

# **Evolution and Implementation of the CDISC Study Data Tabulation Model (SDTM)**

Fred Wood  
Octagon Research Solutions

Tom Guinter  
Take Solutions

PharmaSUG Annual Meeting, May 2007

# SDTM and the IGs

## **Study Data Tabulation Model – “Umbrella Document”**

- Developed by the CDISC SDS Team
- Final v1.1 posted 2005-04-28
- 28 Pages

### **SDTMIG v3.1.1 – Submission Data Standards**

- SDTM Implementation Guide: Human Clinical Trials
- Developed by the CDISC SDS Team
- Final v3.1.1 posted 2005-08-26
- 183 pages (will be close to 300 for v3.1.2)

### **SEND v2.3 – Standard for Exchange of Non-Clinical Data**

- SDTM implementation guide for non-clinical studies
- Developed by the CDISC SEND Team
- Under development
- 70-80 pages

# Current Status of the SDTMIG

- Version 3.1.2 posted for Comment mid-July
  - 300 +/- pages
  - New standard domains
  - More detail
  - Move section 9 examples to section 6
  - More examples and assumptions
  - Remove editorial and legacy text
  - Clarifications that can be used now (v3.1.1 compliant)
  - “Findings about” (more explanation later)

# Standards Regulatory History (1)

- July 21, 2004: The FDA adds Study Data Specifications to the draft eCTD Guidance (which has become final as of April, 2006) for submitting animal and human study data in electronic format. This specification references the CDISC **SDTM** for data tabulation datasets.
- December 13, 2004: HHS Regulatory Agenda. Announcement in Federal Register of Proposed Rule Stage for “Submission of Standardized Electronic Study Data From Clinical Studies Evaluating Human Drugs and Biologics”. This notification was reinforced at the FDA Public Meeting held on the **Study Data Tabulation Model** on February 1, 2005. 69 FR 73119.
- December 30, 2004: CBER announced a pilot project to test viewing tools with data submitted using the **SDTM**. Until this time, the SDTM had mainly been supported by CDER. 69 FR 78468.

# Standards Regulatory History (2)

- September 29, 2006: Federal Register notice of **withdrawal of three Electronic Submission Guidances for eNDA, eANDA, and eAnnual Reports**. This notice designates eCTD as “preferred format for electronic submissions” and notes that beginning Jan 1, 2008 any electronic submission going to CDER must be eCTD.
- December 11, 2006: Proposed Rule Change that impacts all NDAs, BLAs, ANDAs, supplements, and amendments. This notice affects clinical data and bioequivalence data, and mandates that data must be submitted and provided in an electronic format that the FDA can process, review, and archive. It also mandates the use of standardized data structures, terminology, and code sets according to the CDISC **SDTM** guidance, and provides a two-year transition period for implementation.

# SDTM Background

- Provides means to submit tabulation data to the FDA in a standard format.
- Defines specific rules
  - Structures - General Observation Classes
  - Domain (dataset names) - Some Standard
  - Variable Names - Defined in the SDTMIG
  - Some Controlled Terminology
- The sponsor decides what data to submit, based on science and regulation.

# General Observation Classes

- **Interventions**
- **Events**
- **Findings**

**Each of these has defined structure and variables. No new variables may be added to these observation classes.**

# Interventions

**Interventions:** investigational treatments, therapeutic treatments, and surgical procedures administered to the subject.

- Things that cause a physiological change in subject
- One record per constant-dosing/treatment interval
- The Topic variable is --TRT
- SDTMIG Intervention Domains
  - Concomitant and Prior Medications (CM)
  - Study Medications (EX)
  - Substance Use (SU)



# Topic and Qualifier Variables for Interventions

--TRT	Name of Treatment
--MODIFY	Modified Treatment Name
--DECOD	Standardized Treatment Name
--CAT	Category
--SCAT	Subcategory
--OCCUR	Occurrence
--STAT	Status
--REASND	Reason
--INDC	Indication
--CLAS	Class
--CLASCD	Class Code
--DOSE	Dose
--DOSTXT	Dose Description

Topic Variable

--DOSU	Dose Units
--DOSFRM	Dose Form
--DOSFRQ	Dosing Frequency per Interval
--DOSTOT	Total Daily Dose
--DOSRGM	Dose Regimen
--ROUTE	Route of Administration
--LOT	Lot Number
--LOC	Location of Dose Administration
--TRTV	Treatment Vehicle
--ADJ	Reason for Dose Adjustment

-- represents the two-letter domain prefix

# Events

**Events:** occurrences or incidents independent of planned study evaluations occurring during the trial or prior to the trial. Things that happen.

- There can be planned evaluations
- One record per event
- The Topic variable is --TERM
- SDTMIG Event Domains
  - Adverse Events (AE)
  - Disposition (DS)
  - Medical History (MH)
  - Deviations (DV)

# Topic and Qualifier Variables for **Events**

--TERM	Reported Term
--MODIFY	Modified Reported Term
--DECOD	Dictionary-Derived Term
--CAT	Category
--SCAT	Subcategory
--OCCUR	Occurrence
--STAT	Status
--REASND	Reason
--BODSYS	Body System or Organ Class
--LOC	Location of the Reaction
--SEV	Severity/Intensity
--SER	Serious Event
--ACN	Action Taken with Study Treatment
--ACNOTH	Other Action Taken
--REL	Causality

