accenture

consulting | technology | outsourcing

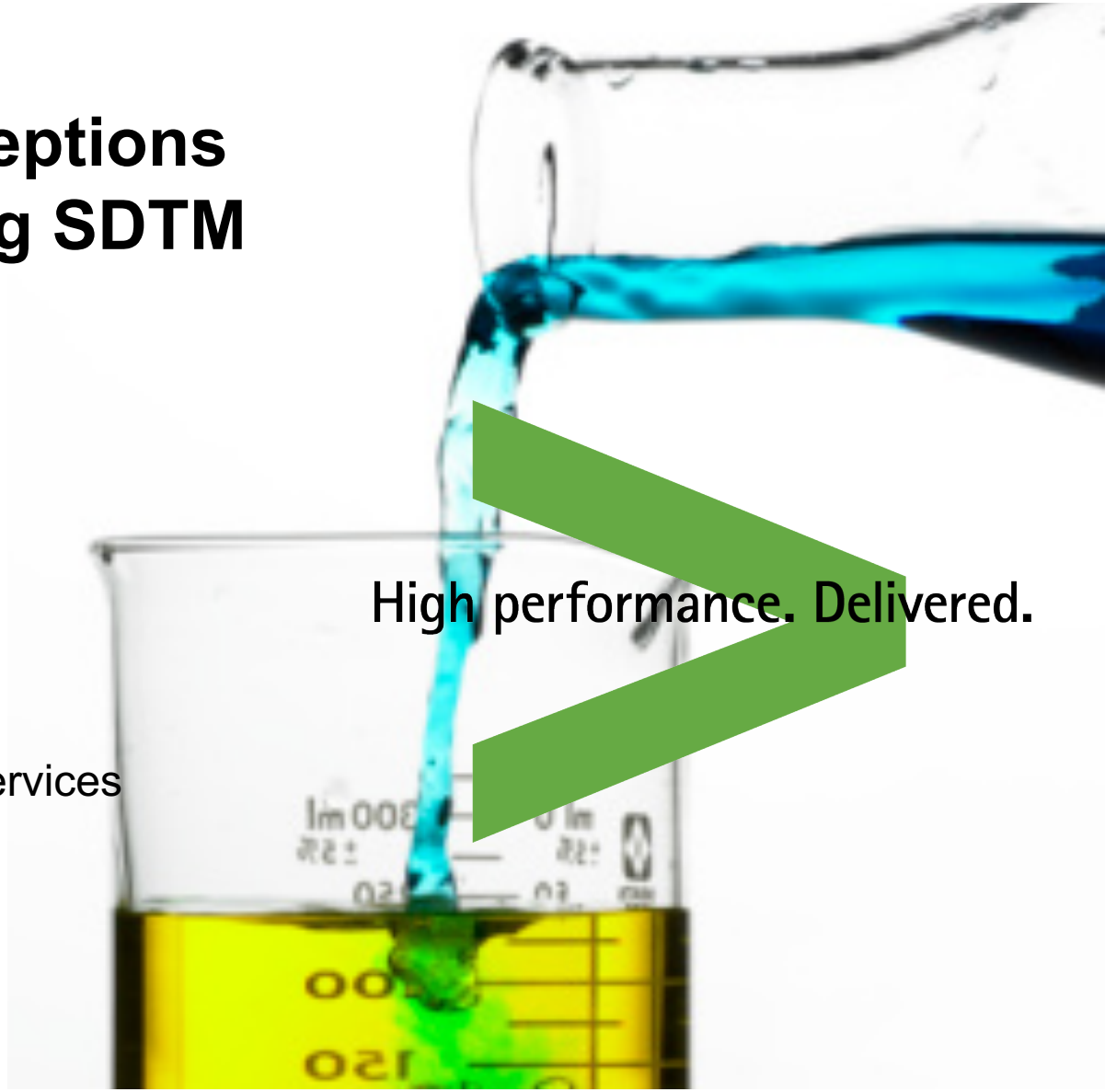
Common Misconceptions when Implementing SDTM

Jerry Salyers
Senior Consultant
Data Standards Consulting

Fred Wood
Senior Manager and Lead
Data Standards Consulting

Accenture Accelerated R&D Services

Boston PhUSE SDE
April 27 2017



High performance. Delivered.

Agenda

- Introduction
- Data Flow
- Compliance
- Extra Data
- Supplemental Qualifiers and Findings About
- Timing Variables
- Exposure Data
- Trial Design

Introduction

- In theory, there is no difference between theory and practice. In practice, there is.

-*Yogi Berra*

In Theory: Companies Should Have and Enforce Data Standards

The Reality:

- Although waning, some believe:
  - Data standards stifle creativity
  - The CRF is a vehicle for expressing creativity
- Many companies outsource legacy-data conversion to multiple CROs/vendors
  - Each has its own way of interpreting the SDTM/SDTMIG
    - Data that has various degrees of compliance
    - Data may not be able to be integrated without additional time and effort
- Many companies do not have standard data-transfer specifications

In Theory: Every Variable Should Have an Unambiguous Meaning

The Reality

- While many variables do, several do not.
- --LOC variable: Can mean different things depending on observation class; Anatomical site of dose administration, anatomical site relevant to event, “Location” used for measurement
- The BRIDG has helped to remediate this ambiguity.
  - PerformedObservation.targetApproachSite
  - PerformedObservation.targetAnatomicSite
  - PerformedObservationResult.targetAnatomicSite

