VERISTAT

ACHIEVING BALANCE IN CLINICAL TRIAL DATABASE DESIGN:

CONSIDERATIONS FOR ENHANCING CLINICAL TRIAL EFFICIENCY

WEBINAR HELD ON DECEMBER 10, 2019

An Impactful CRO,

Committed to Making a Difference

AGENDA

- GENERAL DATABASE BUILD CONSIDERATIONS
- CLINICAL CONSIDERATIONS
- SDTM PRINCIPLES
- FIND THE BALANCE – REAL WORLD EXAMPLES
- QUESTIONS

GENERAL DATABASE BUILD CONSIDERATIONS

Food for thought before you start…

- **Your stakeholders** - who are they? What is important to them?
- **What are your short-term goals?** What will impact these (# sites, end users, duration of study, etc.)?
- **What are your long-term goals?** Will you submit this data to the regulatory agencies?

CONSIDERATIONS FOR DATABASE DESIGN

Who are your stakeholders?

Immediate Stakeholders/ DB Users:

- Site/Study Coordinators (SCs)
- Principle Investigators (PIs)
- Clinical Research Associates (CRAs)
- Read-only roles
- Data Managers

Downstream Stakeholders:

- Biostatisticians
- ADaM and SDTM Programmers
- Medical Writers

The needs of everyone are different when they look at Case Report Form:

* Were any adverse events experienced?	○ Yes ○ No	AE.AEYN
* Adverse Event		AE.AETERM
* Start Date		AE.AESTDAT
* Start Time	:	AE.AESTTIM (HH24:MI)
* Ongoing?	○ Yes ○ No	AE.ONGO
* End Date		AE.AEENDAT
* End Time	:	AE.AEENTIM (HH24:MI)
* Outcome		AE.AEOUT
* Grade		AE.AETOXGR
* Is this an injection site reaction?		AE.AEINJ
* If yes, location of reaction		AE.INJLOC

CONSIDERATIONS FOR DATABASE DESIGN

What are your goals and priorities for this database?

SHORT-TERM

- **Goals:**
  - Build a database that captures clinical data for this protocol that is easy to use?
  - Fastest path to determine protocol endpoints?
  - Easiest to understand for the site coordinators so data entry is quick and accurate?
  - Ability to quickly review data in reports in real-time?
- **Key Considerations:** *speed, price of DB build, "usability" for Clinical stakeholders*

LONG-TERM

- **Goals:**
  - Aggregate data as quickly as possible to produce outputs for ad-hoc or final analysis purposes?
  - Submit data as part of a marketing application to the regulatory bodies (NDA, BLA, MAA)?
  - Determine proof of concept to build a program around that *will* be used for a submission?
- **Key Consideration:** *SDTM Compliance, accuracy, turnaround time, overall price*

SOME REAL-WORLD EXAMPLES OF HOW THIS PLAYS OUT

Sponsor knows data will never be used in a submission, we helped them decide that 100% SDTM compliance isn't a top priority. Built the database to be the most user-friendly as possible to analyze results quickly and access data in real-time.

Sponsor wants to apply as many SDTM principles to the CRFs as possible to minimize time spent later, on migration efforts. They know these CRFs might not be the most user-friendly, but it is a single center study, so sponsor made decision to invest significant time in training site personnel on the nuances of the forms.

Sponsors who didn't focus on SDTM compliance now need significant time, money, and effort spent on migration efforts which were not accounted for in timelines or budgets.

- This was the last study needed in a marketing application submission to the regulatory bodies and timelines were gatekeepers for final timelines for submission.

SO, WHAT NOW?

Now that we've heard about some key considerations, let's hear directly from some folks who are stakeholders and what actually *IS* important to each of them.

- Naveen will share what it important to sites, CRAs, and Clinical Operations from his perspective and experience
- Nirav will share more on SDTM – why is it important, and how do we apply it?
- We'll come together to share some more real-world examples of how we can apply these concepts to find a balance

DATABASE CONSIDERATIONS FOR CLINICAL OPERATIONS

- CRA review of the draft database and/or data entry guidelines
- CRA input of source (if being provided to the site(s))
- Important for CRA's to review/train sites on data entry at the SIV
  - Consistent date formatting/data entry
  - Legitimate auto-queries
  - Ease of addressing queries

WHAT IS STUDY DATA TABULATION MODEL (SDTM)?