--RELNST	Relationship to Non-Study Treatment
--PATT	Pattern of Event
--OUT	Outcome of Event
--SCAN	Involves Cancer
--SCONG	Congenital Anomaly or Birth Defect
--SDISAB	Persist or Signif Disability/Incapacity
--SDTH	Results in Death
--SHOSP	Requires or Prolongs Hospitalization
--SLIFE	Is Life Threatening
--SOD	Occurred with Overdose
--SMIE	Other Medically Important Serious Event
--CONTRT	Concomitant or Additional Trtmnt Given
-TOXGR	Toxicity Grade

Topic Variable

# Findings

**Findings:** observations resulting from planned evaluations.

- Measurements, tests, and questions
- One record per finding result or measurement
- The Topic variable is --TESTCD
- It is likely that ~80% of data could be Findings
- SDTMIG Finding Domains:
  - ECG Test Results (EG)
  - Inclusion/Exclusion Exceptions (IE)
  - Laboratory Test Results (LB)
  - Physical Examinations (PE)
  - Questionnaires (QS)
  - Subject Characteristics (SC)
  - Vital Signs (VS)
  - Drug Accountability (DA)

# Topic and Qualifier Variables for Findings

--TESTCD	Short Name of Test
--TEST	Name of Test
--MODIFY	Modified Term
--CAT	Category
--SCAT	Subcategory
--POS	Position of Subject
--BODSYS	Body System or Organ Class
--ORRES	Result or Finding in Original Units
--ORRESU	Original Units
--ORNRLO	Normal Range Lower Limit-Original Units
--ORNRHI	Normal Range Upper Limit-Original Units
--STRESC	Result or Finding in Standard Format

--STRESN	Numeric Result or Finding in Standard Units
--STRESU	Standard Units
--STNRLO	Normal Range Lower Limit-Standard Units
--STNRHI	Normal Range Upper Limit-Standard Units
--STNRC	Normal Range for Character Results-Standard Units
--NRIND	Normal/Reference Range Indicator
--RESCAT	Result Category
--STAT	Status
--REASND	Reason
--XFN	External Filename

Topic Variable

-- represents the two-letter domain prefix

# Additional Qualifier Variables for Findings

--NAM	Laboratory/Vendor Name
--LOINC	LOINC Code
--SPEC	Specimen Material Type
--SPCCND	Specimen Condition
--LOC	Location Used for the Measurement
--METHOD	Method of Test or Examination
--BLFL	Baseline Flag
--FAST	Fasting Status
--DRVFL	Derived Flag
--EVAL	Evaluator
--TOX	Toxicity
--TOXGR	Toxicity Grade
--SEV	Severity
--DTHREL	Relationship to Death

*-- represents the two-letter domain prefix*

# Domains - Standard

**Interventions**

**Conmeds**

**Exposure**

**Subst Use**

**Events**

**AE**

**Disposition**

**MH**

**Deviations**

**Findings**

**ECG**

**Incl/Excl**

**Labs**

**PhysExam**

**PK Conc  
v3.1.2**

**Micro MB  
v3.1.2**

**Ques'aire**

**SubjChar**

**Vitals**

**DrugAcct**

**PK Param  
v3.1.2**

**Micro MS  
v3.1.2**

**Special Purpose**

**Demog**

**Comments**

**SUPPQUAL**

**RELREC**

**Trial Design  
(7 Tables)**

# Identifier Variables for All Classes

STUDYID	Study Identifier
DOMAIN	Domain Abbreviation
USUBJID	Unique Subject Identifier
--SEQ	Sequence Number
--GRPID	Group ID
--REFID	Reference ID
--SPID	Sponsor ID

*-- represents the two-letter domain prefix;  
note that some Identifiers do not use a prefix.*