Myth: The SDTMIG Dictates What to Collect and Submit

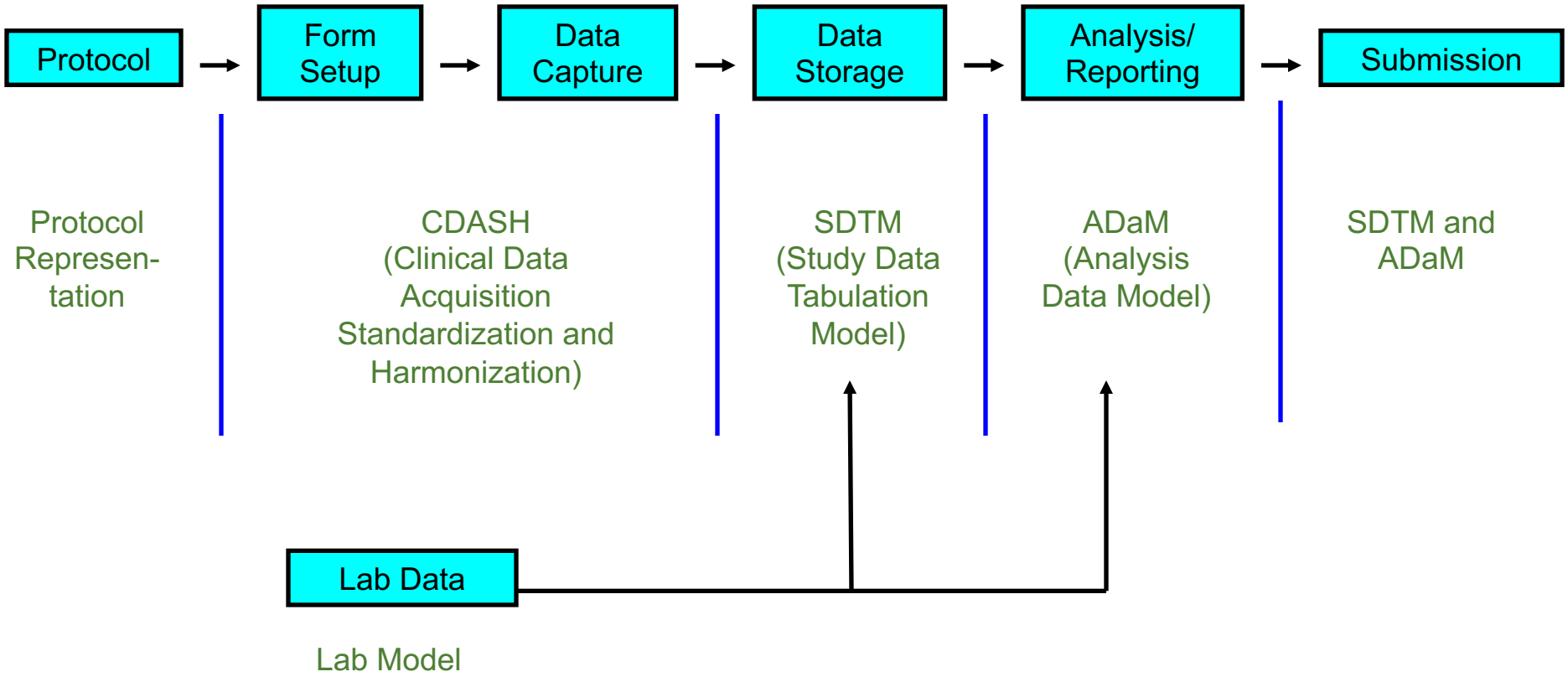
Reality

Science and regulation determine the data needed for a submission.

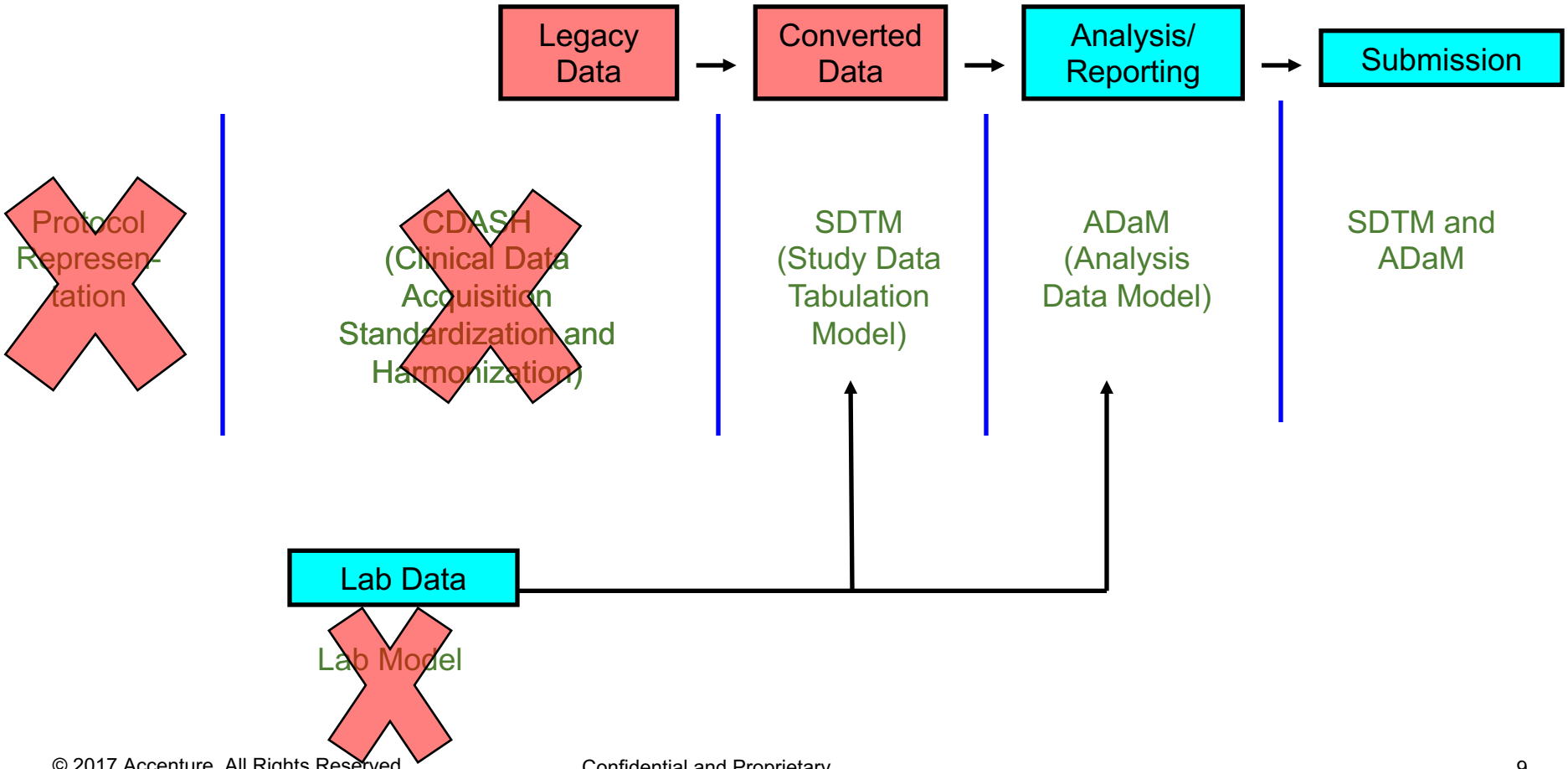
Agenda

- Introduction
- Data Flow
- Compliance
- Extra Data
- Supplemental Qualifiers and Findings About
- Timing Variables
- Exposure Data
- Trial Design