Why Is It Important To Understand in Database Design?

Study Data Tabulation Model (SDTM) defines a standard structure for study data tabulations that are to be submitted as part of a product application to a regulatory authority such as the United States Food and Drug Administration (FDA).

- Data tabulation datasets are one of four ways to represent the human subject Case Report Tabulation (CRT) and equivalent animal data submitted to the FDA

- The standardized datasets will support
  - Repository
  - Standard Review Tools

- Streamlining the flow of data from collection through submission, and facilitating data interchange

OBSERVATIONS AND VARIABLES

Variables can be classified into five major roles:

Identifier variables

Topic variables

Timing variables

Qualifier variables

Rule *variables*

SDTM STANDARD DOMAIN MODELS

SPECIAL-PURPOSE DATASETS

- Domain datasets
- Trial Design Model (TDM) datasets
- Relationship datasets

FINDINGS ABOUT:

- Findings About (FA)
- Skin Response (SR)

GENERAL OBSERVATION CLASSES

Interventions General Observation Class:	Events General Observation Class:	Findings General Observation Class:
• Concomitant Medications (CM) • Exposure as Collected (EC) • Exposure (EX) • Substance Use (SU) • Procedures (PR)	• Adverse Events (AE) • Clinical Events (CE) • Disposition (DS) • Protocol Deviations (DV) • Healthcare Encounters (HO)	• Drug Accountability (DA) • Death Details (DD) • ECG Test Results (EG) • Inclusion/Exclusion Criterion Not Met (IE) • Immunogenicity Specimen Assessments (IS) • Laboratory Test Results (LB) • Microbiology Specimen (MB) • Microscopic Findings (MI) • Morphology (MO) • Microbiology Susceptibility Test (MS) • PK Concentrations (PC) • PK Parameters (PP) • Physical Examination (PE) • Questionnaires (QS) • Reproductive System Findings (RP) • Disease Response (RS) • Subject Characteristics (SC) • Subject Status (SS) • Tumor Identification (TU) • Tumor Results (TR) • Vital Signs (VS)

CLINICAL DATA ACQUISITION STANDARDS HARMONIZATION (CDASH)

The **Clinical Data Acquisition Standards Harmonization (CDASH) Model** describes the foundational structure for the organization, naming, and description of variable and associated attributes to support data collection in clinical trials

- Metadata to help facilitate mapping collected data to their respective SDTM Implementation Guide (SDTMIG) variables
- The CDASH Model organizes data into classes, which represent meaningful groupings of data in clinical research
- It defines CDASH metadata for identifier variables, timing variables, general observation class variables (Events, Interventions, and Findings), domain-specific variables, and special-purpose domain variables, (e.g., DM, CO).

SDTM VS CDASH: WHY YOU NEED BOTH?

SDTM	Why the Difference?	CDASH
Show me the data, not lack of data	SDTM assumes that if there is no record then nothing happened. This works but only if it was checked in data capture, which requires a question and record (e.g., Were there any AEs?)	Absence of evidence is not evidence of absence: must check that missing data is missing
Machine-readable data: -ISO 8601 Dates/Times: 1 variable, YYYY-MMDDThh:mm:ss -Duration: P1M3D	SDTM machine-readable formats for variables such as dates are good for data reusability but are not userfriendly for data capture. Sites recording data in unfamiliar formats increases risk of errors	Human-readable: -Dates/Times: 2 or more variables, DD-MMM-YYYY, HH:MM:SS -Duration: 1 month, 3 days
Variables must be in order by domain; non-standard variables are stored in different datasets (e.g., FA, SUPP--)	Domain-driven organization is critical for standard tools, but data must make sense to the site. This can mean to split domains across CRFs and CRFs across domains, and not split custom and standard variables	Data structure harmonized with SDTM but variables can be arranged to make data capture easier.
Collected relationships between data are represented in RELREC, a separate dataset	RELREC is based on collected data, but data is not captured like that. Entering line numbers in the related datasets is simpler, requiring no derivations (e.g., adding AE line # to related con med)	Links among records are explicit (e.g., this AE related to that CM), or implicit (e.g., AE severity changes going into FA) in data collection
Findings data must be in a normalized or vertical structure; answers are already in variables	Normalized structures can store new tests without changing dataset structures, but most EDC systems can't do this; also, different tests in a domain may need different controlled terms (e.g., different answers for different questions in a survey)	Findings data may be horizontal, letting each test have a different code list; SDTM CT is used for variable names & CRF prompts
Metadata centers on tabulations, e.g., variable labels and roles	SDTM labels identify tabulation data. CDASH has question texts and prompts designed to elicit clear responses on CRFs. CRF instructions convey SDTM and CDASH assumptions in a data capture context	Metadata includes capture needs, e.g., question text/prompt, CRF completion instructions