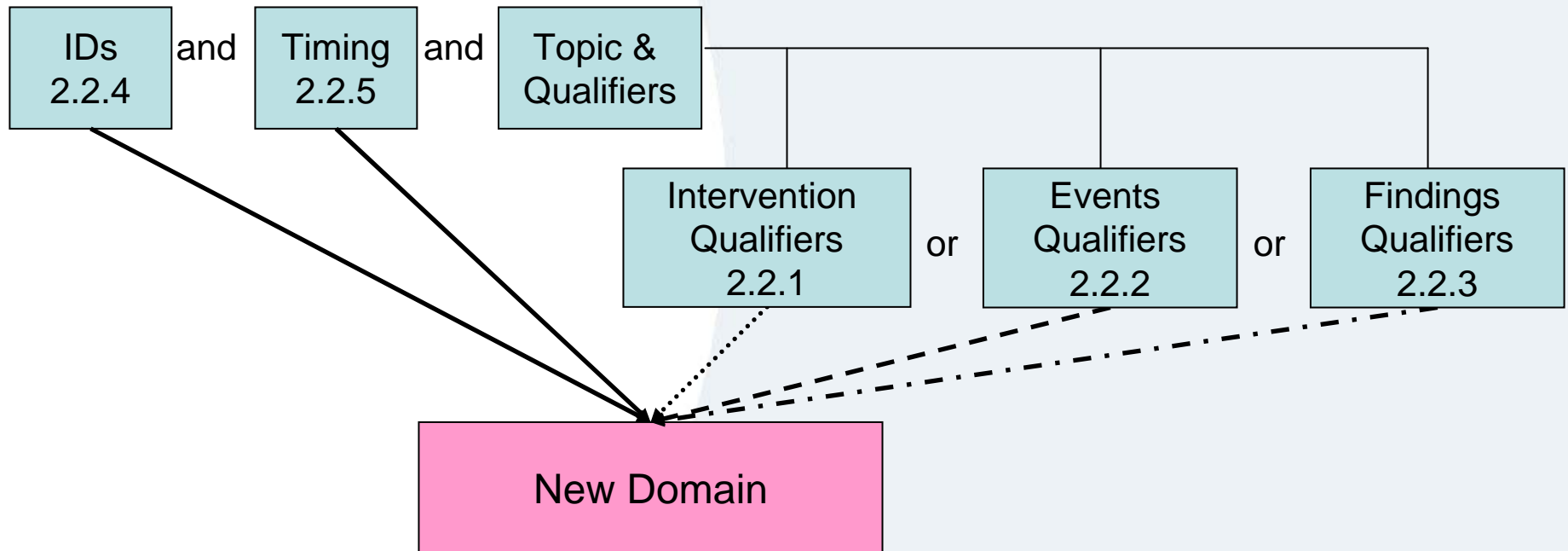
# Timing Variables for All Classes

<b>Visits</b>	VISIT	Visit Name
	VISITNUM	Visit Number
	VISITDAY	Visit (Study) Day
<b>Dates</b>	--DTC	Date/Time of Collection
	--STDTC	Start Date/Time
	--ENDTC	End Date/Time
	--RFTDTC	Reference Time Point Date
<b>Days</b>	--DY	Study Day of collection
	--STDY	Start Study Day
	--ENDY	End Study Day
<b>Time Points</b>	--TPT	Time Point Name
	--TPTNUM	Time Point Number
	--TPTREF	Time Point Reference
<b>Durations</b>	--DUR	Collected duration
	--ELTM	Elapsed Time from Time Points
<b>Reference</b>	--STRF	BEFORE, DURING, AFTER DM Ref Period
	--ENRF	BEFORE, DURING, AFTER DM Ref Period

*-- represents the two-letter domain prefix; note that some Timing variables do not use a prefix.*

# Creating Domains (1)

1. Choose a general observation class.
2. Assign a 2-letter domain code. Check SDTMIG Appendix 10.3.1.
3. Add variables as follows:



4. Modify labels of similar variables in other datasets and follow metadata model for order, name/label/type.

# Creating Domains (2)

- No new sponsor-defined variables can be added to SDTM domains
  - Options for submission data that does not fit:
    - New Domain(s)
    - Supplemental Qualifiers
- Additional SDTM variables can be added as needed:
  - Timing, Identifiers, as well as Qualifiers from the same general observation class
- Variable attributes must conform to the SAS V5 transport file conventions
  - Variables names  $\leq 8$
  - Variable Labels  $\leq 40$
  - Length  $\leq 200$

# Variable Requirements

- Required variables need to be in the domains
  - Their values cannot be null
- Expected variables need to be in the domains
  - Some, but not all, values may be null
- Permissible variables are included in the domains as needed.

# Date/Time Formats

- ISO 8601 is a TEXT string (not a SAS format)
- Submit only those parts of the date/time that were collected
- Has ability to handle incomplete dates (missing middle component[s])

<b>Collected</b>	<b>Certainty</b>	<b>--DTC Value</b>
Dec 15, 2003 13:14:17	Complete	2003-12-15T13:14:17
Dec 15, 2003 13	Unknown minutes	2003-12-15T13
Dec 15, 2003	Unknown time	2003-12-15
Dec 2003	Unknown day	2003-12
2003	Unknown month	2003

# Fitting Data into the Findings Observation Class (1)

Sponsor's Structure for Vital Signs:  
One Record per Subject per Visit

SUBJID	VSDTM	SYSBP in mm	DIABP in mm	PULSE bpm	TEMP in °C
0001	2003-02-01	120	80	65	37

# Fitting Data into the Findings Observation Class (2)

SDTM Structure for Vital Signs:  
One Record per Test

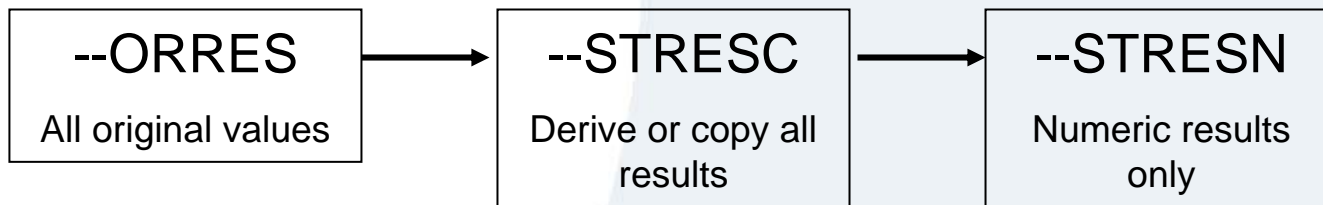
USUBJID	VSTESTCD	VSORRES	VSORRESU	VSDTC
ABC-0001	SYSBP	120	mm	2003-02-01
ABC-0001	DIABP	80	mm	2003-02-01
ABC-0001	PULSE	65	bpm	2003-02-01
ABC-0001	TEMP	37	C	2003-02-01

Former column names  
become values in TESTCD

Units come from  
former SAS label

# Findings Variables – Results (1)

- Each Findings Domain has variables to store:
  - Original results (--ORRES)
  - Standardized character results ( --STRESC)
  - Standardized numeric results (--STRESN)