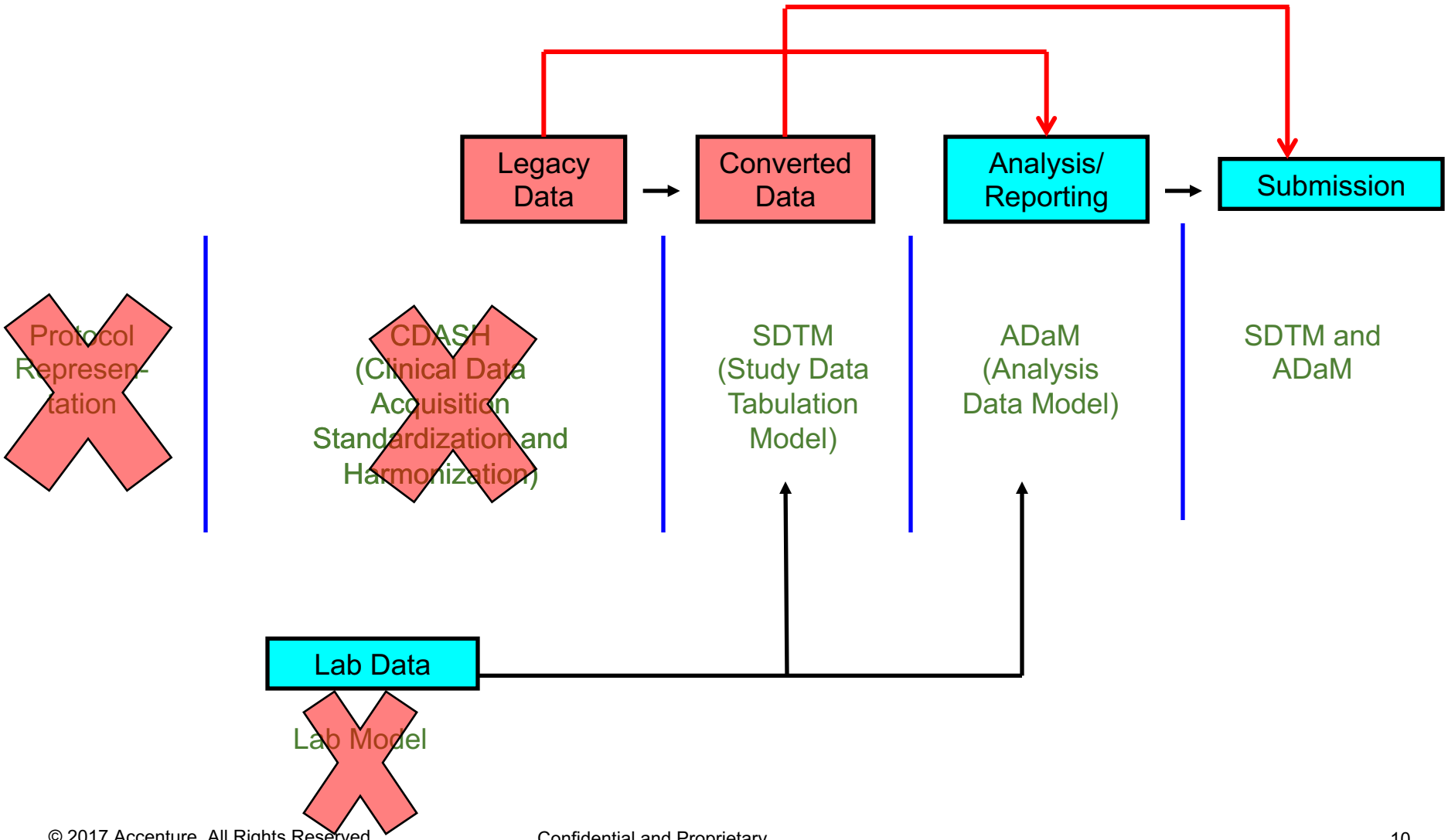
The Ideal: CDISC Standards in Clinical Data Flow



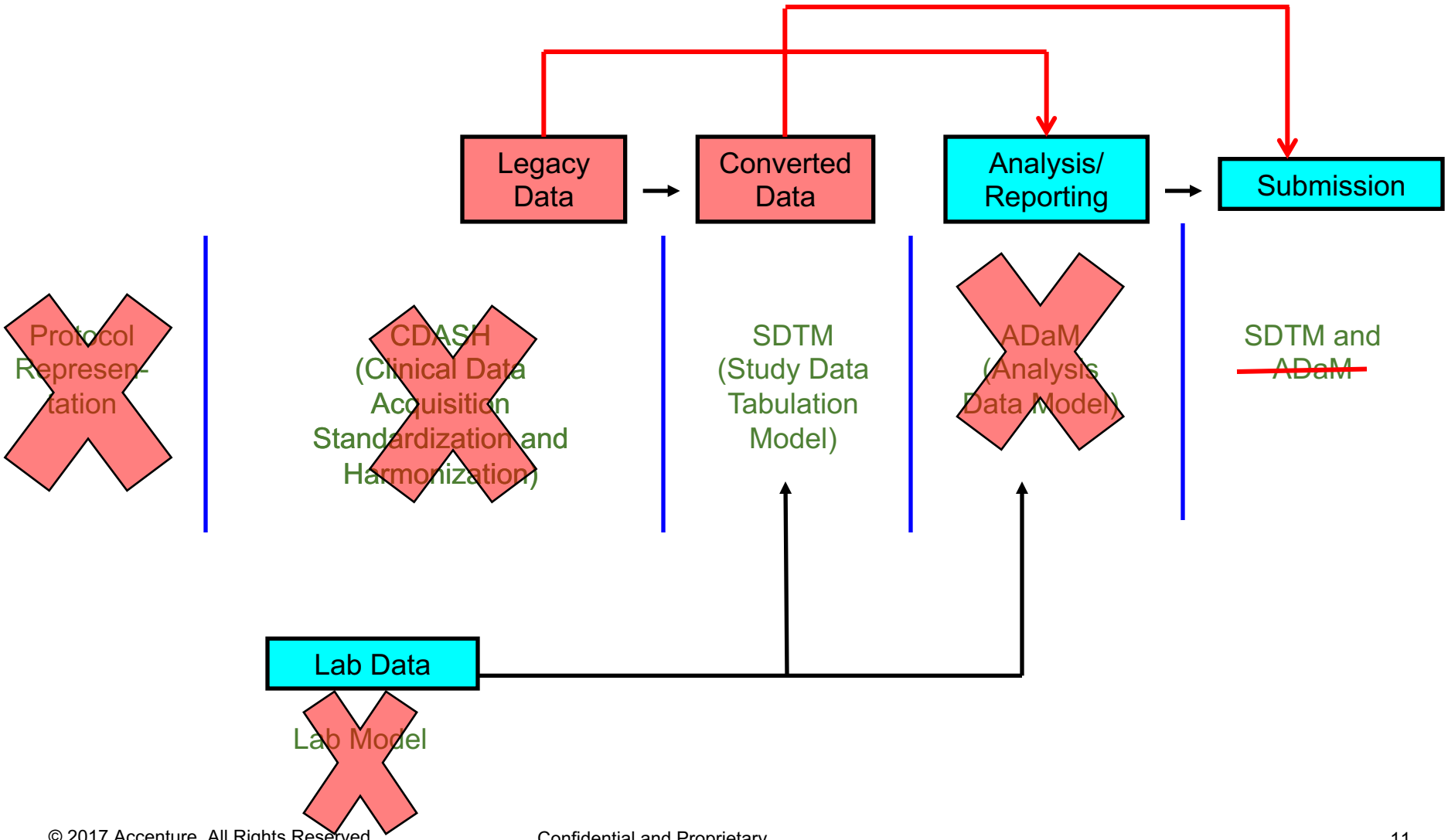
Practice 1: CDISC Standards in Clinical Data Flow



Practice 2: CDISC Standards in Clinical Data Flow



Practice 3: CDISC Standards in Clinical Data Flow



Agenda

- Introduction
- Data Flow
- Compliance
- Extra Data
- Supplemental Qualifiers and Findings About
- Timing Variables
- Exposure Data
- Trial Design

The Ideal: Validation for Compliance Can Be Completely Automated

The Reality

- Automated validation checks play an important role.
  - Not every SDTM/SDTMIG convention can be represented in computer-executable code.
- The variability in clinical trials data makes it virtually impossible for a finite number of checks to identify all potential compliance issues.
  - The number of implementation decisions will always outpace the number of checks.
  - The creativity of humans will always outpace the development of checks to limit that creativity.
- There is no substitute for the involvement of experienced people in the compliance-assessment and remediation processes.