Source: https://www.cdisc.org/system/files/all/standard/CDASH/CDISC_SDTM_and_CDASH_You_Need_Both.pdf

SO, WHAT DO WE DO WITH ALL THIS INFORMATION?

Building CRFs perfectly SDTM compliant often requires compromise. The CRFs might not be as "pretty" or as user friendly. They may require more time and effort by the site coordinators to use the forms.

Building CRFs that fit the wish-list of a site coordinator likely will not conform to SDTM standards. Additional time, effort, and money will be spent performing a mapping to SDTM datasets later as a result.

Our Suggestion: Find the Balance!

Don't let SDTM be an afterthought but find compromise on the areas where it can be easily applied vs. the areas where end "usability" might take precedence.

- In general, generic CRFs can have SDTM principles applied quite easily (Demographics, Informed Consent, Adverse Events etc.)
- Study specific forms are oftentimes the place to prioritize the functionality of the form for end users. Focus on ensure the field or question is clearly stated and an accurate response can be given.

LET'S LOOK AT SOME EXAMPLES

Formatting Date Fields- Each stakeholder may want a date field formatted differently

Site/PI/CRA:

DD/MM/YYYY

Data Manager/Biostatistician:

DD-MMM-YYYY

SDTM Programmer:

ISO8601 YYYY-MM-DD THH:mm:ss

How do we pick?

- EDC system may not offer all these options
- Follow CDASH standard, if ISO format not available and want to be as close to compliant as possible (CDASH formatting is DD-MMM-YYYY)
- Train site personnel on format if it is not the usual for them. Include in training documents like CCGs and HelpText options within the database build.

What happens if we don't pick the SDTM format?

- All formats will get converted to the SDTM format
- Partial dates will be time consuming to convert. Parse out each piece of the date into a separate variable (one for day, one for month, one for year) and then convert into ISO format

