leave as missing when ZIP is missing.
LATITUDE	The latitude of the location	num(8)	Value between -90 and +90 measured in degrees Null values are allowed.

LONGITUDE	The longitude of the location	num(8)	Value between -180 and +180 measured in degrees Null values are allowed.

Primary Key:

FACILITY_CODE. This field should be unique and never missing.

Foreign Key Relationship:

Source Variable (Foreign Key)	Target Table	Target Variable (Primary Key)	Orphans allowed?
[None Defined]			

PHARMACY

Version = 5.0.0 Date = 11/16/2021 StdVar = &_vdw_rx

Subject Area Description

The PHARMACY file contains data on medications dispensed in the outpatient setting. Dispensings in the inpatient setting are not included. Incomplete or unfilled medication orders are also excluded. Rows are unique on the combination of patient, NDC, dispense date, and prescribing provider (the primary key is under review, rolling up data is not required by the work group at this time). When multiple dispensings occur for the same patient for the same drug on the same day from the same provider, the amount dispensed and days supplied of the drug should be summed. Compound drugs may be represented differently across sites. If a drug of interest is a compound drug, users may want to check with individual sites on their handling of these dispensings.

Variable Name	Definition	Type (Len)	Values
MRN	Medical record number is the unique patient identifier within a site and should never leave the site	char(*)	Unique to each patient at each site
RXDATE	Date that medication was dispensed	num(*)	SAS Date
NDC	The identifier of a drug (or pharmacy product). If available, this should be the National Drug Code in the HIPAA/CMS/NCPDP standard 5-4-2 configuration without dashes. If a true NDC is not available, a locally defined unique identifier with up to 11 characters can be used	char(11)	FDA National Drug Code Directory
RXSUP	Number of days of medication supplied	num(8)	Either RXSUP or RXAMT must be a positive value.
RXAMT	Number of units (e.g. pills, tablets) dispensed	num(8)	Either RXSUP or RXAMT must be a positive value.

RXMD	The provider that prescribed this medication	char(*)	

Primary Key:

MRN + RXDATE + NDC + RXMD

Foreign Key Relationship:

Source Variable (Foreign Key)	Target Table	Target Variable (Primary Key)	Orphans allowed?
MRN	DEMOGRAPHICS	MRN	No
NDC	EVERNDC	NDC	No
RXMD	PROVIDER	PROVIDER	No

EVERNDC

Version = 5.0.1 Date = 4/21/2023 StdVar = &_vdw_everndc

Subject Area Description

The EVERNDC table is a lookup table containing all National Drug Codes (NDC) found in the PHARMACY table at each site. That is, pharmacy dispensings occurring at site A will have at least one corresponding record in site A's EVERNDC table but not necessarily any records in site B's EVERNDC table.

Variable Name	Definition	Type(Len)	Values
NDC	The identifier of a drug (or pharmacy product). If available, this should be the National Drug Code in the HIPAA/CMS/NCPDP standard 5-4-2 configuration without dashes. If a true NDC is not available, a locally defined unique identifier with up to 11 characters can be used and the NDC_SITE_SPECIFIC_FLAG should be	char(11)	FDA National Drug Code Directory
NDC_FDA	The original FDA National Drug Code in 4-4-2, 5-3-2, or 5-4-1 format including dashes	char(12)	
GENERIC	The generic name of the drug in all upper case with form, strength, and other labels expunged	char(105)	
BRAND	The brand name of the drug in all upper case with form, strength, and other labels expunged	char(105)	
AHFS1 - AHFS7	The American Society of Health-System Pharmacists drug classification code. If code is in the 6 digit format, pad with two zeros at the end of the code. Store up to seven different AHFS codes across these variables	char(8)	Value set is set by AHFS and considered proprietary.
GPI	The Generic Product Identifier drug classification code	char(14)	Value set is set by Medi-Span and is considered proprietary.
UNIT_OF_MEASURE	The unit of measure as reported in source data without standardization or cleaning except for storing values in all upper case	char(*)	Example values are 2000, U/4ML, GM/15ML, %/5GM, -400 UNIT

STRENGTH	The drug strength as reported in source data without standardization or cleaning except for storing values in all upper case	char(*)	Example values are 99.99%, 9G, 9MCG/0.3ML, 9000 UNIT
DOSAGE_FORM	The dosage form as reported in source data without standardization or cleaning except for storing values in all upper case	char(*)	Example values are VIAL, VIAL PORT, WAFER, WAF, SOLUTION, SOLN SEQ, SPIRIT
OBSOLETE_DT	The date the NDC was made obsolete, only if the drug is reinstated with a different NDC/Generic combination at a later time	num(*)	SAS Date
REINSTATED_DT	The date the NDC was resinstated, only if the drug was reinstated with a different NDC/Generic combination	num(*)	SAS Date
NDC_SITE_SPECIFIC_FLAG	Indicates whether an NDC is a site-specific, non-standard item. This includes but is not limited to compound drugs, stock items, repackaged drugs and study medications. Some site specific NDCs are identifiable via a starting set of characters, such as 11111 or 99999, or a term.	char(1)	Y = Yes N = No U = Unknown

Primary Key:

NDC + GENERIC

*Duplicate NDCs occur when an NDC is retired and reused for a new drug.

Foreign Key Relationship:

Source Variable (Foreign Key)	Target Table	Target Variable (Primary Key)	Orphans allowed?
[None Defined]			

LAB RESULTS

Version = 5.0.0 Date = 11/16/2021 StdVar = &_vdw_lab

Subject Area Description

The LAB RESULTS table contains the results of laboratory tests performed on patients. If a test is not resulted for whatever reason (e.g. specimen not sufficient, patient did not show), then that test should not appear in the table.

Variable Name	Definition	Type(Len)	Values
MRN	Medical record number is the unique patient identifier within a site and should never leave the site	char(*)	Unique to each patient at each site
TEST_TYPE	VDW-specific classification of laboratory results.	char(20)	• The Test_Type, LongName, LOINC, Priority excel spreadsheet maintained by the VDW Lab Workgroup lists all test_types currently defined along with LOINCs associated with each test_type. • The same Test_Type-LOINC association can be found in the EverLOINC SAS table also maintained by the VDW Lab Workgroup. • See Appendix D - Laboratory Test Reference for VDW Laboratory File PDF (The manual serves as a guide to help programmers associate source records with the correct Test_Type. The manual is not up to date.

STAT	Indicates the immediacy of the test. The intent of this value is to determine whether the test was obtained as part of routine care or as an emergent/urgent diagnostic test.	char(1)	E=Expedite S=Stat R=Routine U=Unknown or missing
LOINC	Logical Observation Identifiers Names and Codes (LOINC) is a universal coding system for laboratory tests and measurements developed by Regenstrief Institute. Follow http://loinc.org/ to download a list of all LOINC codes defined by Regenstrief. Not all LOINCs have been mapped to VDW test_types.	char(10)	• The Test_Type, LongName, LOINC, Priority excel spreadsheet maintained by the VDW Lab Workgroup lists all test_types currently defined along with LOINCs associated with each test_type. • The same Test_Type-LOINC association can be found in the EverLOINC SAS table also maintained by the VDW Lab Workgroup. • LOINC codes are in the form nnnnnnnn-n with one to 8 digits preceding the hyphen. Do not fill LOINC codes with leading zeros. Format is left-justified. The last digit of the LOINC code is a check digit and is always preceded by a hyphen. The hyphen, as well as all the numbers, is part of the LOINC and must be included.
PT_LOC	Location of the patient when the lab specimen was obtained.	char(1)	I=Inpatient O=Outpatient E=Emergency Department H=Home U=Unknown or missing

RESULT_LOC	Location where the result was completed. Particularly whether the lab was completed in a certified laboratory.	char(1)	L = Lab P = Point of care (e.g. home, provider office) Code nulls and unknowns as “L”. There are no missing values.
SPECIMEN_ID	Used to uniquely identify a collected specimen which may ultimately be used to obtain multiple lab results. Specimen ID will be used to connect multiple records from the same blood sample.	char(*)	

SPECIMEN_SOURCE	The source or method for collecting the specimen.	char(6)	BLOOD SERUM PLASMA SR_PLS = serum and/or plasma PPP = Platelet Poor Plasma CSF = cerebral spinal fluid URINE STOOL NSWAB = nasal swab (including nose) NWASH = nasal wash NPH = nasopharyngeal swab NPWASH = nasopharyngeal wash THRT = throat, oropharyngeal swab SALIVA SLFSWB = self-swab (similar to NSWAB but with 'self' designation) SPUTUM TR_ASP = tracheal aspirate SPUTUM BAL = bronchoalveolar lavage (BAL) BALBX = BAL biopsy BLDDOT OTHER NS = not specified [Null] = missing value
LOCAL_CD	The lab test type code as reported in source data without standardization or cleaning. This code indicates in the data source system which test was performed.	char(*)	Unique to each site
BATTERY_CD	The battery code as reported in source data without standarization or cleaning. A battery consists of a grouping or series of tests performed and is sometimes referred to as a panel of tests.	char(*)	Unique to each site

PX	Refer to the PX variable on the Procedure file for type and length.	Refer to the PX variable on the Procedure file for value set EXCEPT missing is also an acceptable value on the lab result table.	
PX_CODETYPE	Refer to the PX_CodeType variable on the Procedure file for type and length.	Refer to the PX_CodeType variable on the Procedure file for value set EXCEPT missing is also an acceptable value on the lab result table.	
ORDER_ID	Uniquely identifies the order for this lab test. In the future, may link to a table of orders.	char(*)	Unique to each site
ORDER_DT	Date that the lab test was ordered. It is possible for the order date to be any date from months before the sample was taken (in the case of a recurring test to monitor a chronic condition or drug interaction) to the same day the sample was taken.	num(4)	SAS date
LAB_DT	Date that the specimen was collected.	num(4)	SAS date
LAB_TM	Time that the specimen was collected.	num(4)	SAS time
RESULT_DT	Date that the test was resulted. This date could be the same day the specimen was collected or any date up to weeks later.	num(4)	SAS date
RESULT_TM	Time that the specimen was resulted.	num(4)	SAS time

RESULT_C	The result of the test stored in a character field. This variable works in conjunction with the Modifier variable.	char(20)	• VDW Qualitative Laboratory Test Standards spreadsheet lists expected results for qualitative test types. • See Appendix D - Laboratory Test Reference for VDW Laboratory File PDF (The manual serves as a guide to help programmers correctly associate source records with Test_Types. The manual does not contain an entry for more recently defined test_types. • SEE LAB NOTES TABLE
MODIFIER	Modifies the value stored in the Result_C field.	char(2)	TX = Text EQ = Equal LT = Less than LE = Less than or equal to GT = Greater than GE = Greater than or equal to RA = Range
RESULT_UNIT	The units in which the result is reported after basic standardizations have been applied. At the least, the value should be uppercase and left justified.	char(11)	Examples of possible values: %, U/L, MG/DL, K/UL, 10^6/UL
RESULT_UNIT_RAW	The units in which the result is reported before basic standardizations are applied.	char(11)	Examples of possible values: PER, Percent, units/L, mg/dL, 10*3/uL, 10^9/L, x10e6/uL
NORMAL_LOW_C	The lowest value still considered normal for this test. This variable works in conjunction with the Modifier_Low variable.	char(8)	

MODIFIER_LOW	Modifies the value stored in the Normal_Low_C field.	char(2)	EQ = Equal GT = Greater than GE = Greater than or equal to Null if Normal_Low_C is null The following will probably not be needed for this variable TX = Text LT = Less than LE = Less than or equal to
NORMAL_HIGH_C	The highest value still considered normal for this test. This variable works in conjunction with the Modifer_High variable.	char(8)	
MODIFIER_HIGH	Modifies the value stored in the Normal_High_C field.	char(2)	EQ = Equal LT = Less than LE = Less than or equal to Null if Normal_High_C is null The following will probably not be needed for this variable TX = Text GT = Greater than GE = Greater than or equal to