LBTESTCD	LBORRES	LBORRESU	LBSTRESC	LBSTRESN	LBSTRESU
GLUCOSE	6.0	mg/dL	60.0	60.0	mg/L
URBC	TRACE		1	1	



# Findings Variables – Results (2)

- The Findings domain provides mechanism for measurements/test not done
  - --TESTCD and --TEST are populated
  - --ORRES is null
  - --STAT = NOT DONE
  - --REASND for reason not done
- If entire page is not performed then
  - --TESTCD = --ALL
- --ORRES may also be null for derived records (averages of collected data, or questionnaire subscores)

# Representing Relationships

1. Relating groups of records in a domain for a subject
- 2. Relating non-standard variables to a parent domain**
- 3. Relating independent records in separate domains**
- 4. Relating datasets**
5. Comments

# Relating Non-Standard Variables to a Parent Domain (1)

- SUPPQUAL contains the relationship keys in IDVAR and IDVARVAL
  - Does not require RELREC
  - Contains the name and attributes of the extras items
    - QNAM, QLABEL, QVAL, QORIG, QEVAL
- Use either:
  - a) One SUPPQUAL (Supplemental Qualifiers) dataset or
  - b) Multiple SUPP-- datasets

Do not mix and match

The latter will be recommended in SDTMIG v3.1.2

# Relating Non-Standard Variables to a Parent Domain (2)

Events (HO)  
Dataset

STUDYID	USUBJID	DOMAIN	HOSEQ	HOTERM	HOSTDTC	HOENDTC	HODUR
1999001	0001	HO	1	Hospitalization	2004-01-05	2004-01-12	P1W
1999001	0001	HO	2	Hospitalization	2004-01-23	2004-02-07	P15D
1999001	0002	HO	1	Hospitalization	2004-20-01	2004-01-22	P1D

SUPPHO  
Dataset

STUDYID	USUBJID	RDOMAIN	IDVAR	IDVARVAL	QNAM	QLABEL	QVAL
1999001	0001	HO	HOSEQ	1	AEREPF	AE Reported This Episode	Y
1999001	0001	HO	HOSEQ	1	MEDSFL	Meds Prescribed?	Y
1999001	0001	HO	HOSEQ	1	PROCFL	Procedures Performed?	Y
1999001	0001	HO	HOSEQ	1	PROVNM	Provider Name	General Hosp
1999001	0001	HO	HOSEQ	1	SPUNCD	Specialized Unit Type	ICU
1999001	0001	HO	HOSEQ	1	SPUNFL	Any Time in Spec. Unit	Y
1999001	0001	HO	HOSEQ	1	RLCNDF	Visit Related to Study Med	Y
1999001	0001	HO	HOSEQ	2	AEREPF	AE Reported This Episode	Y
1999001	0001	HO	HOSEQ	2	MEDSFL	Meds Prescribed?	Y
1999001	0001	HO	HOSEQ	2	PROCFL	Procedures Performed?	N
1999001	0001	HO	HOSEQ	2	PROVNM	Provider Name	Univ Hosp
1999001	0001	HO	HOSEQ	2	SPUNCD	Specialized Unit Type	CCU
1999001	0001	HO	HOSEQ	2	SPUNFL	Any Time in Spec. Unit	Y
1999001	0001	HO	HOSEQ	2	RLCNDF	Visit Related to Study Med	Y
1999001	0001	HO	HOSEQ	2	REASON	Reason	ARRAY
1999001	0002	HO	HOSEQ	1	AEREPF	AE Reported This Episode	Y
1999001	0002	HO	HOSEQ	1	MEDSFL	Meds Prescribed?	N
1999001	0002	HO	HOSEQ	1	PROCFL	Procedures Performed?	Y
1999001	0002	HO	HOSEQ	1	PROVNM	Provider Name	St. Mary's
1999001	0002	HO	HOSEQ	1	SPUNCD	Specialized Unit Type	
1999001	0002	HO	HOSEQ	1	SPUNFL	Any Time in Spec. Unit	N

# Relating Non-Standard Variables to a Parent Domain (3)

## Events (HO) Dataset

STUDYID	USUBJID	DOMAIN	HOSEQ	HOTERM	HOSTDTC	HOENDTC
1999001	0001	HO	1	Hospitalization	2004-01-05	2004-01-12

## SUPPHO Dataset

STUDYID	USUBJID	RDOMAIN	IDVAR	IDVARVAL	QNAM	QLABEL	QVAL
1999001	0001	HO	HOSEQ	1	AEREPEF	AE Reported This Episode	Y
1999001	0001	HO	HOSEQ	1	MEDSFL	Meds Prescribed?	Y
1999001	0001	HO	HOSEQ	1	PROCFL	Procedures Performed?	Y
1999001	0001	HO	HOSEQ	1	PROVNM	Provider Name	General Hosp
1999001	0001	HO	HOSEQ	1	SFUNCD	Specialized Unit Type	ICU
1999001	0001	HO	HOSEQ	1	SPUNFL	Any Time in Spec. Unit	Y

Renamed variables, as necessary

The SUPPQUAL records are linked to the Events dataset via STUDYID, USUBJID, DOMAIN, and the key specified in IDVAR

Former variable names

Former values


Former variable labels

# Relating Independent Records in Separate Domains (1)

- Used for **collected** relationships (not post hoc determinations)
- Requires the additional dataset RELREC, which identifies relationships between records
- Examples:
  - Discontinuation due to AE
  - CM/LB related to AE

# Relating Independent Records in Separate Domains (2)

STUDYID	USUBJID	RDOMAIN	IDVAR	IDVARVAL	RELID
1999001	0001	HO	HOSEQ	1	1
1999001	0001	AE	AESEQ	4	1
1999001	0001	HO	HOSEQ	2	2
1999001	0001	AE	AESEQ	11	2



Related records are given same RELID.  
RELID is only unique within subject.