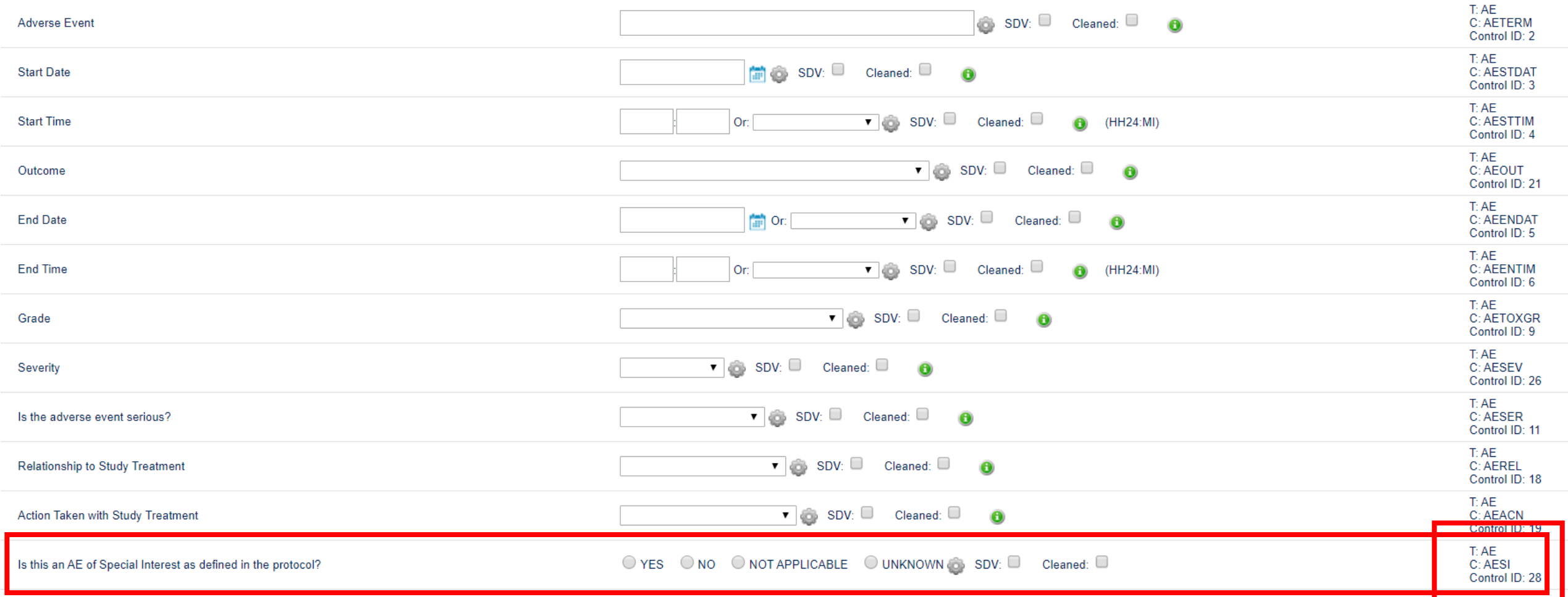
LET'S LOOK AT SOME EXAMPLES

Adding a field that is not standard to a CRF

CDISC has developed Implementation Guides that show you all the available domains, variable names, and controlled terminologies (for things like codelists)

Version	Seq. For Ord	Observation Cla	Domain Pre	Variable Name (minus domain prefix)	Variable Nan	Variable Label	Ty	Controlled Terms or Format	Role	CDISC Notes (for domains) Description (for General Classes)	Co
SDTMIG 3.3	26	Events	AE	SEV	AESEV	Severity/Intensity	Char	(AESEV)	Record Qualifier	The severity or intensity of the event. Examples: "MILD", "MODERATE", "SEVERE".	Perm
SDTMIG 3.3	27	Events	AE	SER	AESER	Serious Event	Char	(NY)	Record Qualifier	Is this a serious event? Valid values are "Y" and "N".	Exp
SDTMIG 3.3	28	Events	AE	ACN	AEACN	Action Taken with Study Treatment	Char	(ACN)	Record Qualifier	Describes changes to the study treatment as a result of the event. AEACN is specifically for the relationship to study treatment. AEACNOTH is for actions unrelated to dose adjustments of study treatment. Examples of AEACN values include ICH E2B values: "DRUG WITHDRAWN", "DOSE REDUCED", "DOSE INCREASED", "DOSE NOT CHANGED", "UNKNOWN" or "NOT APPLICABLE".	Exp
SDTMIG 3.3	29	Events	AE	ACNOTH	AEACNOTH	Other Action Taken	Char		Record Qualifier	Describes other actions taken as a result of the event that are unrelated to dose adjustments of study treatment. Usually reported as free text. Example: "TREATMENT UNBLINDED. PRIMARY CARE PHYSICIAN NOTIFIED".	Perm
SDTMIG 3.3	30	Events	AE	REL	AEREL	Causality	Char		Record Qualifier	Records the investigator's opinion as to the causality of the event to the treatment. ICH E2A and E2B examples include "NOT RELATED", "UNLIKELY RELATED", "POSSIBLY RELATED", "RELATED". Controlled Terminology may be defined in the future. Check with regulatory authority for population of this variable.	Exp
										Records the investigator's opinion as to whether	