ABN_IND	Indicates whether the test result is abnormal.	char(2)	AB = Abnormal C = Critical AH = Abnormal high CH = Critical high AL = Abnormal low CL = Critical low IN = Inconclusive NL = Normal UK = Unknown or missing
ORDER_PROV	Identifies the provider that ordered the lab test.	char(*)	Unique to each site. Value set is same as rxmd in pharmacy or provider in utilization.
ORD_DEPT	The department or specialty in which the order took place.	char(6)	See Appendix A1 - DEPT. These values are the same as for the variable DEPT in encounters.
FACILITY_CODE	A code indiciating the facility, hospital, or clinic in which the lab order originated.	char(*)	Unique to each site. Value set is same as Facility_Code in utilization.
ROW_ID	A site specific identifer that enables linkage between the LAB NOTES and LAB RESULTtables.	type and length are site specific	Unique to each site

Primary Key:

No primary key defined

Foreign Key Relationship:

Source Variable (Foreign Key)	Target Table	Target Variable (Primary Key)	Orphans allowed?
MRN	DEMOGRAPHICS	MRN	No
ORDER_PROV	PROVIDER	PROVIDER	No

LAB NOTES

Version = 5.0.0 Date = 11/16/2021 StdVar = &_vdw_lab_notes

Subject Area Description

The LAB NOTES table includes text notes that may accompany lab test results. If a note is long, it may span over several records in the table. The LAB NOTES table links to the LAB RESULT table by the variable Row_ID. Not all records in the LAB RESULTS table will have corresponding

Variable Name	Definition	Type(Len)	Values
ROW_ID	A site specific identifer that enables linkage between the LAB NOTES and LAB RESULTtables.	type and length are site specific	
RESULT_NOTE	Contains result note (or, theoretically, this field might be used to hold results or lower boundary or upper boundary of normal range values that require more than 8 bytes of storage).	char(80)	
NOTE_TYPE	Identifies the type of note in the Result_Note field.	char(1)	R = Result L = Low normal value H = High normal value N = Note
LINE	A line counter starting at one and incrementing by one as needed to accommodate each Row_ID / Note_Type combination.	num(4)	Positive integer

Primary Key:

ROW_ID + NOTE_TYPE + LINE

Source Variable (Foreign Key)	Target Table	Target Variable (Primary Key)	Orphans allowed?
No Foreign Key			

VITAL SIGNS

Version = 5.0.0 Date = 11/16/2021 StdVar = &_vdw_vitalsigns

Subject Area Description

The VITAL SIGNS table includes physiological (i.e. temperature, pulse, respiratory rate, blood pressure) and anthropometric (i.e. height, weight, BMI) measures taken by health professionals during encounters. It also includes these measures reported by a patient in a Virtual Care Encounter and measures collected through an approved medical device.

Variable Name	Definition	Type(Len)	Values
MRN	Medical record number is the unique patient identifier within a site and should never leave the site	char(*)	Unique to each patient at each site
ENC_ID	Foreign key to the ENCOUNTER table uniquely identifying the encounter; unlinkable records should contain null/missing values	char(*)	Unique to each encounter at each site
MEASURE_DATE	The date on which these vital signs were measured	num(4)	SAS Date
MEASURE_TIME	The time at which these vital signs were measured	num(4)	SAS Time
ENCTYPE	Refer to the ENCTYPE variable on the ENCOUNTER table for definition, type, length, and value set		
HT	The height measured in inches of the patient at this time	num(8)	Positive real number HT missing if: age = 0 and (ht < 3 or ht > 41) age between 1 and 5 and (ht < 12 or ht > 60) age between 6 and 12 and (ht < 20 or ht > 84) age between 13 and 17 and (ht < 30 or ht > 108) age > = 18 and (ht < 36 or ht > 108)

WT	The weight measured in pounds of the patient at this time	num(8)	Positive real number WT missing if: age = 0 and (wt < 0 or wt > 80) age between 1 and 5 and (wt < 9 or wt > 200) age between 6 and 12 and (wt < 20 or wt > 350) age between 13 and 17 and (wt < 25 or wt > 650) age > = 18 and (wt < 50 or wt > 1000)
SYSTOLIC	The systolic blood pressure measured in mmHg of the patient at this time	num(4)	Integer SYSTOLIC missing if < 50 or > 300
DIASTOLIC	The diastolic blood pressure measured in mmHg of the patient at this time	num(4)	Integer DIASTOLIC missing if < 20 or > 160
BP_TYPE	The type of blood pressure taken	char(1)	R = Rooming O = Orthostatic M = Multiple E = Extended
POSITION	The patient's position for orthostatic blood pressure measurements	char(1)	1 = Sitting 2 = Standing 3 = Supine Null = Unknown
HT_RAW	The height measurement as reported in source data without standardization or cleaning	char(*)	Examples may include specific values, ranges, or categories
WT_RAW	The weight measurement as reported in source data without standardization or cleaning	char(*)	Examples may include specific values, ranges, or categories
BMI_RAW	The patient's body mass index measurement as reported in source data without standardization, calculation, or cleaning	char(*)	
HEAD_CIR_RAW	The patient's head circumference measurement as reported in source data without standardization or cleaning	char(*)	

SYSTOLIC_RAW	The systolic blood pressure measurement as reported in source data without standardization or cleaning	char(*)	Examples may include specific values, ranges, or categories
DIASTOLIC_RAW	The diastolic blood pressure measurement as reported in source data without standardization or cleaning	char(*)	Examples may include specific values, ranges, or categories
RESPIR_RAW	The respirations in breaths per minute measurement as reported in source data without standardization or cleaning	char(*)	
TEMP_RAW	The patient's body temperature measurement as reported in source data without standardization or cleaning	char(*)	
PULSE_RAW	The heartbeats per minute measurement as reported in source data without standardization or cleaning	char(*)	

Primary Key:

No primary key defined

Foreign Key Relationship:

Source Variable (Foreign Key)	Target Table	Target Variable (Primary Key)	Orphans allowed?
MRN	DEMOGRAPHICS	MRN	No
ENC_ID	ENCOUNTER	ENC_ID	No

SOCIAL HISTORY

Version = 5.0.1 Date = 07/21/2023 StdVar = &_vdw_social_hx

Subject Area Description

The SOCIAL HISTORY table contains behavioral measures taken by health professionals during clinic visits, over the telephone, or via questionnaires. These measures include the use of tobacco, alcohol, and illegal drugs as well as sexual behavior and contraceptive use. Because of the range of tobacco products available and the variability in the level of detail recorded, precision and capture of tobacco related variables may vary both across and within sites. Consistency of response is also not guaranteed over time.

Social history measures may carry special privacy concerns. Beyond the nature of these variables, the use of free text fields in any content area warrants additional care as they may contain personal health information. Users are encouraged to consult with privacy experts and/or experienced users before extracting data if they are unfamiliar with these additional data sensitivities.

Variable Name	Definition	Type(Len)	Values
MRN	Medical record number is the unique patient identifier within a site and should never leave the site	char(*)	Unique to each patient at each site
CONTACT_DATE	Date of encounter or date of social history data collection	num(4)	SAS date, missing is not allowed
ENC_ID	Foreign key to the ENCOUNTER table uniquely identifying the encounter; unlinkable records should contain null/missing values	char(*)	SAS date, missing values allowed
EDUCATION_YEARS	Number of years of education completed recorded as free text	char(*)	Free text, missing values allowed

SEX_ACTIVE	Whether the person is sexually active	char(1)	Y = Yes N = No W = Not currently (Was) X = Not asked U = Unknown or missing
SEX_FEMALE_PARTNER	Whether the person has female sex partners	char(1)	Y = Yes N = No X = Not asked U = Unknown or missing
SEX_MALE_PARTNER	Whether the person has male sex partners	char(1)	Y = Yes N = No X = Not asked U = Unknown or missing
SEX_COMMENT	Comment about the person's sexual activity	char(*)	Free text, missing values allowed
BC_ABSTINENCE	Whether the person uses abstinence as birth control	char(1)	Y = Yes N = No X = Not asked U = Unknown or missing
BC_CONDOM	Whether the person uses condoms as birth control	char(1)	Y = Yes N = No X = Not asked U = Unknown or missing
BC_DIAPHRAGM	Whether the person uses a diaphragm as birth control	char(1)	Y = Yes N = No X = Not asked U = Unknown or missing
BC_IMPLANT	Whether the person uses an implant as birth control	char(1)	Y = Yes N = No X = Not asked U = Unknown or missing

BC_INJECTION	Whether the person uses injections as birth control	char(1)	Y = Yes N = No X = Not asked U = Unknown or missing
BC_INSERTS	Whether the person uses inserts as birth control	char(1)	Y = Yes N = No X = Not asked U = Unknown or missing
BC_IUD	Whether the person uses an intrauterine device (IUD) as birth control	char(1)	Y = Yes N = No X = Not asked U = Unknown or missing
BC_PILL	Whether the person uses birth control pills	char(1)	Y = Yes N = No X = Not asked U = Unknown or missing
BC_RHYTHM	Whether the person uses the rhythm method as birth control	char(1)	Y = Yes N = No X = Not asked U = Unknown or missing
BC_SPERMICIDE	Whether the preson uses spermicide as birth control	char(1)	Y = Yes N = No X = Not asked U = Unknown or missing
BC_SPONGE	Whether the person uses a sponge as birth control	char(1)	Y = Yes N = No X = Not asked U = Unknown or missing
BC_SURGICAL	Whether a surgery (e.g. vasectomy, hysterectomy) significantly reduces the chance of contraception	char(1)	Y = Yes N = No X = Not asked U = Unknown or missing

BC_OTHER	Whether another type of birth control is used that has not been previously specified	char(1)	Y = Yes N = No X = Not asked U = Unknown or missing
BC_COMMENT	Comment about the person's birth control	char(*)	Free text, missing values allowed
ALCOHOL_USE	Whether the person is an alcohol drinker	char(1)	Y = Yes N = No I = Infrequent Q = Quit/Former X = Not asked U = Unknown or missing
ALCOHOL_BEER_NWK	The number of drinks per week, beer	char(*)	Drinks/wk, beer; could be a range; 0 is valid
ALCOHOL_WINE_NWK	The number of drinks per week, wine	char(*)	Drinks/wk, wine; could be a range; 0 is valid
ALCOHOL_LIQ_NWK	The number of drinks per week, hard liquor	char(*)	Drinks/wk, liquor; could be a range; 0 is valid
ALCOHOL_UNSPEC_NWK	The number of drinks per week, unspecified type	char(*)	Drinks/wk, unspecified; could be a range; 0 is valid
ALCOHOL_COMMENT	Comment about the person's alcohol use	char(*)	Free text, missing values allowed
ILLICIT_DRUG_USE	Whether the person uses illicit drugs	char(1)	Y = Yes N = No I = Infrequent P = Passive Q = Quit/Former X = Not asked U = Unknown or missing
ILLICIT_DRUG_USE_FREQ	Frequency of illicit drug use as free text	char(*)	Free text, missing values allowed
ILLICIT_DRUG_USE_COMMENT	Comment about the person's illicit drug use	char(*)	Free text, missing values allowed