Will the dataset created from this CRF fail a validation check? (1)

GENERAL MEDICAL HISTORY

Does the subject have any significant medical history within the past 6 months?

☐ **Yes, list the condition(s) below** ☐ **No**

Body System			**Condition**	**End Date** (mm/dd/yyyy)	**Check if Ongoing**
1. Eyes, Ears, Nose, Throat	☐ Yes	☐ No	_____	___/___/_____	☐
2. Respiratory	☐ Yes	☐ No	_____	___/___/_____	☐
3. Gastrointestinal	☐ Yes	☐ No	_____	___/___/_____	☐
4. Endocrine/Metabolic	☐ Yes	☐ No	_____	___/___/_____	☐

Does the “Body System” represent a “Pre-Specified” term? Do we have an --OCCUR variable at all in the SDTM MH domain? The next slide shows the sponsor’s original MH dataset.

Will this dataset fail a validation check? (2)

DOMAIN	MHSEQ	MHTERM	MHCAT	MHSCAT	MHOCCUR	MHENRTPT	MHENTPT
MH	1	Allergies	GENERAL MEDICAL HISTORY		N	U	2012-11-19
MH	2	Dermatological Disease	GENERAL MEDICAL HISTORY		N	U	2012-11-19
MH	3	Endocrine/Metabolic Disease	GENERAL MEDICAL HISTORY		N	U	2012-11-19
MH	4	Neurological Disease	GENERAL MEDICAL HISTORY		N	U	2012-11-19
MH	5	APPENDECTOMY - 1965	GENERAL MEDICAL HISTORY	SURGERY	Y	U	2012-11-19
MH	6	BACK PAIN	GENERAL MEDICAL HISTORY	MUSCULOSKELETAL DISEASE	Y	ONGOING	2012-11-19
MH	7	BLADDER CA	GENERAL MEDICAL HISTORY	GENITO-URINARY DISEASE	Y	ONGOING	2012-11-19
MH	8	BILATERAL CATERACTS	GENERAL MEDICAL HISTORY	HEENT	Y	ONGOING	2012-11-19
MH	9	CONSTIPATION	GENERAL MEDICAL HISTORY	GASTROINTESTINAL DISEASE	Y	ONGOING	2012-11-19

- If a body system didn't have a history, the body system was mapped to MHTERM and MHOCCUR was set to "N'. Is this according to SDTM?
- Also of note, the relative timing variable MHENRTPT is set to "U" (for "Unknown") for those records where MHOCCUR = "N".
- Of course, the question is, Should these be MH records in the first place?

The Ideal: Validation for Compliance and for Data Integrity Should be Separate Processes

The Reality

- These types of checks are bundled together to help reviewers identify potential problems in the data.
- The submission of SDTM datasets makes it easy to check for data integrity at the same time.

Potential Problems

- Some sponsors don’t understand the difference.
- Some sponsors request that legacy data be “fixed” during conversion in order to avoid known errors.

Agenda

- Introduction
- Data Flow
- Compliance
- Extra Data
- Supplemental Qualifiers and Findings About
- Timing Variables
- Exposure Data
- Trial Design

Theory: The FDA Might Want "Extra" Data in SUPP-- Datasets (1)

The Reality

- There is no limit to the number of non-standard variables that can be submitted.
- If it wasn't true tabulation data, and the FDA wanted it, they would ask for it.

Theory: The FDA Might Want "Extra" Data in SUPP-- Datasets (2)

Potential Problem Areas

- Analysis data
- Additional dates associated with QNAMs that may be a signal for the creation of a new dataset (different "timing" than the parent record)
- Multiple representations of the same data
  - Numeric and character codes
  - Legacy controlled terminology
  - Separate dates and times
  - SAS dates
- Non-unique QNAMs for relating to the same parent record
- Legacy values that were mapped to CDISC CT

“Extra” Data – Submitting data “as collected” as well as mapped to Controlled Terminology (1)

Adverse Event CRF (Partial)

Onset Date (dd-MMM-yyyy): _____ **AESTDTC**

Resolution Date (dd-MMM-yyyy): _____ **AEENDTC**

Outcome:

Resolved	☐	**AEOUT**
Ongoing	☐	**AEENRTPT**
Died	☐	
Change in Grade or Seriousness	☐	

"Extra" Data – Submitting data "as collected" as well as mapped to Controlled Terminology (2)

- AE Dataset – Outcome mapped to CT

STUDYID	DOMAIN	USUBJID	AESEQ	AETERM	AEOUT	AESTDTC	AEENRTPT	AEENTPT
ABC15-0806	AE	ABC15-0806-001	7	EJECTION FRACTION DECREASED	NOT RECOVERED/NOT RESOLVED	2011-03-22	ONGOING	TRIAL EXIT
ABC15-0806	AE	ABC15-0806-001	20	SINUS TACHYCARDIA	NOT RECOVERED/NOT RESOLVED	2011-03-22	ONGOING	TRIAL EXIT
ABC15-0806	AE	ABC15-0806-003	1	ALOPECIA	NOT RECOVERED/NOT RESOLVED	2011-03-22	ONGOING	TRIAL EXIT
ABC15-0806	AE	ABC15-0806-003	9	MACULAR UPPER CHEST RASH	NOT RECOVERED/NOT RESOLVED	2011-03-17	ONGOING	TRIAL EXIT
ABC15-0806	AE	ABC15-0806-005	1	ALOPECIA	NOT RECOVERED/NOT RESOLVED	2011-03-29	ONGOING	TRIAL EXIT
ABC15-0806	AE	ABC15-0806-005	6	FATIGUE	NOT RECOVERED/NOT RESOLVED	2011-06-16	ONGOING	TRIAL EXIT
ABC15-0806	AE	ABC15-0806-005	9	RIGHT CHEST WALL PAIN	NOT RECOVERED/NOT RESOLVED	2011-05-10	ONGOING	TRIAL EXIT
ABC15-0806	AE	ABC15-0806=005	11	SINUS TACHYCARDIA	NOT RECOVERED/NOT RESOLVED	2011-03-08	ONGOING	TRIAL EXIT

Collected value of "Ongoing" mapped to Controlled Terminology in AEOUT as well as to an "end" relative timing variable.

