

How to Design a Custom SDTM Domain for Nonclinical Data

Standards Roadmap Team

Abstract

Project: How to Design a Custom SDTM Domain for Nonclinical Data

Data Standards are being developed to convey complex clinical and nonclinical data in a standard form across organizations and for regulatory submissions. CDISC SEND is a nonclinical standard for general toxicology studies and carcinogenicity studies. The SEND standard is currently being expanded to include reproductive and developmental toxicology studies and safety pharmacology studies. Though study designs for many nonclinical studies have uniform design and often capture similar endpoints, some design elements and study endpoints have a more unique quality that has not been specifically considered in the standard. The science of study conduct and data interpretation allows data to be captured in various ways across studies. A standard for any given data type is a compromise of the different ways of collecting this data. This built-in flexibility in the data standard is subject to interpretation so that similar data across studies may be tabulated in multiple ways. In such cases, data loses a standard quality. The Standards Roadmap Team created a decision tree to guide a data-driven decision-making process for capturing non-standard data types in a standard dataset. The intent is to improve uniformity in the interpretation of the tenets of the SDTM standard and thereby provide a consistent way of creating a standard domain structure that specifically suits the data.

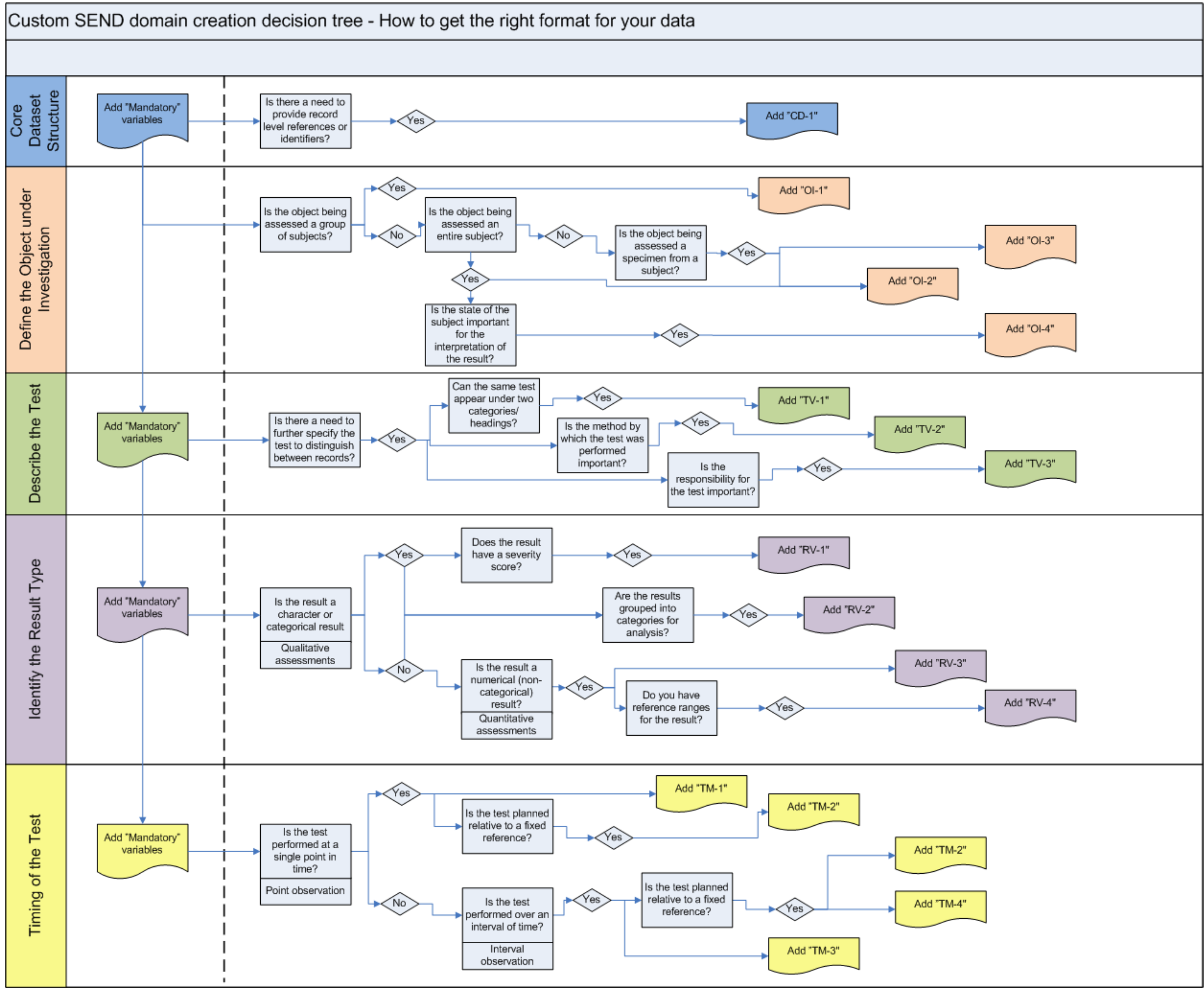
1. Instructions

Using the “How to Design a Domain” resource

The purpose of this resource is to guide users toward appropriate variables that will accommodate data types that are not specified in the standard. The primary reference document is the decision tree which will guide users on which domain variables are most appropriate for their data types.

- Users should begin at the top left of the decision tree at the category “Core Dataset Structure” and follow the “swim lane” going right across the decision tree for this category of variables. The categories on the decision tree have corresponding color-coding on the Domain Table, which is a supplemental reference document that provides more detail on the nature of variables. The questions in the decision tree will guide the user to which groups of variables are needed for their particular data.
- When users encounter a grouping of variables on the decision tree, they should then refer to the Domain Table “Core” column to determine which variables are considered required, expected or permissible for their data type.
- After completing the decisions for a category (i.e. traveling to the end of a lane) the user should go to the lane directly below the one they completed and follow the decision tree (left to right) across that lane.
- The user should consider each category (i.e. Core Dataset Structure, Object Identifiers etc.) and it is essential to use each category in a dataset. In other words, all categories of variables are essential to a dataset but the decision tree will guide the user to select the relevant groups of variables within these categories. Some guidance for selecting variables within a group is provided in the “Core (within a variable group)” column; however the fine-tuning of the domain structure should reflect the data need.