IV_DRUG_USE	Whether the person uses IV drugs	char(1)	Y = Yes N = No I = Infrequent Q = Quit/Former X = Not asked U = Unknown or missing
TOBACCO_USE	Whether the person is a user of tobacco	char(1)	Y = Yes / Current I = Infrequent N = Never Q = Quit / Former P = Passive/Environmental/Second hand X = Not Asked U = Unknown or missing
TOBACCO_USE_YEARS	The number of years the person used tobacco	char(*)	Free text, missing values allowed
TOBACCO_PACKS_DAY	Number of packs smoked per day as free text	char(*)	Free text, missing values allowed
TOBACCO_SMOKELESS_USE	Whether the person uses smokeless tobacco	char(1)	Y = Yes / Current I = Infrequent N = No / Never used Q = Quit / Former X = Not asked U = Unknown or missing
TOBACCO_SMOKELESS_START_DATE	Date the person started using smokeless tobacco	num(4)	SAS Date, missing allowed
TOBACCO_SMOKELESS_QUIT_DATE	Date the person quit using smokeless tobacco	num(4)	SAS Date, missing allowed

TOBACCO_SMOKING_USE	Whether the person uses smoking tobacco	char(1)	E = Current every day S = Current some days H = Heavy smoker L = Light smoker N = Never P = Passive Q = Former U = Unknown X = Never assessed Y = Smoker - current status unknown
TOBACCO_SMOKING_START_DATE	Date the person started using smoked tobacco	num(4)	SAS Date, missing allowed
TOBACCO_SMOKING_QUIT_DATE	Date the person quit using smoked tobacco	num(4)	SAS Date, missing allowed
TOBACCO_CHEW	Whether the person uses chew tobacco	char(1)	Y = Yes N = No X = Not asked U = Unknown or missing
TOBACCO_CIGARETTES	Whether the person uses tobacco cigarettes	char(1)	Y = Yes N = No X = Not asked U = Unknown or missing
TOBACCO_CIGARS	Whether the person uses tobacco cigars	char(1)	Y = Yes N = No X = Not asked U = Unknown or missing
TOBACCO_PIPES	Whether the person uses tobacco pipes	char(1)	Y = Yes N = No X = Not asked U = Unknown or missing
TOBACCO_SNUFF	Whether the person uses snuff tobacco	char(1)	Y = Yes N = No X = Not asked U = Unknown or missing

TOBACCO_UNSPECIFIED	Whether the person uses an unspecified form of tobacco	char(1)	Y = Yes N = No X = Not asked U = Unknown or missing
TOBACCO_COMMENT	Comment about the person's tobacco use	char(*)	Free text, missing values allowed

Primary Key:

No primary key defined

Foreign Key Relationship:

Source Variable (Foreign Key)	Target Table	Target Variable (Primary Key)	Orphans allowed?
MRN	DEMOGRAPHICS	MRN	No
ENC_ID	ENCOUNTER	ENC_ID	No

DEATH

Version = 5.0.0 Date = 11/16/2021 StdVar = &_vdw_death

Subject Area Description

The DEATH table contains one record per person in the DEMOGRAPHICS table for whom there is some belief that the person may be dead. When sources of death provide conflicting information, site data managers should make local determinations as to which source to use, collating information when possible, and reflecting their confidence in the observation overall using the CONFIDENCE variable.

Variable Name	Definition	Type(Len)	Values
MRN	Medical record number is the unique patient identifier within a site and should never leave the site	char(*)	Unique to each patient at each site
DEATHDT	The date that the person most likely died.	num(4)	SAS date, missing values allowed
DTIMPUTE	Indicates whether the death date is imputed and if so, how it was imputed.	char(1)	M = Month of date imputed D = Day of date imputed B = Both month & day imputed N = Date not imputed
SOURCE_LIST	A list of all sources of data that report this death regardless of any discrepencies in other variables (e.g. DEATHDT).	char(8)	A concatenated string of letters in descending order of reliability indicating all sources that report this person's death B = Social Security Admin N = National Death Index S = State Death records T = Tumor registry E = Encounter data P = Patient data M = Membership data O = Other
CONFIDENCE	Based on all information available, the level of confidence that this person is in fact dead.	char(1)	E = Excellent F = Fair P = Poor

Primary Key:

MRN

Foreign Key Relationship:

Source Variable (Foreign Key)	Target Table	Target Variable (Primary Key)	Orphans allowed?
MRN	DEMOGRAPHICS	MRN	No

CAUSE OF DEATH

Version = 5.0.0 Date = 11/16/2021 StdVar = &_vdw_cause_of_death

Subject Area Description

The CAUSE OF DEATH table contains one record per person per known cause of death.

Variable Name	Definition	Type(Len)	Values
MRN	Medical record number is the unique patient identifier within a site and should never leave the site	char(*)	Unique to each patient at each site
COD	The cause of death diagnosis mortality code. For ICD diagnosis coding, include decimal points in the value.	char(8)	
DX_CODETYPE	Refer to the DX_CODETYPE variable on the ENCOUNTER table for definition, type, length, and value set.		
CAUSETYPE	The type of cause of death.	char(1)	I = Immediate/Primary U = Underlying C = Contributory O = Other
SOURCE_LIST	Refer to the SOURCE_LIST variable on the DEATH table for definition, type, length, and value set.		
CONFIDENCE	Refer to the CONFIDENCE variable on the DEATH table for definition, type, length, and value set.		

Primary Key:

MRN + COD + DX_CODETYPE + CAUSETYPE

Foreign Key Relationship:

Source Variable (Foreign Key)	Target Table	Target Variable (Primary Key)	Orphans allowed?
MRN	DEMOGRAPHICS	MRN	No

TUMOR MAIN

Version = 5.1.0 Date = 4/21/2023 StdVar = &_vdw_tumor_main

Subject Area Description

The TUMOR content area contains records of documented neoplasms (typically malignant) diagnosed in patients as indicated in a tumor registry. In the Tumor Main table, there is one record per separately diagnosed neoplasm per reporting registry. Diagnoses of neoplasms in sources other than registries should not be included. Information in this table should be based entirely on information reported by registries (e.g., do not populate demographic variables with information from other sources).

The Tumor Main table contains a subset of the relevant tumor related information, namely those variables that are most likely to be important to the definition of a cohort. Additional information can be retrieved from the Tumor Supplemental table, which links to the Main table via tumorID.

The most current description as of 10/28/2022 is provided. It is always advisable to check http://datadictionary.naaccr.org/default.aspx?c=23 for the official descriptions of variables that correspond directly to NAACCR items as well as their value sets.

Variable Name	Definition	Type(Len)	Values
tumorID	Unique identifier of this tumor record	char(*)	Unique to each tumor record
mrn	Medical record number is the unique patient identifier within a site and should never leave the site	char(*)	Unique to each patient at each site
dataSource	Indicates the source of data for this record.	char(4)	For the first two positions use: LO = Local registry SE = Regional SEER registry ST = State operated registry The final two positions are locally defined to further differentiate source. Use values 'XX' for these positions if they are not needed at your site.

sequence	Central (preferred) or facility sequence number of tumor	char(2)	As defined by NAACCR items #560 (Sequence Number—Hospital) or #380 (Sequence Number—Central). 00=first and only, 01=first of several, 02-35 = 2nd-35th tumor, 60-87 = 1st-27th non-malignant tumor
dxDate	Date of diagnosis (numeric/date version of dateOfDiagnosis)	num(4)	numeric/date version of dateOfDiagnosis (NAACCR item #390)
dxDateImpute	Indicates whether parts of dxDate were imputed	char(1)	M = Month of date imputed D = Day of date imputed B = Both month & day imputed N = Date not imputed
dxYear	Year of diagnosis (from date of diagnosis)	num(4)	Four-digit year (which should correspond to year extracted from dxDate, if populated). Unknown values should be null (not 9999)
primarySite	Primary site	char(4)	As defined by NAACCR item #400. Valid codes are those listed by the WHO (excluding the decimal point) at http://codes.iarc.fr/topography.
seerSummaryStage	Best available SEER Summary Stage	char(1)	0 = In situ 1 = Localized 2 = Regional, direct extension only 3 = Regional, regional lymph nodes only 4 = Regional, direct extension and regional lymph nodes 5 = Regional, NOS 7 = Distant 8 = Not applicable 9 = Unstaged
histologicTypeIcdO3	Histologic type	char(4)	As defined by NAACCR item #522. Refer to https://www.naaccr.org/icdo3/ for the inventory of codes.

behavior	The data item Behavior Code describes the malignant potential of the tumor, ranging from /0 benign to /3 malignant (invasive).	char(1)	As defined by NAACCR item #523 (Behavior Code ICD-O-3; preferred) or #430 (Behavior [92-00] ICD-O-2). Valid values are: 0 = Benign (Reportable for intracranial and CNS sites only) 1 = Uncertain whether benign or malignant, borderline malignancy, low malignant potential, and uncertain malignant potential (Reportable for intracranial and CNS sites only) 2 = Carcinoma in situ; intraepithelial; noninfiltrating; non-invasive (carcinoma) 3 = Malignant, primary site (invasive) 6 = Metastatic site 9 = Unknown
stageAJatDx	AJCC stage group assessed at diagnosis	char(15)	Expected values: 0, 0A, 0IS, 1, 1A, 1A1, 1A2, 1B, 1B1, 1B2, 1C, 1S, 2, 2A, 2A1, 2A2, 2B, 2C, 2BULKY, 3, 3A, 3B, 3C, 3C1, 3C2, 4, 4A, 4A1, 4A2, 4B, 4C, OC, 88, 99, XX
stageAJpostTx	AJCC stage group assessed after neoadjuvant treatment	char(15)	Expected values: 0, 0A, 0IS, 1, 1A, 1A1, 1A2, 1B, 1B1, 1B2, 1C, 1S, 2, 2A, 2A1, 2A2, 2B, 2C, 2BULKY, 3, 3A, 3B, 3C, 3C1, 3C2, 4, 4A, 4A1, 4A2, 4B, 4C, OC, 88, 99, XX
gradeAtDx	Grade (differentiation) at diagnosis	char(1)	As defined by NAACCR items 440, 3843, and 3844: 1-9, A-E, L, H, M, S
gradePostTx	Grade (differentiation) after neoadjuvant treatment	char(1)	As defined by NAACCR items 1068 and 3845: 1-5, 8, 9, A-E, L, H, M, S
dxAge	Age at diagnosis	num(4)	Numeric version of NAACCR item #230. Unknown ages (stored as 999) should be left blank
birthDate	Date of birth	num(4)	numeric/date version of dateOfBirth (NAACCR item #220)

sex	Sex	char(1)	As defined by NAACCR item #220: 1 = Male 2 = Female 3 = Other (intersex, disorders of sexual development/DSD) 4 = Transsexual, NOS 5 = Transsexual, natal male 6 = Transsexual, natal female 9 = Not stated/Unknown
race1-race5	Race	char(2)	As defined by NAACCR item #160, valid values are 01–08, 10–17, 20–22, 25–28, 30–32, and 96–99. For race2-5, code 88 as well as null values are allowed as well.
spanishHispanicOrigin	Hispanic origin	char(1)	As defined by NAACCR item #190: 0 = Non-Spanish; non-Hispanic 1 = Mexican (includes Chicano) 2 = Puerto Rican 3 = Cuban 4 = South or Central American (except Brazil) 5 = Other specified Spanish/Hispanic origin (includes European; excludes Dominican Republic) 6 = Spanish, NOS Hispanic, NOS Latino, NOS 7 = Spanish surname only 8 = Dominican Republic 9 = Unknown whether Spanish or not
classOfCase	Class of case; defines involvement of facility in diagnosis and treatment of the tumor	char(2)	As defined by NAACCR item #610, valid values are 00, 10–14, 20–22, 30–38, 40–43, 49, and 99. To limit an analysis to so-called "analytic" cases, restrict classOfCase to values 00 through 22.
vitalStatus	Vital status of patient as of last contact	char(1)	As defined by NAACCR item #1760: 0 = Dead 1 = Alive