“Extra” Data – Submitting data “as collected” as well as mapped to Controlled Terminology (3)

STUDYID	RDOMAIN	USUBJID	IDVAR	IDVARVAL	QNAM	QLABEL	QVAL	QORIG
ABC15-0806	AE	ABC15-0806-001	AESEQ	7	AEOUT	AE Outcome	Ongoing	CRF
ABC15-0806	AE	ABC15-0806-001	AESEQ	20	AEOUT	AE Outcome	Ongoing	CRF
ABC15-0806	AE	ABC15-0806-003	AESEQ	1	AEOUT	AE Outcome	Ongoing	CRF
ABC15-0806	AE	ABC15-0806-003	AESEQ	9	AEOUT	AE Outcome	Ongoing	CRF
ABC15-0806	AE	ABC15-0806-005	AESEQ	1	AEOUT	AE Outcome	Ongoing	CRF
ABC15-0806	AE	ABC15-0806-005	AESEQ	6	AEOUT	AE Outcome	Ongoing	CRF
ABC15-0806	AE	ABC15-0806-005	AESEQ	9	AEOUT	AE Outcome	Ongoing	CRF
ABC15-0806	AE	ABC15-0806-005	AESEQ	11	AEOUT	AE Outcome	Ongoing	CRF

Collected outcome of “Ongoing” also mapped to SUPPAE with an invalid QNAM (QNAM cannot be the same as a valid SDTM variable name).

Agenda

- Introduction
- Data Flow
- Compliance
- Extra Data
- Supplemental Qualifiers and Findings About
- Timing Variables
- Exposure Data
- Trial Design

In Theory: Supplemental Qualifier Key Principles Would Be Clear

- Non-Standard variables in SUPP-- are no less important than the standard Qualifiers.
- Every record must relate back to at least one valid parent record.
  - If the IDVAR is something other than --SEQ, a single SUPP record may relate to more than one parent record.
- -- QNAMs and QVALs relate only to parent records, but not to each other.
- Data (QVALs) share the timing of the parent record.

Myth: Supplemental Qualifiers are Always the Best Solution for NSVs

Case Study: Physical Exam

Case Study: Physical Exam

Physical Exam CRF

Body System	Description of Abnormality	
Abdomen		
Extremities/Joints		
General Appearance		
Heart		
HEENT		
Lungs		
Lymph Nodes		
Mental Status		
Neurologic		
Reflexes		
Skin		
Disease Relapse Since Last Visit?	□ Yes □ No	

Some sponsors have been tempted to model the last question as a Supplemental Qualifier because it's not a body system..

Case Study: Physical Exam

PE Dataset (Partial)

STUDYID	DOMAIN	USUBJID	PESEQ	PETESTCD	PETEST	PEORRES	PESTRESC
2001-01	PE	2001-01-1008	1	ABDOMEN	Abdomen	NORMAL	NORMAL
2001-01	PE	2001-01-1008	2	EXTRJOIN	Extremities/ Joints	JOINTS SWOLLEN IN FINGERS	JOINTS SWOLLEN IN FINGERS
2001-01	PE	2001-01-1008	3	GENAPP	General Appearance	NORMAL	NORMAL
2001-01	PE	2001-01-1008	4	HEART	Heart	NORMAL	NORMAL
2001-01	PE	2001-01-1008	5	HEENT	HEENT	NORMAL	NORMAL
2001-01	PE	2001-01-1008	6	LUNGS	Lungs	NORMAL	NORMAL
2001-01	PE	2001-01-1008	7	LYMPNODE	Lymph Nodes	NORMAL	NORMAL
2001-01	PE	2001-01-1008	8	MENTSTAT	Mental Status	NORMAL	NORMAL
2001-01	PE	2001-01-1008	9	NEURO	Neurologic	NORMAL	NORMAL
2001-01	PE	2001-01-1008	10	REFLEXES	Reflexes	NORMAL	NORMAL
2001-01	PE	2001-01-1008	11	SKIN	Skin	NORMAL	NORMAL

Case Study: Physical Exam
SUPPPE Representations

Method 1: IDVAR = VISIT

STUDYID	RDOMAIN	USUBJID	IDVAR	IDVARVAL	QNAM	QLABEL	QVAL	QORIG
2001-01	PE	2001-01-1008	VISIT	1	PERELFL	Relapse Since Last Visit	N	CRF

Method 2: IDVAR = PESEQ

STUDYID	RDOMAIN	USUBJID	IDVAR	IDVARVAL	QNAM	QLABEL	QVAL	QORIG
2001-01	PE	2001-01-1008	PESEQ	1	PERELFL	Relapse Since Last Visit	N	CRF
2001-01	PE	2001-01-1008	PESEQ	2	PERELFL	Relapse Since Last Visit	N	CRF
2001-01	PE	2001-01-1008	PESEQ	3	PERELFL	Relapse Since Last Visit	N	CRF
2001-01	PE	2001-01-1008	PESEQ	4	PERELFL	Relapse Since Last Visit	N	CRF
2001-01	PE	2001-01-1008	PESEQ	5	PERELFL	Relapse Since Last Visit	N	CRF
2001-01	PE	2001-01-1008	PESEQ	6	PERELFL	Relapse Since Last Visit	N	CRF
2001-01	PE	2001-01-1008	PESEQ	7	PERELFL	Relapse Since Last Visit	N	CRF
2001-01	PE	2001-01-1008	PESEQ	8	PERELFL	Relapse Since Last Visit	N	CRF
2001-01	PE	2001-01-1008	PESEQ	9	PERELFL	Relapse Since Last Visit	N	CRF
2001-01	PE	2001-01-1008	PESEQ	10	PERELFL	Relapse Since Last Visit	N	CRF
2001-01	PE	2001-01-1008	PESEQ	11	PERELFL	Relapse Since Last Visit	N	CRF

Case Study: Physical Exam

Merged View (Partial) - Both SUPP Representations

STUDYID	DOMAIN	USUBJID	PESEQ	PETESTCD	PETEST	PEORRES	PESTRESC	PERELFL
2001-01	PE	2001-01-1008	1	ABDOMEN	Abdomen	NORMAL	NORMAL	N
2001-01	PE	2001-01-1008	2	EXTRJOIN	Extremities/ Joints	JOINTS SWOLLEN IN FINGERS	JOINTS SWOLLEN IN FINGERS	N
2001-01	PE	2001-01-1008	3	GENAPP	General Appearance	NORMAL	NORMAL	N
2001-01	PE	2001-01-1008	4	HEART	Heart	NORMAL	NORMAL	N
2001-01	PE	2001-01-1008	5	HEENT	HEENT	NORMAL	NORMAL	N
2001-01	PE	2001-01-1008	6	LUNGS	Lungs	NORMAL	NORMAL	N
2001-01	PE	2001-01-1008	7	LYMPNODE	Lymph Nodes	NORMAL	NORMAL	N
2001-01	PE	2001-01-1008	8	MENTSTAT	Mental Status	NORMAL	NORMAL	N
2001-01	PE	2001-01-1008	9	NEURO	Neurologic	NORMAL	NORMAL	N
2001-01	PE	2001-01-1008	10	REFLEXES	Reflexes	NORMAL	NORMAL	N
2001-01	PE	2001-01-1008	11	SKIN	Skin	NORMAL	NORMAL	N

Regardless of which IDVAR is chosen, the merged view is the same. Method 1 just simply collapses 11 SUPPPE records to a single SUPPPE record (limited usefulness).