2. Decision Tree



3. Color-Coded Domain Table

Decision based adding of variables (see decision tree)			SDTM / SEND definitions and rules for the population of the variables				
Variable category	Variable group	Core (within variable group)	Variable Name	Variable Label	Type	Controlled Terms, Code list, or Format	CDISC Notes
Core Dataset structure	Mandatory	Req	STUDYID	Study Identifier	Char		Unique identifier for a study.
Core Dataset structure	Mandatory	Req	DOMAIN	Domain Abbreviation	Char	XX	Two-character abbreviation for the domain. Refer to CDISC SDTM and SEND for reserved domain codes. Domain codes starting with “X”, “Y” and “Z” are always considered.
Object Identifier	OI-2	Req	USUBID	Unique Subject Identifier	Char		Identifier used to uniquely identify a subject across all studies for all applications or observations including all locations.
Object Identifier	OI-1	Req	POOLID	Pool Identifier	Char		Identifier used to uniquely identify a subject across all studies for all applications or observations including all locations.
Core Dataset structure	Mandatory	Req	SEQ	Sequence Number	Num		Sequence number given to ensure uniqueness of subject records within a domain. May be any valid number.
Object Identifier	OI-3	Perm	REFID	Reference Identifier	Char		Optional internal or external identifier such as lab specimen ID or UUID for an ICG workflow or a medical image.
Core Dataset structure	CD-1	Perm	SPID	Sponsor-Defined Identifier	Char		Sponsor-defined reference number. Perhaps presented as an explicit line identifier or defined in the sponsor’s operational database. Example: Line number on the Lab form.
Test Variables	Mandatory	Req	TESTCD	Measurement, Test or Examination Short Name	Char	Controlled terminology is expected	Short name of the measurement, test, or examination described in --TEST. It can be used as a column name when converting a dataset from a vertical to a horizontal format. The value --TESTCD cannot be longer than 40 characters, nor can it start with a number (e.g., “1TEST” is not valid). L1TESTCD cannot contain characters other than letters, numbers, or underscores.
Test Variables	Mandatory	Req	TEST	Measurement, Test or Examination Name	Char	Controlled terminology is expected	Long name for --TESTCD. The value in L1TEST cannot be longer than 40 characters.
Test Variables	TV-1	Exp	CAT	Measurement, Test or Examination Category	Char	Controlled terminology is expected	Used to define a category of related tests across subjects.
Test Variables	TV-1	Perm	SCAT	Measurement, Test or Examination Subcategory	Char	Grouping Qualifier	A further categorization of a test category.
Object Identifier	OI-4	Perm	POS	Position of Subject during Observation	Char	(POSITION)	Position of the subject during a measurement or examination. Examples: SUPINE, STANDING, SITTING.
Result Variable	Mandatory	Exp	ORRES	Result or Finding as Collected	Char		Result of the measurement or finding as originally received or collected.
Results Variables	RV-3	Exp	ORRESU	Unit of the Original Result	Char	(UNIT)	The unit for the original result. The unit of the original result should be mapped to a synonym unit in the Controlled Terminology list.
Results Variables	RV-4	Perm	ORNRLO	Reference Range Lower Limit-Orig Unit	Char		Lower end of reference range used at the time of collection, in original units.
Results Variables	RV-4	Perm	ORNRHI	Reference Range Upper Limit-Orig Unit	Char		Upper end of reference range used at the time of collection, in original units.
Results Variables	Mandatory	Exp	STRESC	Standardized Result in Character Format	Char	Controlled terminology is expected	Contains the result value for all findings, copied or derived from --ORRES in a standard format or standard units. --STRESC should store all results or findings in character format. If results are numeric, they should also be stored in numeric format. --STRESN.
Results Variables	RV-3	Exp	STRESN	Standardized Result in Numeric Format	Num		Result of the measurement or finding as originally received or collected.
Results Variables	RV-3	Exp	STRESU	Unit of the Standardized Result	Char	(UNIT)	The unit for the original result. The unit of the original result should be mapped to a synonym unit in the Controlled Terminology list.
Results Variables	RV-4	Perm	STNRLO	Reference Range Lower Limit-Std Unit	Num		Lower end of reference range for continuous or numeric standardized results (--STRESN) represented in standardized units.