- For Subject 0001, the first hospitalization record (HOSEQ=1) was related to AESEQ=4. These records are assigned a RELID value of 1.
- For Subject 0001, the second hospitalization record (HOSEQ=2) was related to AESEQ=11. These are assigned a RELID value of 2.

# Relating Datasets

- Use when two datasets have related information
- Requires RELTYPE, populated with values of ONE or MANY
- USUBJID not populated because the relationship exists for all subjects. USUBJID will still be part of the merge key.
- IDVARVAL is not populated because the relationship exists for all values of IDVAR. IDVAR will still be part of the merge key.

Row	STUDYID	RDOMAIN	USUBJID	IDVAR	IDVARVAL	RELTYPE	RELID
1	EFC1234	MB		MBSPID		ONE	A
2	EFC1234	MS		MSSPID		MANY	A



# Trial Design (1)

Purpose is to support reviewer to:

- Clearly and quickly grasp the design of a clinical trial
- Compare the designs of different trials
- Search a data warehouse for clinical trials with certain features
- Compare planned and actual treatments and visits for subjects in a clinical trial (future)

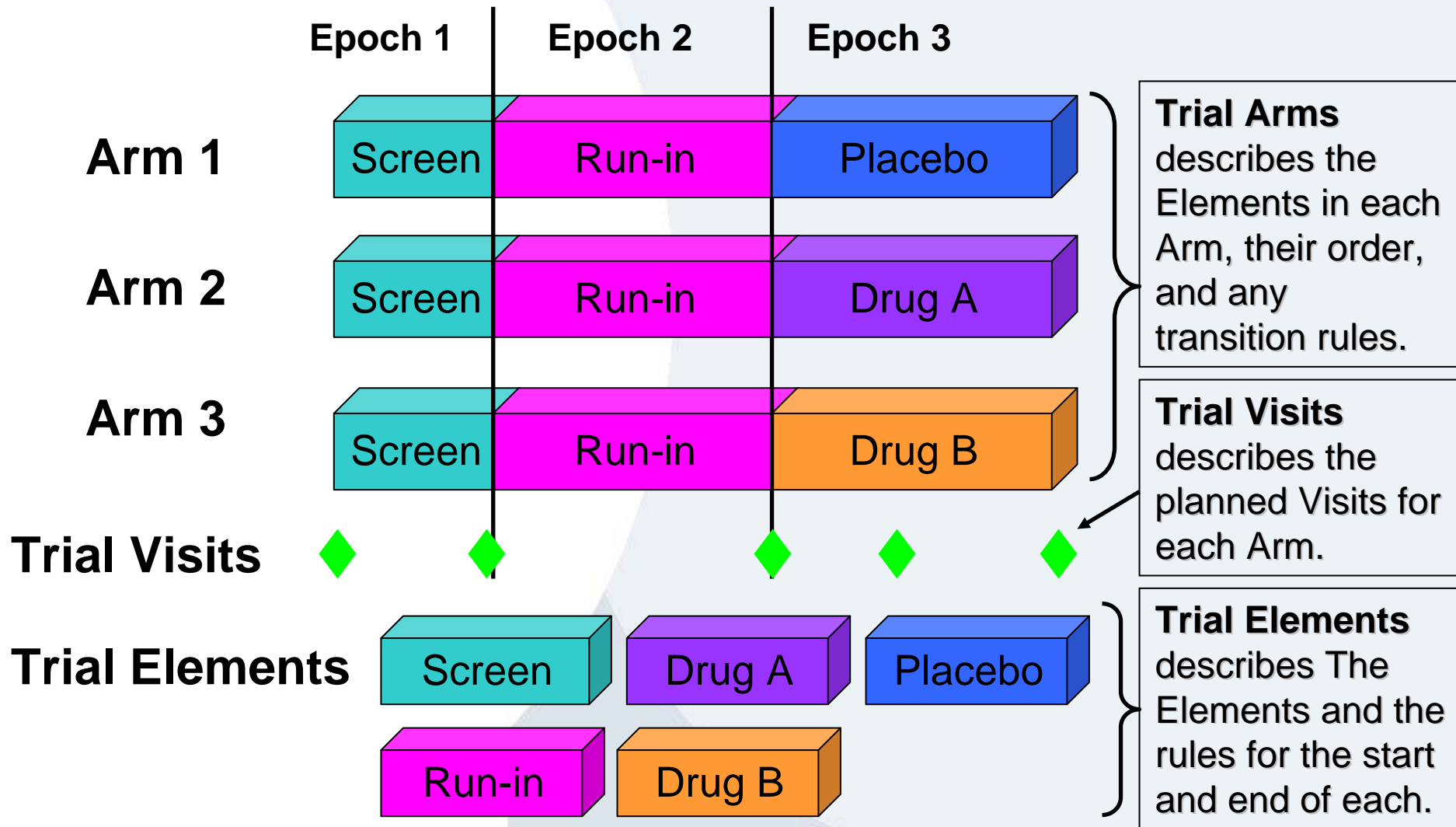
# Trial Design (2)

- Trial Arms (Planned)
- Trial Elements (Planned)
- Trial Visits (Planned)
- Trial Inclusion/Exclusion Criteria (Lookup table)
- Trial Summary (Descriptive attributes of trial)
- Subject Elements (Actual)
- Subject Visits (Actual)

The latter two will become special-purpose domains, moving out of Trial Design, for SDTMIG v3.1.2.