Case Study 1: Physical Exam
Alternate Representation - No SUPPPE

STUDYID	DOMAIN	USUBJID	PESEQ	PETESTCD	PETEST	PEORRES	PESTRESC	PERELFL
2001-01	PE	2001-01-1008	1	ABDOMEN	Abdomen	NORMAL	NORMAL	N
2001-01	PE	2001-01-1008	2	EXTRJOIN	Extremities/ Joints	JOINTS SWOLLEN IN FINGERS	JOINTS SWOLLEN IN FINGERS	N
2001-01	PE	2001-01-1008	3	GENAPP	General Appearance	NORMAL	NORMAL	N
2001-01	PE	2001-01-1008	4	HEART	Heart	NORMAL	NORMAL	N
2001-01	PE	2001-01-1008	5	HEENT	HEENT	NORMAL	NORMAL	N
2001-01	PE	2001-01-1008	6	LUNGS	Lungs	NORMAL	NORMAL	N
2001-01	PE	2001-01-1008	7	LYMPNODE	Lymph Nodes	NORMAL	NORMAL	N
2001-01	PE	2001-01-1008	8	MENTSTAT	Mental Status	NORMAL	NORMAL	N
2001-01	PE	2001-01-1008	9	NEURO	Neurologic	NORMAL	NORMAL	N
2001-01	PE	2001-01-1008	10	REFLEXES	Reflexes	NORMAL	NORMAL	N
2001-01	PE	2001-01-1008	11	SKIN	Skin	NORMAL	NORMAL	N
2001-01	PE	2001-01-1008	12	PERELFL	Relapse Since Last Visit	N	N	

The creation of a new record, rather than the creation of a Supplemental Qualifier, shows that the Relapse Question is related to the physical exam as a whole, and not to every body system.

Reality: Supplemental Qualifiers are Not Always the Best Solution for NSVs

- One can't always conclude that all legacy that doesn't map into a single SDTM-based domain must be a Supplemental Qualifier.
- In some cases, the mapping of non-standard variables into Supplemental Qualifiers resulted in data with different time frames and/or different data structures being shown in merged views of data.
- One must think of what the resulting merged view will look like in assessing:
  - Which NSVs are submitted as Supplemental Qualifiers
  - The selection of IDVAR

In Theory: Findings About Would Be Used Solely for Findings About an Intervention or Event

Reality

- In many cases it's not.
  - Some sponsors view FA as an alternative to custom domains
    - No need to worry about a domain code
  - In many cases, a Findings domain would suffice
    - The --OBJ variable is a duplicate of the --TESTCD, the domain, or some dummy value.
  - Sometimes SUPP-- datasets are an option.

Supplemental Qualifiers vs. Findings About

Characteristic	SUPPQUAL	Findings About	Findings
CDISC Controlled Terminology	Only for a few QNAM values	None for --TESTCD and --TEST values other than that for tests based upon SDTM Qualifiers (e.g., SEV, DUR, OCCUR)	--TESTCD and --TEST values for modeled domains
Timing	Relies on parent record	--DTC for --TESTCD/--TEST	--DTC for --TESTCD/--TEST
Uniqueness defined by other keys	No	Yes	Yes
Uses and Requires --OBJ	No	Yes	No
Ability to relate multiple qualifiers (e.g., results and units) to each other	No	Yes	Yes
Relates to an Event or Intervention record	Always	Likely	Possible

Supplemental Qualifiers – The Better Option? Or is FA?

This example is shown in the CV Therapeutic Area User Guide (TAUG). The subject was judged to have suffered a stroke which was determined to be an ischemic stroke. Is this better shown as a record in FA, or is it simply a supplemental qualifier? What determines the best approach?

ce.xpt

STUDYID	DOMAIN	USUBJID	CESEQ	CELNKID	CETERM	CESTDTC	CEENDTC
2005-ABC	CE	ABC-0101	1	STRK-01	STROKE	2009-01-20	2009-01-20

face.xpt

STUDYID	DOMAIN	USUBJID	FASEQ	FALNKID	FATESTCD	FAOBJ	FAORRES	FADTC
2005-ABC	FA	ABC-0101	1	STRK-01	STRKTYPE	STROKE	ISCHEMIC STROKE	2009-01-21

OR

suppce.xpt

STUDYID	RDOMAIN	USUBJID	IDVAR	IDVARVAL	QNAM	QVAL	QORIG
2005-ABC	CE	ABC-0101	CESEQ	1	STRKTYPE	ISCHEMIC STROKE	CRF

This example points to the important difference between the --STDTC in the Events domain and the --DTC in the Findings domain.

Agenda

- Introduction
- Data Flow
- Compliance
- Extra Data
- Supplemental Qualifiers and Findings About
- Timing Variables
- Exposure Data
- Trial Design

In Theory: Relative Timing Variables Should Be Easy

In Reality

- They are not.
  - The use of --ENRF and --STRF is limited to only pertaining to the relationship to the Study reference Period.
  - The addition of --STTPT, --STRTPT, --ENTPT, and --ENRTPT overcame the limitations of the first two variables.
    - There is no standard format for the time-point variables
  - Sometimes they are poorly understood, and sponsors want to populate them even if they have a date.

Supplemental Qualifiers is not a place to submit ambiguous or "coded" data or when sponsors aren't sure how to map relative timing variables

STUDYID	RDOMAIN	USUBJID	IDVAR	IDVARVAL	QNAM	QLABEL	QVAL	QORIG
2005-ABC	CM	ABC-0101	CMSEQ	1	CMCONT	Continuing	X	CRF
2005-ABC	CM	ABC-0102	CMSEQ	1	CMCONT	Continuing	X	CRF
2005-ABC	CM	ABC-0102	CMSEQ	2	CMCONT	Continuing	X	CRF
2005-ABC	CM	ABC-0104	CMSEQ	1	CMCONT	Continuing	X	CRF
2005-ABC	CM	ABC-0104	CMSEQ	3	CMCONT	Continuing	X	CRF

This sponsor is attempting to submit the actual act of "checking the box" either in addition to mapping to a relative timing variable and utilizing controlled terminology or even instead of mapping to an appropriate relative timing variable

Theory: Representing Timing of Observations Should be Easy (1)

In Reality

- There are challenges.
  - Findings domains don't use --STDTC.
  - Many diseases of interest capture more than one date:
    - Date of symptom onset
    - Date of diagnosis
    - Findings recorded at the time of diagnosis, but included in the study data
    - Date of last exacerbation