dateLastFollowUp	Date of last contact; identical to date of death if vital status=dead	num(4)	SAS date version of NAACCR item #1750 (dateOfLastContact)
laterality	Laterality	char(1)	As defined by NAACCR item #410: 0 = Not a paired site 1 = Right: origin of primary 2 = Left: origin of primary 3 = Only one side involved, right or left origin unspecified 4 = Bilateral involvement at time of diagnosis, lateral origin unknown for a single primary; or both ovaries involved simultaneously, single histology; bilateral retinoblastomas; bilateral Wilms' tumors 5 = Paired site: midline tumor 9 = Paired site, but no information concerning laterality
rxSummSurgery	Cancer-directed surgery code as part of first course of treatment	char(4)	Site specific codes; As defined by NAACCR items #1290,#1646,#1640 (00-99 codes) as well as #1291 (A000-A990 codes)
rxSummRadiation	Codes for the type of radiation therapy performed as part of the first course of treatment.	char(2)	As defined in NAACCR item #1360
rxSummChemo	Codes for chemotherapy given as part of the first course of treatment or the reason chemotherapy was not given. Includes treatment given at all facilities as part of the first course.	char(2)	As defined in NAACCR item #1390
rxSummHormone	Records whether systemic hormonal agents were administered as first-course treatment at any facility, or the reason they were not given	char(2)	As defined in NAACCR item #1400

rxSummBrm	Records whether immunotherapeutic (biologic response modifiers) agents were administered as first-course treatment at all facilities or the reason they were not given	char(2)	As defined in NAACCR item #1410
rxSummOther	Identifies other treatment given at all facilities that cannot be defined as surgery, radiation, or systemic therapy according to the defined data items in this manual	char(2)	As defined in NAACCR item #1420
rxSummTransplntEndocr	Identifies systemic therapeutic procedures administered as part of the first course of treatment at this and all other facilities	char(2)	As defined in NAACCR item #3250
rxSummPalliativeProc	Palliative Care at any facility. Identifies any care provided in an effort to palliate or alleviate symptoms	char(2)	As defined in NAACCR item #3270
recurrenceType1st	Recurrence type (of first recurrence)	char(2)	As defined in NAACCR item #1880
recurrenceDate	Date of first recurrence	num(4)	SAS date version of NAACCR item #1860 (recurrenceDate1st)
surgeryDate	Date of cancer-directed surgery (as part of 1st course of treatment)	num(4)	SAS date version of NAACCR item #1200 (rxDateSurgery)
chemoDate	Date of chemotherapy start (as part of 1st course of treatment)	num(4)	SAS date version of NAACCR item #1220 (rxDateChemo)
radiationDate	Date of radiation therapy start (as part of 1st course of treatment)	num(4)	SAS date version of NAACCR item #1210 (rxDateRadiation)
hormoneDate	Date of hormone therapy start (as part of 1st course of treatment)	num(4)	SAS date version of NAACCR item #1230 (rxDateHormone)
brmDate	Date of immunotherapy (biological response modifier therapy) start (as part of 1st course of treatment)	num(4)	SAS date version of NAACCR item #1240 (rxDateBrm)
otherDate	Date of other cancer therapy start (as part of 1st course of treatment)	num(4)	SAS date version of NAACCR item #1250 (rxDateOther)
systemicDate	Date of systemic therapy start (as part of 1st course of treatment)	num(4)	SAS date version of NAACCR item #3230 (rxDateSystemic)

transplntEndocrDate	Date of Hematologic Transplant and Endocrine Procedure (as part of 1st course of treatment)	num(4)	valid SAS date

Primary Key:

TumorID

Foreign Key Relationship:

Source Variable (Foreign Key)	Target Table	Target Variable (Primary Key)	Orphans Allowed?
MRN	DEMOGRAPHICS	MRN	No

TUMOR SUPPLEMENTAL

Version = 5.1.0 Date = 4/21/2023 StdVar = &_vdw_tumor_supp

Subject Area Description

The TUMOR content area contains records of documented neoplasms (typically malignant) diagnosed in patients as indicated in a tumor registry. One Tumor ID corresponds one diagnosed neoplasm in one reporting registry. (see Tumor Main for more information)

The Tumor Supplemental table has one record per NAACCR item per tumorID. It contains detailed tumor information on staging, prognosis, treatment and outcomes.

There should be no missing values. If information is not available, there should be no record in this table.

The most current description as of 10/28/2022 is provided, it's always advised to check http://datadictionary.naaccr.org/default.aspx?c=23 for the official descriptions of variables that correspond directly to NAACCR items.

Variable Name	Definition	Type(Len)	Values
tumorID	Unique identifier of the Tumor record	char(*)	
tumorAttribute	Attribute of the tumor record for which information is provided on this line	char(32)	For NAACCR items, the NAACCR XML tag; other, locally defined attributes can be added, but their name must not be a NAACCR XML tag
tumorValue	The value corresponding to the attribute	char(30)	as defined by NAACCR depending on the attribute (NAACCR item); must not be null

Primary Key:

TumorID + TumorAttribute

Source Variable (Foreign Key)	Target Table	Target Variable (Primary Key)	Orphans allowed?
TumorID	TumorMain	TumorID	No

Version = 5.1.0 Date = 4/21/2023

EAV Attributes to be populated (to the extent possible)

NAACCR

Item #	TumorAttribute (NAACCR XML Tag)	Description	Comment
80	addrAtDxState	Address at Dx--State	
89	countyAtDxAnalysis	County at Dx Analysis	
90	countyAtDx	County at Dx	
95	countyAtDxGeocode2000	County at DX Geocode2000	
96	countyAtDxGeocode2010	County at DX Geocode2010	
97	countyAtDxGeocode2020	County at DX Geocode2020	
191	nhiaDerivedHispOrigin	NHIA Derived Hispanic Origin	
230	ageAtDiagnosis	Age at Diagnosis	
240	dateOfBirth	Date of Birth	
344	tobaccoUseSmokingStatus	Tobacco Use Smoking Status	
364	censusTrCert19708090	Census Tr Cert 1970/80/90	
365	censusTrCertainty2000	Census Tr Certainty 2000	
367	censusTrCertainty2010	Census Tr Certainty 2010	
369	censusTractCertainty2020	Census Tract Certainty 2020	
380	sequenceNumberCentral	Sequence Number--Central	
390	dateOfDiagnosis	Date of Diagnosis	
420	histologyIcdO2	Histology (92-00) ICD-O-2	
430	behaviorIcdO2	Behavior (92-00) ICD-O-2	
440	grade	Grade	
490	diagnosticConfirmation	Diagnostic Confirmation	
550	accessionNumberHosp	Accession Number--Hosp	
560	sequenceNumberHospital	Sequence Number--Hospital	
670	rxHospSurgPrimSite	RX Hosp--Surg Prim Site	
682	dateRegionalLNDissection	Date Regional Lymph Node Dissection	
690	rxHospRadiation	RX Hosp--Radiation	
700	rxHospChemo	RX Hosp--Chemo	
710	rxHospHormone	RX Hosp--Hormone	

720	rxHospBrm	RX Hosp--BRM
730	rxHospOther	RX Hosp--Other
746	rxHospSurgPrimSite2023	RX Hosp--Surg Prim Site 2023
752	tumorSizeClinical	Tumor Size Clinical
754	tumorSizePathologic	Tumor Size Pathologic
756	tumorSizeSummary	Tumor Size Summary
759	seerSummaryStage2000	SEER Summary Stage 2000
760	seerSummaryStage1977	SEER Summary Stage 1977
762	derivedSummaryStage2018	Derived Summary Stage 2018
764	summaryStage2018	Summary Stage 2018
772	eodPrimaryTumor	EOD Primary Tumor
774	eodRegionalNodes	EOD Regional Nodes
776	eodMets	EOD Mets
779	extentOfDisease10Dig	Extent of Disease 10-Dig
780	eodTumorSize	EOD--Tumor Size
785	derivedEod2018T	Derived EOD 2018 T
790	eodExtension	EOD--Extension
795	derivedEod2018M	Derived EOD 2018 M
800	eodExtensionProstPath	EOD--Extension Prost Path
810	eodLymphNodeInvolv	EOD--Lymph Node Involv
815	derivedEod2018N	Derived EOD 2018 N
818	derivedEod2018StageGroup	Derived EOD 2018 Stage Group
820	regionalNodesPositive	Regional Nodes Positive
830	regionalNodesExamined	Regional Nodes Examined
832	dateSentinelLymphNodeBiopsy	Date of Sentinel Lymph Node Biopsy
834	sentinelLymphNodesExamined	Sentinel Lymph Nodes Examined
835	sentinelLymphNodesPositive	Sentinel Lymph Nodes Positive
840	eodOld13Digit	EOD--Old 13 Digit
850	eodOld2Digit	EOD--Old 2 Digit
860	eodOld4Digit	EOD--Old 4 Digit
870	codingSystemForEod	Coding System for EOD
880	tnmPathT	TNM Path T
890	tnmPathN	TNM Path N
900	tnmPathM	TNM Path M

910	tnmPathStageGroup	TNM Path Stage Group
920	tnmPathDescriptor	TNM Path Descriptor
940	tnmClinT	TNM Clin T
950	tnmClinN	TNM Clin N
960	tnmClinM	TNM Clin M
970	tnmClinStageGroup	TNM Clin Stage Group
980	tnmClinDescriptor	TNM Clin Descriptor
995	ajccId	AJCC ID
1001	ajccTnmClinT	AJCC TNM Clin T
1002	ajccTnmClinN	AJCC TNM Clin N
1003	ajccTnmClinM	AJCC TNM Clin M
1004	ajccTnmClinStageGroup	AJCC TNM Clin Stage Group
1011	ajccTnmPathT	AJCC TNM Path T
1012	ajccTnmPathN	AJCC TNM Path N
1013	ajccTnmPathM	AJCC TNM Path M
1014	ajccTnmPathStageGroup	AJCC TNM Path Stage Group
1021	ajccTnmPostTherapyT	AJCC TNM Post Therapy Path (yp) T
1022	ajccTnmPostTherapyN	AJCC TNM Post Therapy Path (yp) N
1023	ajccTnmPostTherapyM	AJCC TNM Post Therapy Path (yp) M
1024	ajccTnmPostTherapyStageGroup	AJCC TNM Post Therapy Path (yp) Stage Group
1031	ajccTnmClinTSuffix	AJCC TNM Clin T Suffix
1032	ajccTnmPathTSuffix	AJCC TNM Path T Suffix
1033	ajccTnmPostTherapyTSuffix	AJCC TNM Post Therapy Path (yp) T Suffix
1034	ajccTnmClinNSuffix	AJCC TNM Clin N Suffix
1035	ajccTnmPathNSuffix	AJCC TNM Path N Suffix
1036	ajccTnmPostTherapyNSuffix	AJCC TNM Post Therapy Path (yp) N Suffix
1060	tnmEditionNumber	TNM Edition Number
1062	ajccTnmPostTherapyClinT	AJCC TNM Post Therapy Clin (yc) T
1063	ajccTnmPostTherapyClinTSuffix	AJCC TNM Post Therapy Clin (yc) T Suffix
1064	ajccTnmPostTherapyClinN	AJCC TNM Post Therapy Clin (yc) N
1065	ajccTnmPostTherapyClinNSuffix	AJCC TNM Post Therapy Clin (yc) N Suffix
1066	ajccTnmPostTherapyClinM	AJCC TNM Post Therapy Clin (yc) M
1067	ajccTnmPostTherapyClinStageGrp	AJCC TNM Post Therapy Clin (yc) Stage Group
1068	gradePostTherapyClin	Grade Post Therapy Clin (yc)