# Trial Design (3)

## Trial Elements, Arms, Visits



# Current Status of the SDTM (1)

## New Domains in SDTMIG v3.1.2

- Deviations (DV)
- Drug Accountability (DA)
- Pharmacokinetics (PC & PP)
- Pharmacogenomics (PG)
- Microbiology (MB and MS)

### Work in Progress

- Oncology Domains
- Device Domains
- Image Information

# Current Status of the SDTM (2)

## Enhancements in SDTMIG v3.1.2

- Multiple qualifiers for an observation (LOC)
  - Share the same timing
- Other “Specify” \_\_\_\_\_
- Findings about Events
  - Qualifiers with multiple values, such as severity of an Event collected at each visit

# One-to-Many Relationships

## The Use of “Multiple”

**AE Dataset**

AETERM	AELOC
RASH	MULTIPLE

**SUPPAE  
Dataset**

QNAM	QLABEL	QVAL
AELOC1	Location of the Reaction 1	FACE
AELOC2	Location of the Reaction 2	NECK
AELOC3	Location of the Reaction 3	CHEST

*Only major variables are shown in this example. SDTM domains must contain additional required variables.*

# “Other, Specify”

- Other, Specify: \_\_\_\_\_
  - “OTHER” goes in domain variable.
  - The specified value goes in SUPP--.QVAL
  - The rule is being updated in SDTMIG v3.1.2
    - When the specify value is for:
      - The topic variable ( --TRT, --TERM), the specified value goes in the topic variable.
      - A result variable, the specified value goes into --ORRES, while either “OTHER” or the specified value can go into --STRESC, depending on the sponsor’s conventions and needs.
      - A qualifier variable, use SUPP-- as described above.