Theory: Representing Timing of Observations Should be Easy (2)

MHSEQ	MHTERM	MHCAT	MHPRESP	MHOCCUR	MHDTC	MHSTDTC	MHENRTPT	MHENTPT
1	ATYPICAL	PARK	Y	N	2006-05-01			
2	BALANCE	PARK	Y	Y	2006-05-01	2006-09-21		
3	BRADYKINESIA	PARK	Y	Y	2006-05-01	2006-09-21		
4	DYSKINESIA	PARK	Y	Y	2006-05-01	2007-02-12		
5	ABNORMAL GAIT	PARK	Y	Y	2006-05-01	2007-02-12		
7	L-DOPA TREATMENT	PARK	Y	Y	2006-05-01	2006-12-15		
8	RIGIDITY	PARK	Y	Y	2006-05-01	2007-08-21		
9	TREMOR	PARK	Y	Y	2006-05-01	2006-04-01		
10	BPH	GEN			2010-09-30		ONGOING	SCREENING
11	DEPRESSION	GEN			2010-09-30		ONGOING	SCREENING
12	GLAUCOMA	GEN			2010-09-30		ONGOING	SCREENING

The MHDTC variable is being used to represent the date of original diagnosis of Parkinsons. A better solution would have been to create a separate MH record to represent the date of original diagnosis of Parkinson's disease, rather than having the date appear on every record as above, while also "hijacking" the meaning of MHDTC.

MHSEQ	MHTERM	MHCAT	MHPRESP	MHOCCUR	MHDTC	MHSTDTC	MHENRTPT	MHENTPT
1	PARKINSON'S DISEASE	PARK	Y	Y	2010-09-30	2006-05-01		

Alternatively, possible use case for MHEVDTYP

- Concept being introduced in soon-to-be-published SDTM v1.5
- MHEVDTYP – Medical History Event Type; Proposed “domain-specific” variable
  - Specifies the aspect of the medical condition as shown in MHTERM by which MHSTDTC is defined
    - Date of Diagnosis
    - Date of Relapse
    - Date of Initial Symptoms
    - Most recent episode
- Used in the Schizophrenia, Hepatitis C, and COPD TAUGs

COPD – Date of Onset of Symptoms, Date of Diagnosis

Row	STUDYID	DOMAIN	USUBJID	MHSEQ	MHTERM	MHEVDTYP	MHCAT
1	ABC-123	MH	101	1	CHRONIC OBSTRUCTIVE PULMONARY DISEASE	SYMPTOMS	COPD HISTORY
2	ABC-123	MH	101	2	CHRONIC OBSTRUCTIVE PULMONARY DISEASE	DIAGNOSIS	COPD HISTORY

Row	MHPRESP	MHOCCUR	MHDTC	MHSTDTC
1 (Con't)	Y	Y	2012-09-28	2010-04-01
2 (Con't)	Y	Y	2012-09-28	2011-10-31

Agenda

- Introduction
- Data Flow
- Compliance
- Extra Data
- Supplemental Qualifiers and Findings About
- Timing Variables
- Exposure Data
- Trial Design

In Theory: The EX Domain Should Provide an Accurate Representation of Exposure to Treatment(s)

Reality

- Despite its importance in understanding a drug's safety, our experience indicates that drug exposure data is one of the most vulnerable to poor and incomplete data collection and representation.
  - Frequently derived from protocol
  - Sometimes derived from drug accountability data
  - Sometimes submitted as blinded data
  - Sometimes doesn't provide entire picture

Reality 1: EX Derived from Protocol

- No data collected
- Assumes all subjects took what was intended
- It's easy, but reliability may be questionable

Reality 2: EX from Drug Accountability Data (1)

- Drug accountability data is generally collected for use in the calculation of compliance.
- Although the SDTMIG Method 4 indicates it can be used as a source for creating an EX dataset, our experience indicates this is generally not a good practice for the accurate representation of exposure.
- Dates dispensed and returned may serve as “anchors” from which to create a “constant dosing interval.”
- Using drug accountability data for dosing amounts on specific dates or dosing frequencies could lead to misleading information.

Reality 2: EX from Drug Accountability Data (2)

- Case Study:
  - 28 tablets dispensed for QD (daily) dosing
  - 14 tablets returned at end of the four-week period.
- The actual exposure could have been one of many scenarios including, but not limited to the following:
  - one tablet QOD (every other day)
  - one tablet QD for the first two weeks
  - one tablet QD for the last two weeks
  - one tablet QD every other week
- Conclusions:
  - The EX dataset is not intended to be another representation of DA data
  - Creating a meaningful Exposure dataset from the information above is impossible

Reality 3: EC from Blinded Data

- CRF data collected is usually blinded data.
- Some sponsors want to submit this blinded data.
  - They think it is useful to reviewers.
  - They have SOPs requiring its submission.
- The SDTMIG now provides for the EC (Exposure as Collected) domain and is intended to represent the “blinded” dose administrations.

STUDYID	DOMAIN	USUBJID	ECSEQ	ECTRT	ECDOSE	ECDOSU	ECSTDTC	ECENDTC
DEF0001	EC	0001-101	1	DRUG B	8	INJECTIONS	2012-01-15	2012-01-23
DEF0001	EC	0001-102	1	DRUG B	8	INJECTIONS	2012-01-17	2012-01-25
DEF0001	EC	0001-103	1	DRUG B	8	INJECTIONS	2012-01-17	2012-01-25

- One of the problems with this representation was that DRUG B was given at three different strengths. EX, by itself, would not show this difference if only number of injections is displayed.
- EX is intended to provide a more granular representation of the treatments within the Elements of an Arm and should “unblind” a reviewer to a subject’s actual dose. The above was not considered to be a valid approach.

Reality 4: EX Partial Data Collection

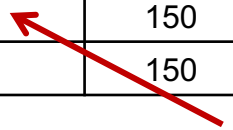
On-Site Dosing

STUDYID	DOMAIN	USUBJID	EXSEQ	EXTRT	EXCAT	EXDOSE	EXDOSFRQ	EXSTDTC	EXENDTC
ABC0001	EX	0001-101	1	DRUG A	AT SITE	150	QD	2012-01-08	
ABC0001	EX	0001-101	2	DRUG A	AT SITE	150	QD	2012-01-15	
ABC0001	EX	0001-101	3	DRUG A	AT SITE	150	QD	2012-01-22	
ABC0001	EX	0001-102	1	DRUG A	AT SITE	150	QD	2012-01-08	
ABC0001	EX	0001-102	2	DRUG A	AT SITE	150	QD	2012-01-15	
ABC0001	EX	0001-102	3	DRUG A	AT SITE	150	QD	2012-01-22	

Above, how many doses of study medication did each subject receive? Do these records reflect a constant dosing interval?

Contrast with the representation below. How many doses did each subject receive?