1112	metsAtDxBone	Mets at DX-Bone
1113	metsAtDxBrain	Mets at DX-Brain
1114	metsAtDxDistantLn	Mets at Dx-Distant LN
1115	metsAtDxLiver	Mets at DX-Liver
1116	metsAtDxLung	Mets at DX-Lung
1117	metsAtDxOther	Mets at DX-Other
1120	pediatricStage	Pediatric Stage
1130	pediatricStagingSystem	Pediatric Staging System
1150	tumorMarker1	Tumor Marker 1
1160	tumorMarker2	Tumor Marker 2
1170	tumorMarker3	Tumor Marker 3
1182	lymphVascularInvasion	Lymphovascular Invasion
1200	rxDateSurgery	RX Date Surgery
1210	rxDateRadiation	RX Date Radiation
1220	rxDateChemo	RX Date Chemo
1230	rxDateHormone	RX Date Hormone
1240	rxDateBrm	RX Date BRM
1250	rxDateOther	RX Date Other
1260	dateInitialRxSeer	Date Initial RX SEER
1270	date1stCrsRxCoc	Date 1st Crs RX CoC
1280	rxDateDxStgProc	RX Date DX/Stg Proc
1285	rxSummTreatmentStatus	RX Summ--Treatment Status
1290	rxSummSurgPrimSite	RX Summ--Surg Prim Site
1291	rxSummSurgPrimSite2023	RX Summ--Surg Prim Site 2023
1292	rxSummScopeRegLnSur	RX Summ--Scope Reg LN Sur
1294	rxSummSurgOthRegDis	RX Summ--Surg Oth Reg/Dis
1296	rxSummRegLnExamined	RX Summ--Reg LN Examined
1310	rxSummSurgicalApproch	RX Summ--Surgical Approch
1320	rxSummSurgicalMargins	RX Summ--Surgical Margins
1330	rxSummReconstruct1st	RX Summ--Reconstruct 1st
1340	reasonForNoSurgery	Reason for No Surgery
1370	rxSummRadToCns	RX Summ--Rad to CNS
1380	rxSummSurgRadSeq	RX Summ--Surg/Rad Seq
1430	reasonForNoRadiation	Reason for No Radiation

1440	reasonForNoChemo	Reason for no Chemo
1450	reasonforNoHormone	Reason for no Hormone
1501	phase1DosePerFraction	Phase I Dose per Fraction
1502	phase1RadiationExternalBeamTech	Phase I Radiation External Beam Planning Tech
1503	phase1NumberOfFractions	Phase I Number of Fractions
1504	phase1RadiationPrimaryTxVolume	Phase I Radiation Primary Treatment Volume
1505	phase1RadiationToDrainingLN	Phase I Radiation to Draining Lymph Nodes
1506	phase1RadiationTreatmentModality	Phase I Radiation Treatment Modality
1507	phase1TotalDose	Phase I Total Dose
1510	radRegionalDoseCgy	Rad--Regional Dose: cGy
1511	phase2DosePerFraction	Phase II Dose per Fraction
1512	phase2RadiationExternalBeamTech	Phase II Radiation External Beam Planning Tech
1513	phase2NumberOfFractions	Phase II Number of Fractions
1514	phase2RadiationPrimaryTxVolume	Phase II Radiation Primary Treatment Volume
1515	phase2RadiationToDrainingLN	Phase II Radiation to Draining Lymph Nodes
1516	phase2RadiationTreatmentModality	Phase II Radiation Treatment Modality
1517	phase2TotalDose	Phase II Total Dose
1520	radNoOfTreatmentVol	Rad--No of Treatment Vol
1521	phase3DosePerFraction	Phase III Dose per Fraction
1522	phase3RadiationExternalBeamTech	Phase III Radiation External Beam Planning Tech
1523	phase3NumberOfFractions	Phase III Number of Fractions
1524	phase3RadiationPrimaryTxVolume	Phase III Radiation Primary Treatment Volume
1525	phase3RadiationToDrainingLN	Phase III Radiation to Draining Lymph Nodes
1526	phase3RadiationTreatmentModality	Phase III Radiation Treatment Modality
1527	phase3TotalDose	Phase III Total Dose
1531	radiationTxDiscontinuedEarly	Radiation Treatment Discontinued Early
1532	numberPhasesOfRadTxToVolume	Number of Phases of Rad Treatment to this Volume
1533	totalDose	Total Dose
1540	radTreatmentVolume	Rad--Treatment Volume
1550	radLocationOfRx	Rad--Location of RX
1570	radRegionalRxModality	Rad--Regional RX Modality
1632	neoadjuvantTherapy	Neoadjuvant Therapy
1633	neoadjuvTherapyClinicalResponse	Neoadjuvant Therapy-Clinical Response
1634	neoadjuvTherapyTreatmentEffect	Neoadjuvant Therapy-Treatment Effect

1639	rxSummSystemicSurSeq	RX Summ--Systemic/Sur Seq
1640	rxSummSurgeryType	RX Summ--Surgery Type
1646	rxSummSurgSite9802	RX Summ--Surg Site 98-02
1647	rxSummScopeReg9802	RX Summ--Scope Reg 98-02
1648	rxSummSurgOth9802	RX Summ--Surg Oth 98-02
1750	dateOfLastContact	Date of Last Contact
1770	cancerStatus	Cancer Status
1772	dateOfLastCancerStatus	Date of Last Cancer (tumor) Status
1860	recurrenceDate1st	Recurrence Date--1st
1861	recurrenceDate1stFlag	Recurrence Date--1st Flag
1880	recurrenceType1st	Recurrence Type--1st
1910	causeOfDeath	Cause of Death
1914	seerCauseSpecificCod	SEER Cause Specific COD
1915	seerOtherCod	SEER Other COD
1960	siteIcdO1	Site (73-91) ICD-O-1
1970	morphIcdO1	Morph (73-91) ICD-O-1
1971	histologyIcdO1	Histology (73-91) ICD-O-1
1972	behaviorIcdO1	Behavior (73-91) ICD-O-1
1973	gradeIcdO1	Grade (73-91) ICD-O-1
2800	csTumorSize	CS Tumor Size
2810	csExtension	CS Extension
2820	csTumorSizeExtEval	CS Tumor Size/Ext Eval
2830	csLymphNodes	CS Lymph Nodes
2840	csLymphNodesEval	CS Lymph Nodes Eval
2850	csMetsAtDx	CS Mets at DX
2851	csMetsAtDxBone	CS Mets at Dx-Bone
2852	csMetsAtDxBrain	CS Mets at Dx-Brain
2853	csMetsAtDxLiver	CS Mets at Dx-Liver
2854	csMetsAtDxLung	CS Mets at Dx-Lung
2860	csMetsEval	CS Mets Eval
2861	csSiteSpecificFactor7	CS Site-Specific Factor 7
2862	csSiteSpecificFactor8	CS Site-Specific Factor 8
2863	csSiteSpecificFactor9	CS Site-Specific Factor 9
2864	csSiteSpecificFactor10	CS Site-Specific Factor10

2865	csSiteSpecificFactor11	CS Site-Specific Factor11
2866	csSiteSpecificFactor12	CS Site-Specific Factor12
2867	csSiteSpecificFactor13	CS Site-Specific Factor13
2868	csSiteSpecificFactor14	CS Site-Specific Factor14
2869	csSiteSpecificFactor15	CS Site-Specific Factor15
2870	csSiteSpecificFactor16	CS Site-Specific Factor16
2871	csSiteSpecificFactor17	CS Site-Specific Factor17
2872	csSiteSpecificFactor18	CS Site-Specific Factor18
2873	csSiteSpecificFactor19	CS Site-Specific Factor19
2874	csSiteSpecificFactor20	CS Site-Specific Factor20
2875	csSiteSpecificFactor21	CS Site-Specific Factor21
2876	csSiteSpecificFactor22	CS Site-Specific Factor22
2877	csSiteSpecificFactor23	CS Site-Specific Factor23
2878	csSiteSpecificFactor24	CS Site-Specific Factor24
2879	csSiteSpecificFactor25	CS Site-Specific Factor25
2880	csSiteSpecificFactor1	CS Site-Specific Factor 1
2890	csSiteSpecificFactor2	CS Site-Specific Factor 2
2900	csSiteSpecificFactor3	CS Site-Specific Factor 3
2910	csSiteSpecificFactor4	CS Site-Specific Factor 4
2920	csSiteSpecificFactor5	CS Site-Specific Factor 5
2930	csSiteSpecificFactor6	CS Site-Specific Factor 6
2935	csVersionInputOriginal	CS Version Input Original
2936	csVersionDerived	CS Version Derived
2937	csVersionInputCurrent	CS Version Input Current
2940	derivedAjcc6T	Derived AJCC-6 T
2950	derivedAjcc6TDescript	Derived AJCC-6 T Descript
2960	derivedAjcc6N	Derived AJCC-6 N
2970	derivedAjcc6NDescript	Derived AJCC-6 N Descript
2980	derivedAjcc6M	Derived AJCC-6 M
2990	derivedAjcc6MDescript	Derived AJCC-6 M Descript
3000	derivedAjcc6StageGrp	Derived AJCC-6 Stage Grp
3010	derivedSs1977	Derived SS1977
3020	derivedSs2000	Derived SS2000
3030	derivedAjccFlag	Derived AJCC--Flag

3040	derivedSs1977Flag	Derived SS1977--Flag
3050	derivedSs2000Flag	Derived SS2000--Flag
3200	radBoostRxModality	Rad--Boost RX Modality
3210	radBoostDoseCgy	Rad--Boost Dose cGy
3220	rxDateRadiationEnded	RX Date Rad Ended
3230	rxDateSystemic	RX Date Systemic
3280	rxHospPalliativeProc	RX Hosp--Palliative Proc
3400	derivedAjcc7T	Derived AJCC-7 T
3402	derivedAjcc7TDescript	Derived AJCC-7 T Descript
3410	derivedAjcc7N	Derived AJCC-7 N
3412	derivedAjcc7NDescript	Derived AJCC-7 N Descript
3420	derivedAjcc7M	Derived AJCC-7 M
3422	derivedAjcc7MDescript	Derived AJCC-7 M Descript
3430	derivedAjcc7StageGrp	Derived AJCC-7 Stage Grp
3440	derivedPrerx7T	Derived PreRx-7 T
3442	derivedPrerx7TDescrip	Derived PreRx-7 T Descrip
3450	derivedPrerx7N	Derived PreRx-7 N
3452	derivedPrerx7NDescrip	Derived PreRx-7 N Descrip
3460	derivedPrerx7M	Derived PreRx-7 M
3462	derivedPrerx7MDescrip	Derived PreRx-7 M Descrip
3470	derivedPrerx7StageGrp	Derived PreRx-7 Stage Grp
3480	derivedPostrx7T	Derived PostRx-7 T
3482	derivedPostrx7N	Derived PostRx-7 N
3490	derivedPostrx7M	Derived PostRx-7 M
3492	derivedPostrx7StgeGrp	Derived PostRx-7 Stge Grp
3600	derivedNeoadjuvRxFlag	Derived Neoadjuv Rx Flag
3605	derivedSeerPathStgGrp	Derived SEER Path Stg Grp
3610	derivedSeerClinStgGrp	Derived SEER Clin Stg Grp
3614	derivedSeerCmbStgGrp	Derived SEER Cmb Stg Grp
3616	derivedSeerCombinedT	Derived SEER Combined T
3618	derivedSeerCombinedN	Derived SEER Combined N
3620	derivedSeerCombinedM	Derived SEER Combined M
3622	derivedSeerCmbTSrc	Derived SEER Cmb T Src
3624	derivedSeerCmbNSrc	Derived SEER Cmb N Src