# Findings About Events

2007-01-01

2007-03-01

Nausea

Severe

Mild

Moderate

Moderate

V0

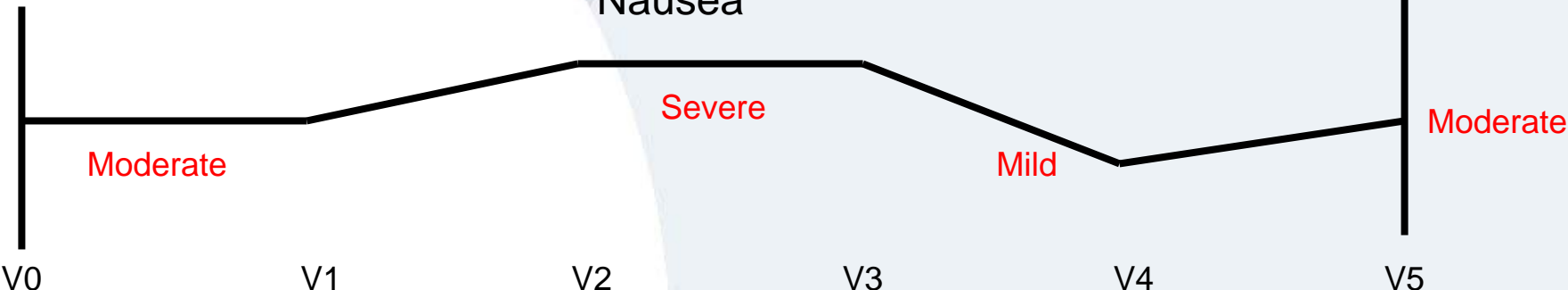
V1

V2

V3

V4

V5





# Findings About Events

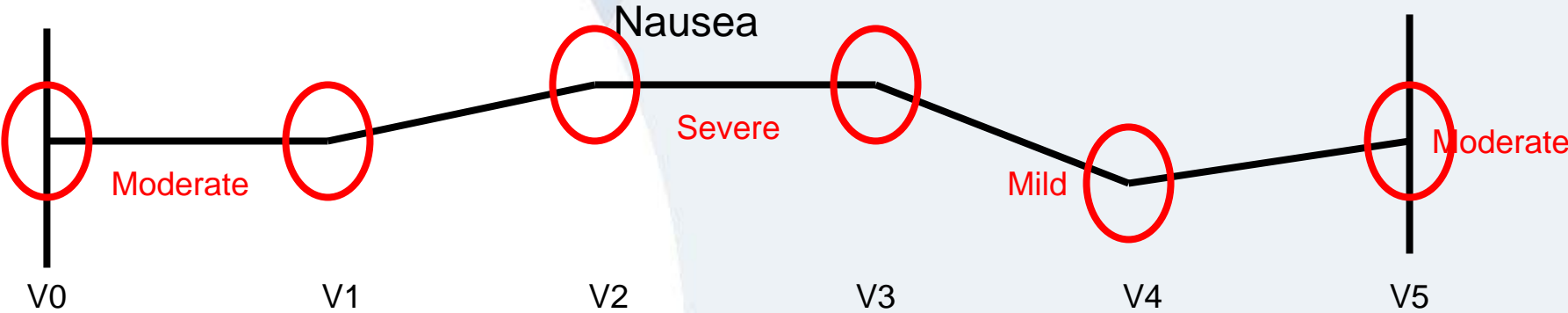


DOMAIN	USUBJID	AESEQ	AETERM	AESTDTC	AEENDTC	AESEV
AE	123-01-01	1	Nausea	2007-01-01	2007-03-01	Severe

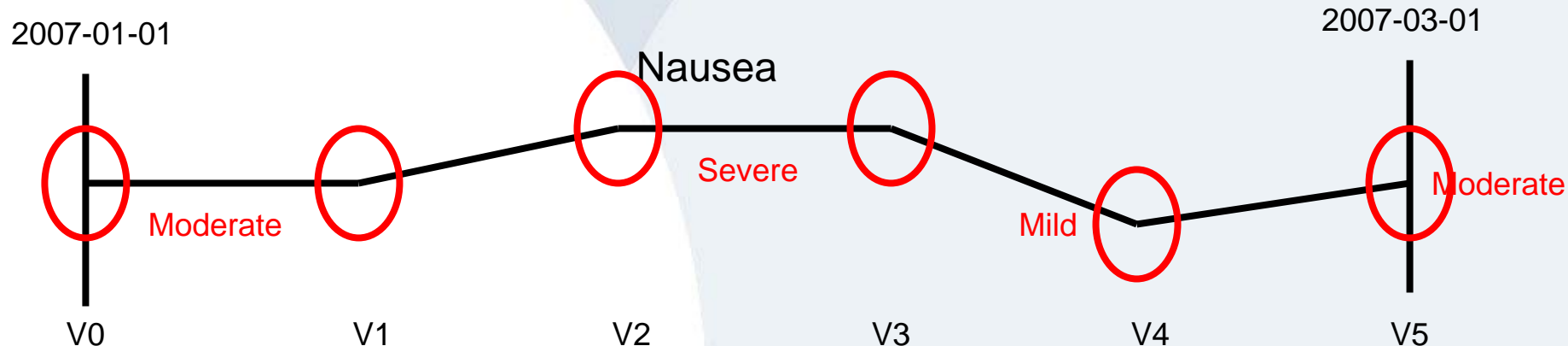
# Findings About Events

2007-01-01

2007-03-01

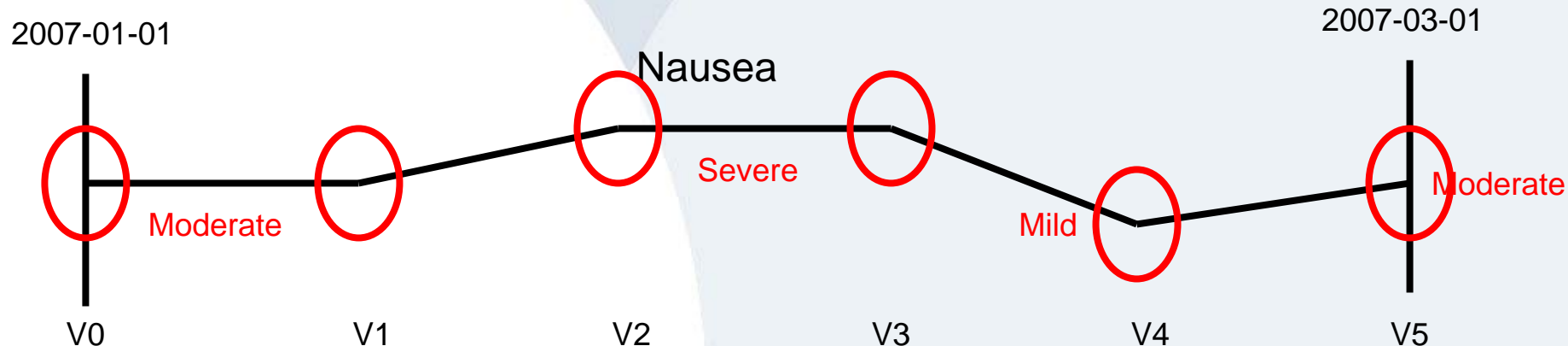


# Findings About Events



DOMAIN	USUBJID	CFSEQ	CFTESTCD	CFTEST	CFOBJ	CFORRES	VISIT
CF	123-01-01	1	SEV	Severity	Nausea	Moderate	V0
CF	123-01-01	2	SEV	Severity	Nausea	Moderate	V1
CF	123-01-01	3	SEV	Severity	Nausea	Severe	V2
CF	123-01-01	4	SEV	Severity	Nausea	Severe	V3
CF	123-01-01	5	SEV	Severity	Nausea	Mild	V4
CF	123-01-01	6	SEV	Severity	Nausea	Moderate	V5