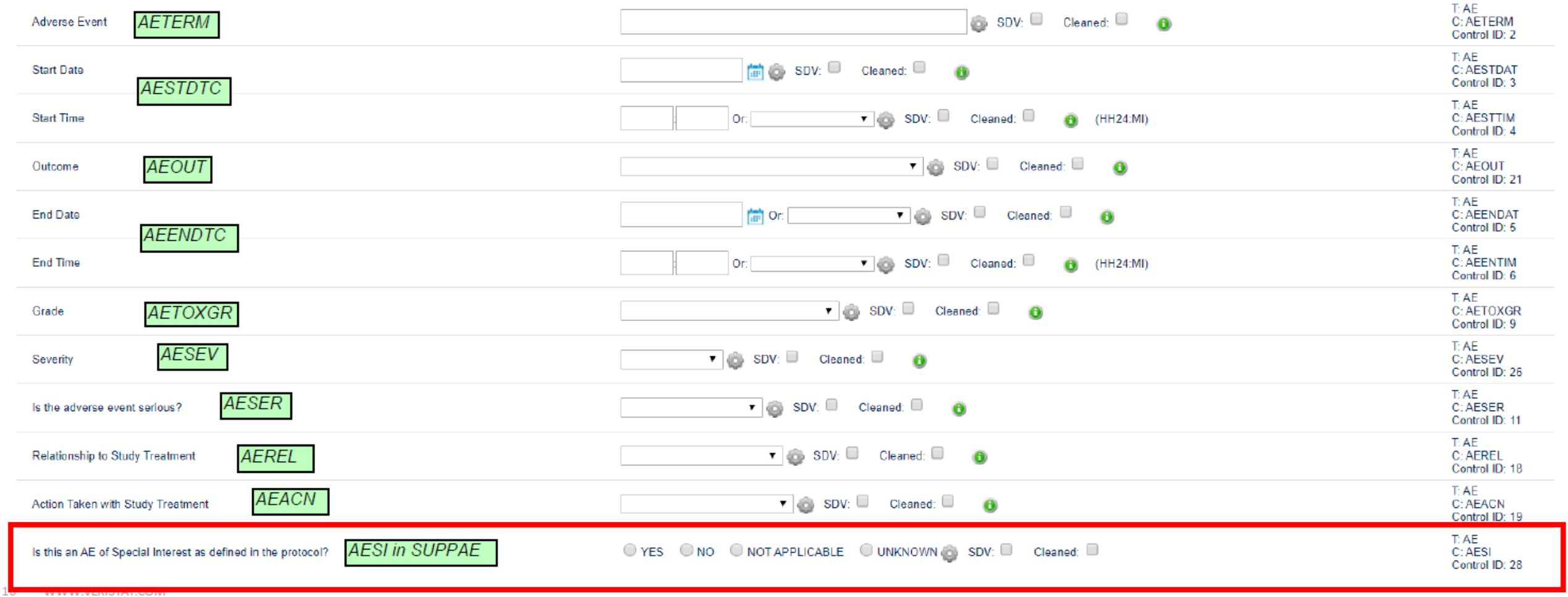
EXAMPLE: ADDING A NONSTANDARD VARIABLE



EXAMPLE: ADDING A NONSTANDARD VARIABLE

Adding a field that is not standard to a CRF

SDTM annotations added below to show where this variable gets mapped to in SDTM datasets



EXAMPLE: AMENDING CONTROLLED TERMINOLOGY

Units for Concomitant Medications

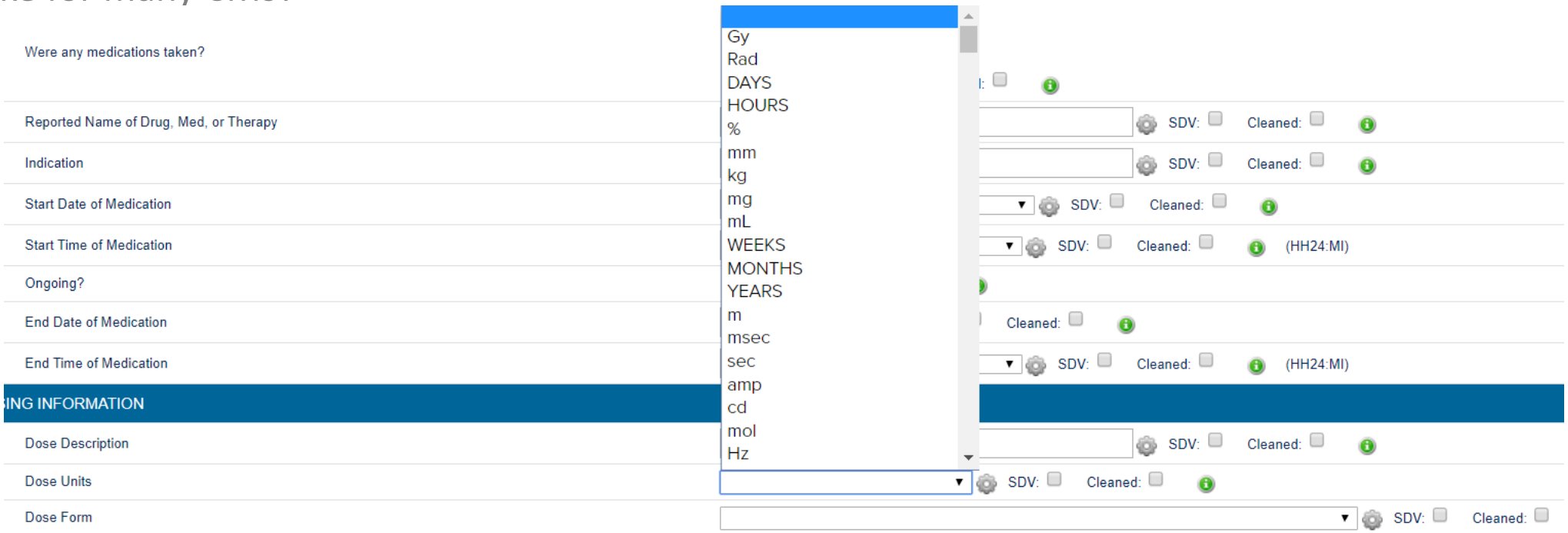
We want to use the standard controlled terminology code list as the options for Unit on our ConMed CRF so that we're prioritizing SDTM compliance.