Results Variables	RV-4	Perm	STNRHI	Reference Range Upper Limit-Std Unit	Num		Upper end of reference range for continuous or numeric standardized results (--STRESN) represented in standardized units.
Results Variables	RV-4	Perm	STNRC	Reference Range for Char Rpt-Std Unit	Char		Reference range for results stored in --STRESC that are characters in ordinal or categorical scale. Example: Negative to 1 once.
Results Variables	RV-4	Perm	RIND	Reference Range Indicator	Char		Indicates where value falls with respect to reference range defined by --ORNRLO and --ORNRHI by --STNRLO and --STNRHI, or by --STNRC. Examples: NORMAL, ABNORMAL, HIGH, LOW. This should not be used to indicate biological significance.
Results Variables	RV-2	Exp	RESCAT	Result Category	Char	Controlled terminology is expected	Used to categorize the result of a finding post collection.
Result Variable	Mandatory	Perm	STAT	Completion Status	Char	(ND)	Used to indicate when a test is not done or result is missing. Should be null if a result exists in --ORRES.
Result Variable	Mandatory	Perm	REASND	Reason Not Done	Char		Describes why --STAT is NOT DONE, such as BROKEN EQUIPMENT or SICK ANIMAL.
Core Dataset structure	CD-1	Perm	XFN	External File Name	Char		Reference for an external file associated with the result.
Test Variables	TV-3	Exp	NAM	Vendor Name	Char		Name or identifier of the laboratory or vendor that provided the test result.
Object Identifier	OI-3	Exp	SPEC	Specimen Material Type	Char	(SPEC)	Defines the type of tissue, organ, or fluid specimen used for the object for the finding. Examples: GLAND, ADIPITAL, KIDNEY, VESSEL, LYMPHATIC.
Object Identifier	OI-3	Exp	ANTREG	Anatomical Region of Specimen	Char		Defines the specific anatomical or biological region of a tissue, organ specimen or the region from which the specimen was obtained, such as a section or part of what is defined in the --SPEC variable. If the anatomical region is not included in the specimen description --SPEC, it may be included in this variable. This field can be a combination of terms where needed. This field can be null if not applicable. Examples: CORTIS, MEDULLA, MUCOSA, SEROSA, SILET, ZONA, FASCICULATA, ZONA RETICULARIS, CRANIAL, MEDIAN, ACCESSORY, SPINAL, LUMBAL, PRENATAL.
Object Identifier	OI-3	Perm	SPCCND	Specimen Condition	Char		Free or standardized text describing the condition of the specimen. Example: AUTOLYZED, HEMOLYZED, ICTERIC, LIPEMIC, etc.
Object Identifier	OI-3	Perm	SPCUFL	Specimen Usability for the Test	Char	(NY)	Describes the usability of the specimen for the test. Example: N = the specimen is not usable, otherwise null.
Test Variables	TV-2	Perm	LOC	Location associated with a result or finding	Char		Anatomical location associated with the test. Example: rectal for temperature.
Object Identifier	OI-3	Perm	LAT	Specimen Laterality within Subject	Char	(LAT)	Qualifier for laterality of the specimen within the subject for paired specimens. Examples: LEFT, RIGHT, BILATERAL.
Object Identifier	OI-3	Perm	DIR	Specimen Directionality within Subject	Char	(DIR)	Qualifier for directionality of the specimen within the subject. Examples: DORSAL, BRONCHIAL.
Object Identifier	OI-3	Perm	PORTOT	Portion or Totality	Char	(PORTOT)	Qualifier for anatomical location or specimen further detailing the portion or totality which means arrangement of, or appositioning of, for example, ENTIRE, SINGLE, SEGMENT, MANY.
Test Variables	TV-2	Perm	METHOD	Method of Test or Examination	Char		Method of the test or examination.
Object Identifier	OI-4	Perm	CSTATE	Consciousness State	Char	(CSTATE)	The consciousness state of the subject at the time of measurement. Examples: CONSCIOUS, SEMI-CONSCIOUS, UNCONSCIOUS.
Result Variables	RV-3	Exp	BLFL	Baseline Flag	Char	(NY)	A baseline indicator may be used to calculate differences or changes from baseline value should be Y or null. The baseline flag is a person-defined.
Object Identifier	OI-4	Perm	EAST	Fasting Status	Char	(NY)	Indicates used to identify fasting status. Null if not relevant.
Test Variables	TV-3	Perm	EVAL	Evaluator	Char		Role of the person who provided the evaluation. Used only for results that are collected or derived data. Examples: RESPONSIBLE SCIENTIST, PRINCIPAL INVESTIGATOR, PEER REVIEWER.
Results Variables	RV-1	Perm	SEV	Severity	Char	(SEV)	Describes the severity or intensity of a particular finding.
Results Variables	TV-2	Exp	LLQ	Lower Limit of Quantitation	Num		Indicates the lower limit of quantitation for an assay. Units should be those used in --STRESU.