3626	derivedSeerCmbMSrc	Derived SEER Cmb M Src
3645	npcrDerivedAjcc8TnmClinStgGrp	NPCR Derived AJCC 8 TNM Clin Stg Grp
3646	npcrDerivedAjcc8TnmPathStgGrp	NPCR Derived AJCC 8 TNM Path Stg Grp
3647	npcrDerivedAjcc8TnmPostStgGrp	NPCR Derived AJCC 8 TNM Post Therapy Stg Grp
3650	npcrDerivedClinStgGrp	NPCR Derived Clin Stg Grp
3655	npcrDerivedPathStgGrp	NPCR Derived Path Stg Grp
3801	chromosome1pLossHeterozygosity	Chromosome 1p: Loss of Heterozygosity (LOH)
3802	chromosome19qLossHeterozygosity	Chromosome 19q: Loss of Heterozygosity (LOH)
3803	adenoidCysticBasaloidPattern	adenoidCysticBasaloidPattern
3804	adenopathy	adenopathy
3805	afpPostOrchiectomyLabValue	afpPostOrchiectomyLabValue
3806	afpPostOrchiectomyRange	afpPostOrchiectomyRange
3807	afpPreOrchiectomyLabValue	afpPreOrchiectomyLabValue
3808	afpPreOrchiectomyRange	afpPreOrchiectomyRange
3809	afpPretreatmentInterpretation	afpPretreatmentInterpretation
3810	afpPretreatmentLabValue	afpPretreatmentLabValue
3811	anemia	Anemia
3812	bSymptoms	B symptoms
3813	bilirubinPretxTotalLabValue	Bilirubin Pretreatment Total Lab Value
3814	bilirubinPretxUnitOfMeasure	Bilirubin Pretreatment Unit of Measure
3815	boneInvasion	Bone Invasion
3816	brainMolecularMarkers	Brain Molecular Markers
3817	breslowTumorThickness	Breslow Tumor Thickness
3818	ca125PretreatmentInterpretation	CA-125 Pretreatment Interpretation
3819	ceaPretreatmentInterpretation	CEA Pretreatment Interpretation
3820	ceaPretreatmentLabValue	CEA Pretreatment Lab Value
3821	chromosome3Status	Chromosome 3 Status
3822	chromosome8qStatus	Chromosome 8q Status
3823	circumferentialResectionMargin	Circumferential Resection Margin (CRM)
3824	creatininePretreatmentLabValue	Creatinine Pretreatment Lab Value
3825	creatininePretxUnitOfMeasure	Creatinine Pretreatment Unit of Measure
3826	estrogenReceptorPercntPosOrRange	Estrogen Receptor Percent Positive or Range
3827	estrogenReceptorSummary	Estrogen Receptor Summary
3828	estrogenReceptorTotalAllredScore	Estrogen Receptor Total Allred Score

3829	esophagusAndEgjTumorEpicenter	Esophagus and EGJ Tumor Epicenter
3830	extranodalExtensionClin	Extranodal Extension Clin (non-Head and Neck)
3831	extranodalExtensionHeadNeckClin	Extranodal Extension Head and Neck Clinical
3832	extranodalExtensionHeadNeckPath	Extranodal Extension Head and Neck Pathological
3833	extranodalExtensionPath	Extranodal Extension Path (non-Head and Neck)
3834	extravascularMatrixPatterns	Extravascular Matrix Patterns
3835	fibrosisScore	Fibrosis Score
3836	figoStage	FIGO Stage
3837	gestationalTrophoblasticPxIndex	Gestational Trophoblastic Prognostic Scoring Index
3838	gleasonPatternsClinical	Gleason Patterns Clinical
3839	gleasonPatternsPathological	Gleason Patterns Pathological
3840	gleasonScoreClinical	Gleason Score Clinical
3841	gleasonScorePathological	Gleason Score Pathological
3842	gleasonTertiaryPattern	Gleason Tertiary Pattern
3843	gradeClinical	Grade Clinical
3844	gradePathological	Grade Pathological
3845	gradePostTherapy	Grade Post Therapy Path (yp)
3846	hcgPostOrchiectomyLabValue	hCG Post-Orchiectomy Lab Value
3847	hcgPostOrchiectomyRange	hCG Post-Orchiectomy Range
3848	hcgPreOrchiectomyLabValue	hCG Pre-Orchiectomy Lab Value
3849	hcgPreOrchiectomyRange	hCG Pre-Orchiectomy Range
3850	her2IhcSummary	HER2 IHC Summary
3851	her2IshDualProbeCopyNumber	HER2 ISH Dual Probe Copy Number
3852	her2IshDualProbeRatio	HER2 ISH Dual Probe Ratio
3853	her2IshSingleProbeCopyNumber	HER2 ISH Single Probe Copy Number
3854	her2IshSummary	HER2 ISH Summary
3855	her2OverallSummary	HER2 Overall Summary
3856	heritableTrait	Heritable Trait
3857	highRiskCytogenetics	High Risk Cytogenetics
3858	highRiskHistologicFeatures	High Risk Histologic Features
3859	hivStatus	HIV Status
3860	iNRProthrombinTime	International Normalized Ratio Prothrombin Time
3861	ipsilateralAdrenalGlandInvolve	Ipsilateral Adrenal Gland Involvement
3862	jak2	JAK2

3863	ki67	Ki-67
3864	invasionBeyondCapsule	Invasion Beyond Capsule
3865	kitGeneImmunohistochemistry	KIT Gene Immunohistochemistry
3866	kras	KRAS
3867	ldhPostOrchiectomyRange	LDH Post-Orchiectomy Range
3868	ldhPreOrchiectomyRange	LDH Pre-Orchiectomy Range
3869	ldhPretreatmentLevel	LDH Level
3870	ldhUpperLimitsOfNormal	LDH Upper Limits of Normal
3871	lnAssessMethodFemoralInguinal	LN Assessment Method Femoral-Inguinal
3872	lnAssessMethodParaaortic	LN Assessment Method Para-Aortic
3873	lnAssessMethodPelvic	LN Assessment Method Pelvic
3874	lnDistantAssessMethod	LN Distant Assessment Method
3875	lnDistantMediastinalScalene	LN Distant: Mediastinal, Scalene
3876	lnHeadAndNeckLevels1To3	LN Head and Neck Levels I-III
3877	lnHeadAndNeckLevels4To5	LN Head and Neck Levels IV-V
3878	lnHeadAndNeckLevels6To7	LN Head and Neck Levels VI-VII
3879	lnHeadAndNeckOther	LN Head and Neck Other
3880	lnIsolatedTumorCells	LN Isolated Tumor Cells (ITC)
3881	lnLaterality	LN Laterality
3882	lnPositiveAxillaryLevel1To2	LN Positive Axillary Level I-II
3883	lnSize	LN Size
3884	lnStatusFemorInguinParaaortPelv	LN Status Femoral-Inguinal, Para-Aortic, Pelvic
3885	lymphocytosis	Lymphocytosis
3886	majorVeinInvolvement	Major Vein Involvement
3887	measuredBasalDiameter	Measured Basal Diameter
3888	measuredThickness	Measured Thickness
3889	methylationOfO6MGMT	Methylation of O6-Methylguanine-Methyltransferase
3890	microsatelliteInstability	Microsatellite Instability (MSI)
3891	microvascularDensity	Microvascular Density
3892	mitoticCountUvealMelanoma	Mitotic Count Uveal Melanoma
3893	mitoticRateMelanoma	Mitotic Rate Melanoma
3894	multigeneSignatureMethod	Multigene Signature Method
3895	multigeneSignatureResults	Multigene Signature Results
3896	nccnInternationalPrognosticIndex	NCCN International Prognostic Index (IPI)

3897	numberOfCoresExamined	Number of Cores Examined
3898	numberOfCoresPositive	Number of Cores Positive
3899	numberOfExaminedParaAorticNodes	Number of Examined Para-Aortic Nodes
3900	numberOfExaminedPelvicNodes	Number of Examined Pelvic Nodes
3901	numberOfPositiveParaAorticNodes	Number of Positive Para-Aortic Nodes
3902	numberOfPositivePelvicNodes	Number of Positive Pelvic Nodes
3903	oncotypeDxRecurrenceScoreDcis	Oncotype Dx Recurrence Score-DCIS
3904	oncotypeDxRecurrenceScoreInvasiv	Oncotype Dx Recurrence Score-Invasive
3905	oncotypeDxRiskLevelDcis	Oncotype Dx Risk Level-DCIS
3906	oncotypeDxRiskLevelInvasive	Oncotype Dx Risk Level-Invasive
3907	organomegaly	Organomegaly
3908	percentNecrosisPostNeoadjuvant	Percent Necrosis Post Neoadjuvant
3909	perineuralInvasion	Perineural Invasion
3910	peripheralBloodInvolvement	Peripheral Blood Involvement
3911	peritonealCytology	Peritoneal Cytology
3913	pleuralEffusion	Pleural Effusion
3914	progesteroneRecepPrcntPosOrRange	Progesterone Receptor Percent Positive or Range
3915	progesteroneRecepSummary	Progesterone Receptor Summary
3916	progesteroneRecepTotalAllredScor	Progesterone Receptor Total Allred Score
3917	primarySclerosingCholangitis	Primary Sclerosing Cholangitis
3918	profoundImmuneSuppression	Profound Immune Suppression
3919	prostatePathologicalExtension	EOD Prostate Pathologic Extension
3920	psaLabValue	PSA (Prostatic Specific Antigen) Lab Value
3921	residualTumVolPostCytoreduction	Residual Tumor Volume Post Cytoreduction
3922	responseToNeoadjuvantTherapy	Response to Neoadjuvant Therapy
3923	sCategoryClinical	S Category Clinical
3924	sCategoryPathological	S Category Pathological
3925	sarcomatoidFeatures	Sarcomatoid Features
3929	separateTumorNodules	Separate Tumor Nodules
3930	serumAlbuminPretreatmentLevel	Serum Albumin Pretreatment Level
3931	serumBeta2MicroglobulinPretxLvl	Serum Beta-2 Microglobulin Pretreatment Level
3932	ldhPretreatmentLabValue	LDH Lab Value
3933	thrombocytopenia	Thrombocytopenia
3934	tumorDeposits	Tumor Deposits

3935	tumorGrowthPattern	Tumor Growth Pattern	
3936	ulceration	Ulceration	
3937	visceralParietalPleuralInvasion	Visceral and Parietal Pleural Invasion	
3938	alkRearrangement	ALK Rearrangement	
3939	egfrMutationalAnalysis	EGFR Mutational Analysis	
3940	brafMutationalAnalysis	BRAF Mutational Analysis	
3941	nrasMutationalAnalysis	NRAS Mutational Analysis	
3942	ca199PretxLabValue	CA 19-9 PreTX Lab Value	
3950	macroscopicEvalOfTheMesorectum	Macroscopic Evaluation of Mesorectum	
3956	p16	p16	
3957	lnStatusPelvic	LN Status Pelvic	
3958	lnStatusParaAortic	LN Status Para-Aortic	
3959	lnStatusFemoralInguinal	LN Status Femoral-Inguinal	
N/A	stageAJ	Best AJCC Stage (ignoring complications of neoadjuvant treatment)	constructed from multiple NAACCR items; historical only
N/A	reasonForNoBrm	Reason no Immunotherapy was administered	no NAACCR item; historical only
N/A	reasonForNoOther	Reason no other therapy was administered	no NAACCR item; historical only
N/A	reasonForNoTransplntEndocr	Reason hematologic transplant or endocrine procedure was	no NAACCR item; historical only

CENSUS LOCATION

Version = 5.1.1 Date = 5/10/2022 StdVar = &_vdw_census_loc

Subject Area Description

The CENSUS LOCATION table contains geographic location markers for patient residencies over time.
Projects are encouraged to contact sites regarding the version of Census data available.