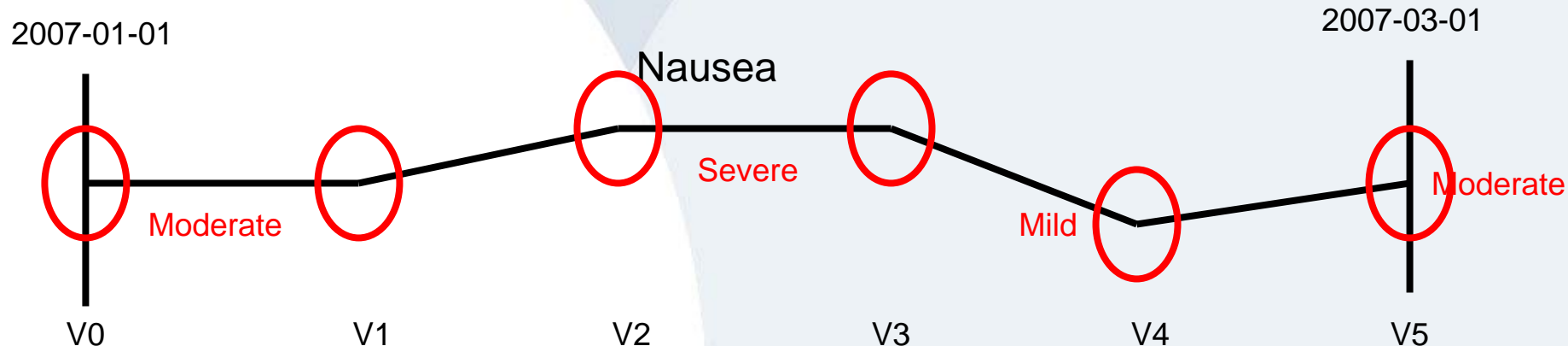
# Findings About Events



DOMAIN	USUBJID	AESEQ	AETERM	AESTDTC	AEENDTC	AESEV
AE	123-01-01	1	Nausea	2007-01-01	2007-03-01	Severe

DOMAIN	USUBJID	CFSEQ	CFTESTCD	CFTEST	CFOBJ	CFORRES	VISIT
CF	123-01-01	1	SEV	Severity	Nausea	Moderate	V0
CF	123-01-01	2	SEV	Severity	Nausea	Moderate	V1
CF	123-01-01	3	SEV	Severity	Nausea	Severe	V2
CF	123-01-01	4	SEV	Severity	Nausea	Severe	V3
CF	123-01-01	5	SEV	Severity	Nausea	Mild	V4
CF	123-01-01	6	SEV	Severity	Nausea	Moderate	V5

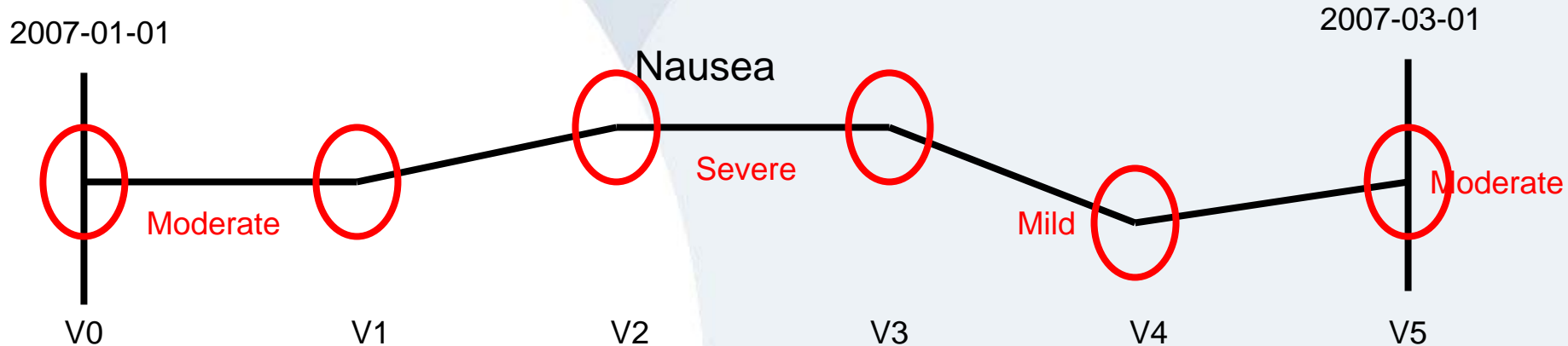
# Findings About Events



DOMAIN	USUBJID	AESEQ	AETERM	AESTDTC	AEENDTC	AESEV
AE	123-01-01	1	Nausea	2007-01-01	2007-03-01	Severe

DOMAIN	USUBJID	CFSEQ	CFTESTCD	CFTEST	CFOBJ	CFORRES	VISIT
CF	123-01-01	1	SEV	Severity	Nausea	Moderate	V0
CF	123-01-01	2	SEV	Severity	Nausea	Moderate	V1
CF	123-01-01	3	SEV	Severity	Nausea	Severe	V2
CF	123-01-01	4	SEV	Severity	Nausea	Severe	V3
CF	123-01-01	5	SEV	Severity	Nausea	Mild	V4
CF	123-01-01	6	SEV	Severity	Nausea	Moderate	V5

# Findings About Events



DOMAIN	USUBJID	AESEQ	AETERM	AESTDTC	AEENDTC	AESEV
AE	123-01-01	1	Nausea	2007-01-01	2007-03-01	Severe

DOMAIN	USUBJID	CFSEQ	CFTESTCD	CFTEST	CFOBJ	CFORRES	VISIT
CF	123-01-01	1	SEV	Severity	Nausea	Moderate	V0
CF	123-01-01	2	SEV	Severity	Nausea	Moderate	V1
CF	123-01-01	3	SEV	Severity	Nausea	Severe	V2
CF	123-01-01	4	SEV	Severity	Nausea	Severe	V3
CF	123-01-01	5	SEV	Severity	Nausea	Mild	V4
CF	123-01-01	6	SEV	Severity	Nausea	Moderate	V5

RELREC records may have been created to document the relationship.

**Get ready to review V3.1.2!**