Results Variables	Mandatory	Perm	EXCLFL	Exclusion Flag	Char	(NY)	If the result should be excluded from all calculations, otherwise null.
Results Variable	Mandatory	Perm	REASEX	Reason for Exclusion	Char		The reason the result should be excluded from all calculations. Used only when --EXCLFL is Y.
Timing Variables	TM-1	Exp	VISITDY	Planned Study Day	Num		This is the planned study day of collection.
Timing Variables	TM-1	Exp	DTC	Date/Time of Test	Char	ISO 8601	Date/time of the measurement, test or observation in ISO 8601 format.
Timing Variables	TM-3	Exp	STDTCT	Start Date/Time of Test	Char	ISO 8601	Date/time of the start of the measurement, test, or observation in ISO 8601 format. Should be populated only for continuous sample collection.
Timing Variables	TM-3	Perm	ENDTCT	End Date/Time of Test	Char	ISO 8601	Date/time of the end of the measurement, test, or observation in ISO 8601 format. Should be populated only for continuous sample collection.
Timing Variables	TM-1	Exp	SDY	Study Day of Test	Num		Study day of the measurement, test, or observation, in integer days. The algorithm for calculation must be relative to the sponsor-defined R1STDTCT variable in the Domain Table (DM domain).
Timing Variables	TM-3	Exp	STDY	Study Day of Start of Test	Num		Study day of the start of the measurement, test, or observation, in integer days. The algorithm for calculation must be relative to the sponsor-defined R1STDTCT variable in the Domain Table (DM domain).
Timing Variables	TM-3	Perm	ENDY	Study Day of End of Test	Num		Study day of the end of the measurement, test, or observation, in integer days. The algorithm for calculation must be relative to the sponsor-defined R1STDTCT variable in the Domain Table (DM domain).
Timing Variables	TM-3	Perm	DUR	Duration of Test	Char	ISO 8601	Collected duration of a measurement, test, or observation, represented in ISO8601 format. This should not be used to report a derived duration.
Timing Variables	TM-2	Exp	TPT	Planned Time Point Name	Char		Text description of the planned time of the measurement, test, or observation.
Timing Variables	TM-2	Exp	TPTNUM	Planned Time Point Number	Num		This may be represented as an elapsed time relative to a fixed reference point, such as time of last dose. See R1PTNUM and L1PTREF. Examples: Start, 5 min post.
Timing Variables	TM-2	Exp	ELTM	Planned Elapsed Time From Time Point Ref	Char	ISO 8601	Nominal version of --TPT used in sorting.
Timing Variables	TM-2	Exp	TPTREF	Time Point Reference	Char		Planned elapsed time (in ISO 8601 format) relative to a planned fixed reference (L1PTREF). This variable is useful where there are repetitive measures. Not a clock time or a date time variable. Represented as an ISO 8601 duration. Examples: “-PT1M” to represent the period of 15 minutes prior to the reference point indicated by L1PTREF, or “PT1H” to represent the period of 1 hour after the reference point indicated by L1PTREF.
Timing Variables	TM-2	Exp	RFTDTC	Date/Time of Reference Time Point	Char	ISO 8601	Name of the fixed reference point referred to by --ELTM, if used for L1PTNUM, and L1PTREF. Examples: PREVIOUS DOSE, PREVIOUS MEAL.
Timing Variables	TM-4	Perm	EVLINT	Evaluation Interval	Char	ISO 8601	Duration of interval associated with an observation such as a finding --TESTCD, represented in ISO 8601 character format. Example: P2M to represent a period of the past 2 months at the evaluation interval for a question from a questionnaire, such as SF-36.
Timing Variables	TM-4	Perm	STINT	Planned Start of Assessment Interval	Char		The start of a planned evaluation or assessment interval relative to the Time Point Reference.
Timing Variables	TM-4	Perm	ENINT	Planned End of Assessment Interval	Char		The end of a planned evaluation or assessment interval relative to the Time Point Reference.

4. Result

“Unmodeled” data captured using both a streamlined approach and existing data standards

Acknowledgements: Gitte Frausing, Novo Nordisk A/S, Denmark; Debra Oetzman, Covance Laboratories, Inc, Madison, WI; Kathy Powers, Pfizer; Sarah Obbers, Janssen Research & Development, Beerse, Belgium; Anisa Scott, JMP Life Sciences, SAS Institute, Cary, NC; Lynda Sands, GlaxoSmithKline, King of Prussia, PA; Donna Danduone, Instem, Conshohocken, PA; Robert T. Dorsam, FDA, CDER, USA

Note: The opinions expressed in this poster are those of the authors and do not necessarily represent the opinions of their respective companies or organizations.