Variable Name	Definition	Type(Len)	Values
MRN	Medical record number is the unique patient identifier within a site and should never leave the site	char(*)	Unique to each patient at each site
LOC_START	The date on which the person's tenure at this location began.	num(4)	SAS date
LOC_END	The date on which the person's tenure at this location ended.	num(4)	SAS date
GEOCODE	The concatenation of FIPS codes for State, County, and Census Tract.	char(11)	FIPS values are set by the National Institute of Standards & Technology and are public information
GEOCODE_BOUNDARY_YEAR	The decade of the boundary files used to perform geocoding.	num(8)	Year values (e.g. 2010, 2020)
GEOLEVEL	The most granular geographic level of the GEOCODE indicating the specificity of the match made. Also called 'map flag'.	char(1)	T = Tract C = County Z = Zip code U = Unable to be appended P = Address is post office

MATCH_STRENGTH	A code indicating the type of match achieved by the geocoding software.	char(1)	0 = No coordinates 1 = Zip +0 centroid 2 = Zip +2 centroid 3 = Zip +4 centroid 4 = Shape path centroid 5 = Street address position 6 = Point zip centroid X = Street intersection
LATITUDE	The latitude of the location	num(8)	Value between -90 and +90 measured in degrees
LONGITUDE	The longitude of the location	num(8)	Value between -180 and +180 measured in degrees
ZIP	Zip Code--a five-character numeric code assigned by the US Postal Service to various regions where it delivers mail.	char(5)	Any valid zip code, as assigned by the US Postal Service
GEOCODE_APP	The name of the application used to geocode this location	char(*)	Free text field

Primary Key:

MRN + GEOCODE_BOUNDARY_YEAR + LOC_START (no overlapping time periods are permitted)

Foreign Key Relationship:

Source Variable (Foreign Key)	Table Target	Target Variable (Primary Key)	Orphans allowed?
MRN	DEMOGRAPHICS	MRN	No
GEOCODE + GEOCODE_BOUNDARY_YEAR	CENSUS ACS DEMOGRAPHICS	GEOCODE + GEOCODE_BOUNDARY_YEAR	No
GEOCODE + GEOCODE_BOUNDARY_YEAR	CENSUS DECENNIAL DEMOGRAPHICS	GEOCODE + CENSUS_YEAR	No

CENSUS ACS DEMOGRAPHICS

Version = 5.1.0 Date = 2/21/2022 StdVar = &_vdw_census_demog_acs

Subject Area Description

The CENSUS ACS DEMOGRAPHICS table contains data from the American Community Survey (ACS) administered by the US Census Bureau. The ACS is an ongoing survey with an annual release covering a 5-year lookback period; this table includes data at the census tract level. It is recommended that sites implement the full ACS data file released by the HCSRN.
Projects are encouraged to contact sites regarding the version of Census data available.

Variable Name	Definition	Type(Len)	Values
CENSUS_YEAR	Year the ACS data was released by the Census Bureau	Num(8)	Year values (e.g. 2016, 2017)
GEOCODE	Concatenation of the FIPS codes for State, County, and Census Tract	Char(11)	The first two characters signify the state, the next three signify the county, the next six signify the tract.
GEOCODE_BOUNDARY_YEAR	The decade of the boundary year used to create the census tracts of the ACS data.	num(8)	Year values (e.g. 2010, 2020)
STATE	State Code--a two-character numeric code assigned to US states, districts, territories and protectorates.	char(2)	Any valid state code, as used by the USCB.
COUNTY	County Code--a 3-character numeric code assigned to census counties.	char(3)	Any valid county code, as used by the USCB.
TRACT	Tract Code--a six-character numeric code assigned to census tracts.	char(6)	Any valid tract code, as used by the USCB.
EDUCATION1	Less than 9th grade	num(8)	Any proportion between 0 and 1.
EDUCATION2	9th - 12th grade	num(8)	Any proportion between 0 and 1.
EDUCATION3	high school graduate	num(8)	Any proportion between 0 and 1.

EDUCATION4	some college, no degree	num(8)	Any proportion between 0 and 1.
EDUCATION5	associate degree	num(8)	Any proportion between 0 and 1.
EDUCATION6	bachelor degree	num(8)	Any proportion between 0 and 1.
EDUCATION7	graduate or professional degree	num(8)	Any proportion between 0 and 1.
EDUCATION8	Doctorate degree	num(8)	Any proportion between 0 and 1.
MEDFAMINCOME	Median Family Income	num(8)	Any integer.
FAMINCOME1	less than $10,000	num(8)	Any proportion between 0 and 1.
FAMINCOME2	$10,000 - $14,999	num(8)	Any proportion between 0 and 1.
FAMINCOME3	$15,000-$19,999	num(8)	Any proportion between 0 and 1.
FAMINCOME4	$20,000-$24,999	num(8)	Any proportion between 0 and 1.
FAMINCOME5	$25,000-$29,999	num(8)	Any proportion between 0 and 1.
FAMINCOME6	$30,000-$34,999	num(8)	Any proportion between 0 and 1.
FAMINCOME7	$35,000-$39,999	num(8)	Any proportion between 0 and 1.
FAMINCOME8	$40,000-$44,999	num(8)	Any proportion between 0 and 1.
FAMINCOME9	$45,000-$49,999	num(8)	Any proportion between 0 and 1.
FAMINCOME10	$50,000-$59,999	num(8)	Any proportion between 0 and 1.
FAMINCOME11	$60,000-$74,999	num(8)	Any proportion between 0 and 1.
FAMINCOME12	$75,000-$99,999	num(8)	Any proportion between 0 and 1.
FAMINCOME13	$100,000-$124,999	num(8)	Any proportion between 0 and 1.
FAMINCOME14	$125,000-$149,999	num(8)	Any proportion between 0 and 1.
FAMINCOME15	$150,000-$199,999	num(8)	Any proportion between 0 and 1.
FAMINCOME16	$200,000+	num(8)	Any proportion between 0 and 1.

FAMPOVERTY	Proportion of family households in the geography with below-poverty level income.	num(8)	Any proportion between 0 and 1.
MEDHOUSINCOME	median household income	num(8)	Any integer.
HOUSINCOME1	< $10,000	num(8)	Any proportion between 0 and 1.
HOUSINCOME2	$10,000-$14,999	num(8)	Any proportion between 0 and 1.
HOUSINCOME3	$15,000-$19,999	num(8)	Any proportion between 0 and 1.
HOUSINCOME4	$20,000-$24,999	num(8)	Any proportion between 0 and 1.
HOUSINCOME5	$25,000-$29,999	num(8)	Any proportion between 0 and 1.
HOUSINCOME6	$30,000-$34,999	num(8)	Any proportion between 0 and 1.
HOUSINCOME7	$35,000-$39,999	num(8)	Any proportion between 0 and 1.
HOUSINCOME8	$40,000-$44,999	num(8)	Any proportion between 0 and 1.
HOUSINCOME9	$45,000-$49,999	num(8)	Any proportion between 0 and 1.
HOUSINCOME10	$50,000-$59,999	num(8)	Any proportion between 0 and 1.
HOUSINCOME11	$60,000-$74,999	num(8)	Any proportion between 0 and 1.
HOUSINCOME12	$75,000-$99,999	num(8)	Any proportion between 0 and 1.
HOUSINCOME13	$100,000-$124,999	num(8)	Any proportion between 0 and 1.
HOUSINCOME14	$125,000-$149,999	num(8)	Any proportion between 0 and 1.
HOUSINCOME15	$150,000-$199,999	num(8)	Any proportion between 0 and 1.
HOUSINCOME16	$200,000+	num(8)	Any proportion between 0 and 1.
HOUSPOVERTY	%Households with below-poverty level income	num(8)	Any proportion between 0 and 1.
POV_LT_50	<50% of poverty level	num(8)	Any proportion between 0 and 1.
POV_50_74	Between 50 and 74.99% of poverty level	num(8)	Any proportion between 0 and 1.
POV_75_99	Between 75 and 99.99% of poverty level	num(8)	Any proportion between 0 and 1.

POV_100_124	Between 100 and 124.99% of poverty level	num(8)	Any proportion between 0 and 1.
POV_125_149	Between 125 and 149.99% of poverty level	num(8)	Any proportion between 0 and 1.
POV_150_174	Between 150 and 174.99% of poverty level	num(8)	Any proportion between 0 and 1.
POV_175_184	Between 175 and 184.99% of poverty level	num(8)	Any proportion between 0 and 1.
POV_185_199	Between 185 and 199.99% of poverty level	num(8)	Any proportion between 0 and 1.
POV_GT_200	> 200% of poverty level	num(8)	Any proportion between 0 and 1.
ENGLISH_SPEAKER	Proportion of people over age 5 that speak only English or speak english “very well”	Num(8)	Any proportion between 0 and 1.
SPANISH_SPEAKER	Proportion of people over age 5 who speak only spanish or speak spanish “very well”	Num(8)	Any proportion between 0 and 1.
BORNINUS	Proportion of people over age 5 that were born in the US	Num(8)	Any proportion between 0 and 1.
MOVEDINLAST12MON	Proportion of households that moved in the last 12 months	Num(8)	Any proportion between 0 and 1.
MARRIED	Proportion of people over age 15 who are married	Num(8)	Any proportion between 0 and 1.
DIVORCED	Proportion of people over age 15 who are divorced	Num(8)	Any proportion between 0 and 1.
DISABILITY	Proportion of people over age 18 living with any disability	Num(8)	Any proportion between 0 and 1.
UNEMPLOYMENT	Proportion of civilian noninstitutionalized population between 18 and 64 who are unemployed	Num(8)	Any proportion between 0 and 1.
UNEMPLOYMENT_MALE	Proportion of civilian noninstitutionalized males between 18 and 64 who are unemployed	Num(8)	Any proportion between 0 and 1.
INS_MEDICARE	Proportion of people covered by Medicare	Num(8)	Any proportion between 0 and 1.
INS_MEDICAID	Proportion of people covered by Medicaid	Num(8)	Any proportion between 0 and 1.
HH_NOCAR	Proportion of households with no car (owner and renter occupied)	Num(8)	Any proportion between 0 and 1.

HH_PUBLIC_ASSISTANCE	Proportion of households on public assistance	Char(1)	Any proportion between 0 and 1.
HMOWNER_COSTS_MORT	Proportion of households with monthly owner costs > 50% HH income, in homes with mortgages	Num(8)	Any proportion between 0 and 1.
HMOWNER_COSTS_NO_MORT	Proportion of households with monthly owner costs > 50% HH income, in homes without mortgages	Num(8)	Any proportion between 0 and 1.
HOMES_MEDVALUE	Median value of homes	Num(8)	Any integer.
PCT_CROWDING	Proportion of households with >= 1 person per room	Num(8)	Any proportion between 0 and 1.
FEMALE_HEAD_OF_HH	Proportion of households headed by females (no male present)	Num(8)	Any proportion between 0 and 1.
MGR_FEMALE	Proportion of female management occupations	Num(8)	Any proportion between 0 and 1.
MGR_MALE	Proportion of male management occupations	Num(8)	Any proportion between 0 and 1.
RESIDENTS_65	Proportion of population over 65	Num(8)	Any proportion between 0 and 1.
SAME_RESIDENCE	Proportion of persons in same residence since year 2005	Num(8)	What years does ACS provide for this?

Primary Key:

GEOCODE + CENSUS_YEAR

Foreign Key Relationship:

Source Variable (Foreign Key)	Target Table	Target Variable (Primary Key)	Orphans allowed?
[None Defined]			

CENSUS DECENNIAL DEMOGRAPHICS

Version = 5.0.0 Date= 11/16/2021 StdVar = &_vdw_census_demog_dec

Subject Area Description

The CENSUS DECENNIAL DEMOGRAPHICS table contains data from the Decennial survey administered by the US Census Bureau. The Decennial survey is conducted every ten years; this table includes data at the census tract level. It is recommended that sites implement the most recent Decennial data Projects are encouraged to contact sites regarding the version of Census data available.