Code	Codelist Code	Codelist Extensible (Yes/No)	Codelist Name	CDISC Submission Value	CDISC Preferred Term	CDISC Synonym(s)	CDISC Definition
C71620	C71620	Yes	Unit	UNIT	UNIT	Unit	Terminology codelist used for units within CDISC.
C18063	C71620		Unit	Gy	Gy	Gray	A unit of absorbed radiation dose. One gray is equal to an absorbed dose of one joule per kilogram of matter, or to 100 rads.(NCI)
C18064	C71620		Unit	Rad	Rad	Rad	The special unit for absorbed radiation dose, which is the amount of energy from any type of ionizing radiation (e.g., alpha, beta, gamma, neutrons, etc.) deposited in any medium (e.g., water, tissue, air). A dose of one rad means the absorption of 100 ergs per gram of absorbing tissue. One rad is equal to 0.01 gray.(NCI)
C25301	C71620		Unit	DAYS	DAYS	Day	The time for Earth to make a complete rotation on its axis; ordinarily divided into twenty-four hours. This also refers to a specific day. (NCI)
C25529	C71620		Unit	HOURS	HOURS	Hour	A unit measure of time equal to 3,600 seconds or 60 minutes. It is approximately 1/24 of a median day.(NCI)
C25613	C71620		Unit	%	%	Percentage	A fraction or ratio with 100 understood as the denominator.(NCI)
C28251	C71620		Unit	mm	mm	Millimeter	A unit of measure equal to one thousandth of a meter.(NCI)
C28252	C71620		Unit	kg	kg	Kilogram	The basic SI unit of mass. It is defined as the mass of an international prototype in the form of a platinum-iridium cylinder kept at Sevres in France. It is the only basic unit still defined in terms of a material object, and also the only one with a prefix [kilo] already in place. A kilogram is equal to 1,000 grams and 2.204 622 6 pounds. (NCI)
C28253	C71620		Unit	mg	mg	Milligram	The unit of mass equal to one thousandth of a gram or 1000 micrograms. One milligram equals approximately 0.015432 grain or 35.274 x 10E-6 ounce.(NCI)
C28254	C71620		Unit	mL	mL	Milliliter	The unit of volume equal to one thousandth of a liter, one cubic centimeter, 10E-6 cubic meter, or approximately to 0.061 023 7 cubic inch.(NCI)
C29844	C71620		Unit	WEEKS	WEEKS	Week	Any period of seven consecutive days.(NCI)
C29846	C71620		Unit	MONTHS	MONTHS	Month	One of the 12 divisions of a year as determined by a calendar. It corresponds to the unit of time of approximately to one cycle of the moon's phases, about 30 days or 4 weeks.(NCI)
C29848	C71620		Unit	YEARS	YEARS	Year	The period of time that it takes for earth to make a complete revolution around the sun, approximately 365 days; a specific one year period.(NCI)

EXAMPLE: AMENDING CONTROLLED TERMINOLOGY

Units for Concomitant Medications

Do we really want ALL these options available on the CRF (there's 310 options!)? The site will need to scroll/sort through this for a long time just to enter/select a unit! Can you imagine the time this would take for many CMs?