Variable Name	Definition	Type(Len)	Values
CENSUS_YEAR	Year of the Decennial survey	Num(8)	Year values (e.g. 2010, 2020)
GEOCODE	Concatenation of the FIPS codes for State, County, and Census Tract	Char(11)	The first two characters signify the state, the next three signify the county, the next six signify the tract.
STATE	State Code--a two-character numeric code assigned to US states, districts, territories and protectorates.	char(2)	Any valid state code, as used by the USCB.
COUNTY	County Code--a 3-character numeric code assigned to census counties.	char(3)	Any valid county code, as used by the USCB.
TRACT	Tract Code--a six-character numeric code assigned to census tracts.	char(6)	Any valid tract code, as used by the USCB.
RA_NHS_WH	%white, non-hispanic	num(8)	Any proportion between 0 and 1.
RA_NHS_BL	%black or african american, non-hispanic	num(8)	Any proportion between 0 and 1.
RA_NHS_AM	%american indian or alaska native, non-hispanic	num(8)	Any proportion between 0 and 1.
RA_NHS_AS	%asian, non-hispanic	num(8)	Any proportion between 0 and 1.
RA_NHS_HA	%native hawaiian or other pacific islander, non-hispanic	num(8)	Any proportion between 0 and 1.
RA_NHS_OT	%other, non-hispanic	num(8)	Any proportion between 0 and 1.

RA_NHS_ML	%two or more races, non-hispanic	num(8)	Any proportion between 0 and 1.
RA_HIS_WH	%white, hispanic	num(8)	Any proportion between 0 and 1.
RA_HIS_BL	%black or african american, hispanic	num(8)	Any proportion between 0 and 1.
RA_HIS_AM	%amercian indian or alaska native, hispanic	num(8)	Any proportion between 0 and 1.
RA_HIS_AS	%asian, hispanic	num(8)	Any proportion between 0 and 1.
RA_HIS_HA	%native hawaiian or other pacific islander, hispanic	num(8)	Any proportion between 0 and 1.
RA_HIS_OT	%other, hispanic	num(8)	Any proportion between 0 and 1.
RA_HIS_ML	%two or more races, hispanic	num(8)	Any proportion between 0 and 1.
HOUSES_N	Number of housing units in geography	num(8)	Any integer.
HOUSES_OCCUPIED	Proportion of houses_n that are occupied.	num(8)	Any proportion between 0 and 1.
HOUSES_OWN	Proportion of occupied housing units that are occupied by owners.	num(8)	Any proportion between 0 and 1.
HOUSES_RENT	Proportion of occupied housing units that are occupied by renters.	num(8)	Any proportion between 0 and 1.
HOUSES_UNOCC_FORRENT	Proportion of unoccupied housing units that are for rent.	num(8)	Any proportion between 0 and 1.
HOUSES_UNOCC_FORSALE	Proportion of unoccupied housing units that are for sale	num(8)	Any proportion between 0 and 1.
HOUSES_UNOCC_RENTSOLD	Proportion of unoccupied housing units that are rented or sold, but still unoccupied.	num(8)	Any proportion between 0 and 1.
HOUSES_UNOCC_SEASONAL	Proportion of unoccupied housing units that are used seasonally or other occasional use.	num(8)	Any proportion between 0 and 1.
HOUSES_UNOCC_MIGRANT	Proportion of unoccupied housing units that are used for migrant workers.	num(8)	Any proportion between 0 and 1.
HOUSES_UNOCC_OTHER	Proportion of unoccupied housing units that are vacant for other reasons.	num(8)	Any proportion between 0 and 1.

Primary Key:

GEOCODE + CENSUS_YEAR

Foreign Key Relationship:

Source Variable (Foreign Key)	Target Table	Target Variable (Primary Key)	Orphans allowed?
[None Defined]			

PRO TYPES

Version = 5.1.0 Date = 1/4/2022 StdVar = &_vdw_pro_types

Subject Area Description

The PRO_TYPES table serves as a lookup dataset for patient surveys based on the survey type & subtype. Every pro_id, pro_type, and pro_subtype appearing in other VDW files should appear in the PRO_TYPES table.

Variable Name	Definition	Type(Len)	Values
PRO_ID	Unique identifier to a questionnaire type and subtype combination	Char(10)	Use designated values from the PRO Workgroup where they exist. For site specific surveys, choose an appropriate number, prefixed by your site abbreviation (Ex: HFHS_101)
PRO_TYPE	Name of the questionnaire type	Char(15)	Example: PHQ
PRO_TYPE_DESCRIPTION	Description of questionnaire type	Char(80)	Example: Patient Health Questionnaire
PRO_SUBTYPE	Name of the questionnaire subtype	Char(15)	Example: PHQ-9
PRO_SUBTYPE_DESCRIPTION	Description of questionnaire subtype	Char(80)	
PRO_TYPE_NOTES	Additional notes for the questionaire type	Char(255)	

Primary Key:

PRO_ID

Foreign Key Relationship:

Source Variable (Foreign Key)	Target Table	Target Variable (Primary Key)	Orphans allowed?
[None Defined]			

PRO SURVEYS

Version = 5.1.0 Date = 1/4/2022 StdVar = &_vdw_pro_surveys

Subject Area Description

The PRO_SURVEYS table serves as a lookup dataset for the PRO_SURVEY_RESPONSES table, and contains survey metadata, such as the survey name, question text, and question version.

Variable Name	Definition	Type(Len)	Values
PRO_ID	Unique identifier to a questionnaire type and subtype combination	Char(10)	Use values from the PRO_Types table.
SURVEY_DATA_SOURCE	Data source of the survey	Char(100)	Varies by site. Ex. Questionnaires, Vendor, etc.
SURVEY_ID	Unique identifier to a collection of questions.	Char(18)	
SURVEY_NAME	The name of the survey	Char(100)	Ex: DEPRESSION SCREEN - PHQ 9
QUESTION_ID	Unique identifier for a question	Char(18)	
QUESTION_VER	Sequence number of available versions of each	Num	Ex: 1, 2, 3
QUESTION_DATE	Date the question was published	Num	
QUESTION_TEXT	Standardized question text	Char(255)	
QUESTION_TEXT_ORIG	Source question text	Char(255)	
QUESTION_LANG	Language of the source question text	Char(3)	Values are consistent with lang_iso from the VDW

Primary Key:

PRO_ID + SURVEY_ID +
QUESTION_ID +
QUESTION_VER

Foreign Key Relationship:

Source Variable (Foreign Key)	Target Table	Target Variable (Primary Key)	Orphans allowed?
[None Defined]			

PRO SURVEY RESPONSES

Version = 5.2.0 Date = 4/21/2023 StdVar = &_vdw_pro_survey_responses

Subject Area Description

The PRO_SURVEY_RESPONSES table holds question level responses to patient surveys, one record per response.

Variable Name	Definition	Type(Len)	Values
PRO_ID	Unique identifier to a questionnaire type and subtype	Char(10)	Use values from the PRO_Types table.
SURVEY_ID	Unique identifier to a questionnaire	Char(18)	Use values from the PRO_Surveys table.
QUESTION_ID	Unique identifier for a question	Char(18)	Use values from the PRO_Surveys table.
QUESTION_VER	Sequence number of available versions of each question	Num	Use values from the PRO_Surveys table.
MRN	Identifier unique to an individual. Used to link across VDW ta	Char(Varies)	Use values from the VDW Demographics table.
RESPONSE_DATE	Date associated with the entered response	Num	
RESPONSE_TIME	Time associated with the entered response	Num	
RESPONSE_TEXT	Standardized response text	Char(255)	
RESPONSE_TEXT_ORIG	Source response text	Char(255)	
ENC_ID	Unique identifier used for linking table to the VDW Encounter table	Char(Varies)	Values should be populated and mapped to the VDW Encounter table when possible. However, missing and/or unlinked values are acceptable.
SURVEY_MODE	How the survey question was administered.	Char(2)	TV = Telephone/Video OV = Office/Hospital Visit HV = Home Visit GV = Group Visit MA = Mail WE = Web/Electronic OT = Other UN = Unknown

Primary Key:

None

Foreign Key Relationship:

Source Variable (Foreign Key)	Target Table	Target Variable (Primary Key)	Orphans allowed?
[None Defined]			

Version = 5.0.1 Date = 4/21/2023 Foreign Key Tables (Copied as Linked Pictures)

ENROLLMENT

Source Variable (Foreign Key)	Target Table	Target Variable (Primary Key)	Orphans allowed?
MRN	DEMOGRAPHICS	MRN	No
PCP	PROVIDER	PROVIDER	No

PHARMACY

Source Variable (Foreign Key)	Target Table	Target Variable (Primary Key)	Orphans allowed?
MRN	DEMOGRAPHICS	MRN	No
NDC	EVERNDC	NDC	No
RXMD	PROVIDER	PROVIDER	No

DEMOGRAPHICS

Source Variable (Foreign Key)	Target Table	Target Variable (Primary Key)	Orphans allowed?
[None Defined]			

ENCOUNTER

Source Variable (Foreign Key)	Target Table	Target Variable (Primary Key)	Orphans allowed?
MRN	DEMOGRAPHICS	MRN	No
PROVIDER	PROVIDER	PROVIDER	No

DIAGNOSIS

Source Variable (Foreign Key)	Target Table	Target Variable (Primary Key)	Orphans allowed?
MRN	DEMOGRAPHICS	MRN	No
ENC_ID	ENCOUNTER	ENC_ID	No
PROVIDER	PROVIDER	PROVIDER	No
DIAG_PROVIDER	PROVIDER	PROVIDER	No

PROCEDURE

Source Variable (Foreign Key)	Target Table	Target Variable (Primary Key)	Orphans allowed?
MRN	DEMOGRAPHICS	MRN	No
ENC_ID	ENCOUNTER	ENC_ID	No
PROVIDER	PROVIDER	PROVIDER	No
PERFORMINGPROVIDER	PROVIDER	PROVIDER	No

PROVIDER

Source Variable (Foreign Key)	Target Table	Target Variable (Primary Key)	Orphans allowed?
[None Defined]			

LANGUAGE

Source Variable (Foreign Key)	Target Table	Target Variable (Primary Key)	Orphans allowed?
MRN	DEMOGRAPHICS	MRN	No

DEATH

Source Variable (Foreign Key)	Target Table	Target Variable (Primary Key)	Orphans allowed?
MRN	DEMOGRAPHICS	MRN	No

CAUSE OF DEATH

Source Variable (Foreign Key)	Target Table	Target Variable (Primary Key)	Orphans allowed?
MRN	DEMOGRAPHICS	MRN	No

VITALS

Source Variable (Foreign Key)	Target Table	Target Variable (Primary Key)	Orphans allowed?
MRN	DEMOGRAPHICS	MRN	No
ENC_ID	ENCOUNTER	ENC_ID	No

EVERNDC

Source Variable (Foreign Key)	Target Table	Target Variable (Primary Key)	Orphans allowed?
[None Defined]			

LAB RESULTS

Source Variable (Foreign Key)	Target Table	Target Variable (Primary Key)	Orphans allowed?
MRN	DEMOGRAPHICS	MRN	No
ORDER_PROV	PROVIDER	PROVIDER	No

LAB NOTES

Source Variable (Foreign Key)	Target Table	Target Variable (Primary Key)	Orphans allowed?
No Foreign Key			

TUMOR MAIN

Source Variable (Foreign Key)	Target Table	Target Variable (Primary Key)	Orphans allowed?
MRN	DEMOGRAPHICS	MRN	No

SOCIAL HISTORY

Source Variable (Foreign Key)	Target Table	Target Variable (Primary Key)	Orphans allowed?
MRN	DEMOGRAPHICS	MRN	No
ENC_ID	ENCOUNTER	ENC_ID	No

Census Location

Source Variable (Foreign Key)	Target Table	Target Variable (Primary Key)	Orphans allowed?
MRN	DEMOGRAPHICS	MRN	No
GEOCODE + CENSUS_YEAR	CENSUS DEMOG	GEOCODE + CENSUS_YEAR	No

Census Demographics

Source Variable (Foreign Key)	Target Table	Target Variable (Primary Key)	Orphans allowed?
[None Defined]			