- Luckily, this codelist is *extensible*. This means we can remove the units we do not want to use.

EXAMPLE: AMENDING CONTROLLED TERMINOLOGY

Units for Concomitant Medications

We want to use the standard controlled terminology code list as the options for Unit on our ConMed CRF so that we're prioritizing SDTM compliance.

Code	Codelist Code	Codelist Extensible (Yes/No)	Codelist Name	CDISC Submission Value	CDISC Preferred Term	CDISC Synonym(s)	CDISC Definition
C71620	C71620	Yes	Unit	UNIT	UNIT	Unit	Terminology codelist used for units within CDISC.
C18063	C71620		Unit	Gy	Gy	Gray	A unit of absorbed radiation dose. One gray is equal to an absorbed dose of one joule per kilogram of matter, or to 100 rads.(NCI)
C18064	C71620		Unit	Rad	Rad	Rad	The special unit for absorbed radiation dose, which is the amount of energy from any type of ionizing radiation (e.g., alpha, beta, gamma, neutrons, etc.) deposited in any medium (e.g., water, tissue, air). A dose of one rad means the absorption of 100 ergs per gram of absorbing tissue. One rad is equal to 0.01 gray.(NCI)
C25301	C71620		Unit	DAYS	DAYS	Day	The time for Earth to make a complete rotation on its axis; ordinarily divided into twenty-four hours. This also refers to a specific day. (NCI)
C25529	C71620		Unit	HOURS	HOURS	Hour	A unit measure of time equal to 3,600 seconds or 60 minutes. It is approximately 1/24 of a median day.(NCI)
C25613	C71620		Unit	%	%	Percentage	A fraction or ratio with 100 understood as the denominator.(NCI)
C28251	C71620		Unit	mm	mm	Millimeter	A unit of measure equal to one thousandth of a meter.(NCI)
C28252	C71620		Unit	kg	kg	Kilogram	The basic SI unit of mass. It is defined as the mass of an international prototype in the form of a platinum-iridium cylinder kept at Sevres in France. It is the only basic unit still defined in terms of a material object, and also the only one with a prefix [kilo] already in place. A kilogram is equal to 1,000 grams and 2.204 622 6 pounds. (NCI)
C28253	C71620		Unit	mg	mg	Milligram	The unit of mass equal to one thousandth of a gram or 1000 micrograms. One milligram equals approximately 0.015432 grain or 35.274 x 10E-6 ounce.(NCI)
C28254	C71620		Unit	mL	mL	Milliliter	The unit of volume equal to one thousandth of a liter, one cubic centimeter, 10E-6 cubic meter, or approximately to 0.061 023 7 cubic inch.(NCI)
C29844	C71620		Unit	WEEKS	WEEKS	Week	Any period of seven consecutive days.(NCI)
C29846	C71620		Unit	MONTHS	MONTHS	Month	One of the 12 divisions of a year as determined by a calendar. It corresponds to the unit of time of approximately to one cycle of the moon's phases, about 30 days or 4 weeks.(NCI)
C29848	C71620		Unit	YEARS	YEARS	Year	The period of time that it takes for earth to make a complete revolution around the sun, approximately 365 days; a specific one year period.(NCI)

WRAPPING UP | KEY TAKEAWAYS

- ✓ Think about this *before* you embark on your database build journey
- ✓ Weigh the pros and cons of satisfying each stakeholders needs against your overall short and long-term goals
- ✓ Identify the “quick wins” – places where you can easily satisfy the needs of one stakeholder without much effort (ex: using Controlled Terminology codelists but pairing down extensible ones)
- ✓ Seek advice from experts who can see the big picture

QUESTIONS ??

CONTACT INFORMATION

Kathleen Boruchowski, MS
Director of Data Management, Veristat

Kathleen.boruchowski@veristat.com

Naveen Minocha
Senior Clinical Research Associate, Veristat

Naveen.minocha@veristat.com

Nirav Patel, MS
Senior Manager, Statistical Programming, Veristat

nirav.patel@veristat.com