STUDYID	DOMAIN	USUBJID	EXSEQ	EXTRT	EXCAT	EXDOSE	EXDOSFRQ	EXSTDTC	EXENDTC
ABC0001	EX	0001-101	4	DRUG A		150	QD	2012-01-08	2012-01-22
ABC0001	EX	0001-102	4	DRUG A		150	QD	2012-01-08	2012-01-22



Possibly Use EXCAT = DOSING PERIOD

Agenda

- Introduction
- Data Flow
- Compliance
- Extra Data
- Supplemental Qualifiers and Findings About
- Timing Variables
- Exposure Data
- Trial Design

In Theory: Trial Design Should Be Easy to Create

In Reality

For many studies it's created retrospectively, and this can can be difficult.

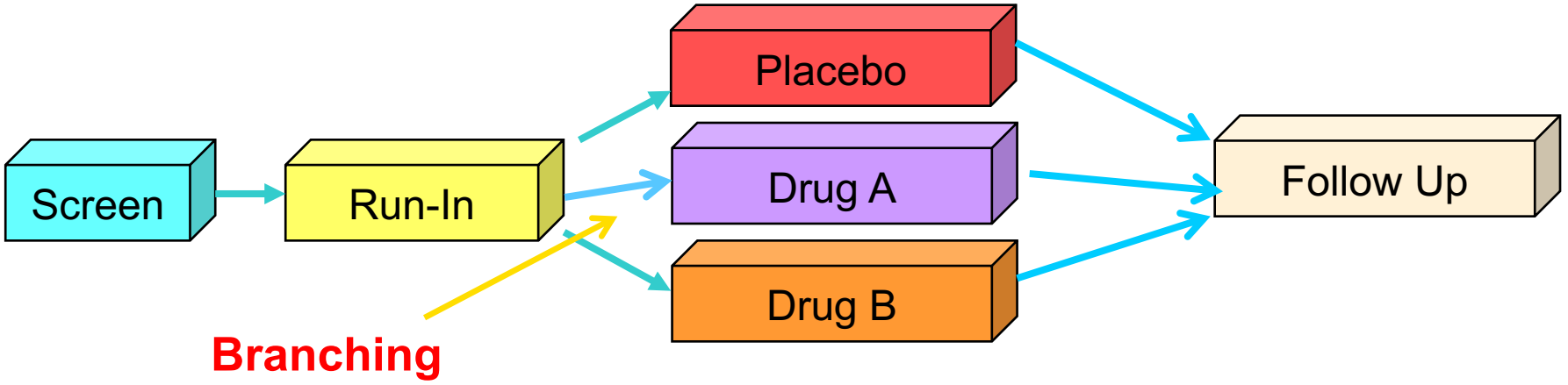
- Finding the necessary information
  - Protocol
  - CRFs
  - Actual Subject Data
- Inconsistency in above information

In Theory: Trial Design Should Be Easy to Understand

In Reality

- Sponsors sometimes over-complicate the design or present only a partial representation of the study design.
  - Sponsors and CROs confuse what is an element within an ARM and what is an Epoch
- Element start rules are often poorly defined
  - Elements may have been defined by events whose dates were not recorded
    - Inability to create proper Element start and end rules.
    - Proper start and end rules are necessary when developing Subject Elements (SE)

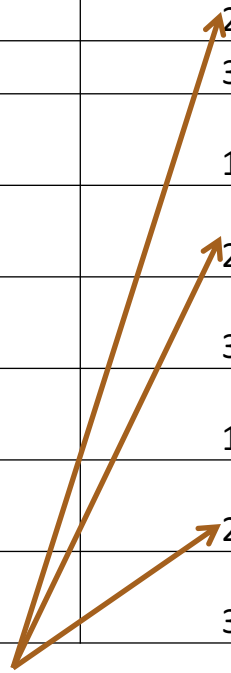
Trial Arms – A Review



- For each ARM, TA contains one record for each occurrence of an element within the ARM
- TABRANCH highlights “decision points” at the end of elements from which subjects “branch” into an element unique to their assigned ARM
- An Epoch is defined in TA as a vertical slice of time that is independent of Arm; identifies a way to tell what is happening across elements while a trial is blinded

Trial Arms – Making an Arm Unique (1)

DOMAIN	ARMCD	ARM	TAETORD	ETCD	ELEMENT	TABRANCH	EPOCH
TA	A	DOCETAXEL	1	SCRN	Screening	Randomized to DOCETAXEL	SCREEN
TA	A	DOCETAXEL	2	**TRT**	Treatment		TREATMENT
TA	A	DOCETAXEL	3	FU	Follow-Up		FOLLOW-UP
TA	B	DRUG A + DOCETAXEL	1	SCRN	Screening	Randomized to Drug A + DOCETAXEL	SCREEN
TA	B	DRUG A + DOCETAXEL	2	**TRT**	Treatment		TREATMENT
TA	B	DRUG A + DOCETAXEL	3	FU	Follow-Up		FOLLOW-UP
TA	C	DRUG B + DOCETAXEL	1	SCRN	Screening	Randomized to Drug B + DOCETAXEL	SCREEN
TA	C	DRUG B + DOCETAXEL	2	**TRT**	Treatment		TREATMENT
TA	C	DRUG B + DOCETAXEL	3	FU	Follow-Up		FOLLOW-UP



Within each Arm, is there an element that makes that Arm unique? Having only a single treatment element doesn't differentiate one Arm from another.

Trial Arms – Making an Arm Unique (2)

DOMAIN	ARMCD	ARM	TAETORD	ETCD	ELEMENT	TABRANCH	EPOCH
TA	A	DOCETAXEL	1	SCRN	Screening	Randomized to DOCETAXEL	SCREEN
TA	A	DOCETAXEL	2	**DOCET**	Docetaxel		TREATMENT
TA	A	DOCETAXEL	3	FU	Follow-Up		FOLLOW-UP
TA	B	DRUG A + DOCETAXEL	1	SCRN	Screening	Randomized to Drug A plus DOCETAXEL	SCREEN
TA	B	DRUG A + DOCETAXEL	2	**DRGADOC**	Drug A plus Docetaxel		TREATMENT
TA	B	DRUG A + DOCETAXEL	3	FU	Follow-Up		FOLLOW-UP
TA	C	DRUG B + DOCETAXEL	1	SCRN	Screening	Randomized to Drug B plus DOCETAXEL	SCREEN
TA	C	DRUG B + DOCETAXEL	2	**DRGBDOC**	Drug B plus Docetaxel		TREATMENT
TA	C	DRUG B + DOCETAXEL	3	FU	Follow-Up		FOLLOW-UP

Within each Arm, now we have a treatment element that makes each Arm unique. Again, having only a single treatment element doesn't differentiate one Arm from another.

Trial Design – Trial Elements (1)

- An element may appear multiple times in Trial Arms (TA) but appears only once in TE
- "Rules" describe how a subject transitions into and out of the element
- There can be no "gaps" in trial elements
- One element always leads right into the next with no gap in between. The start rule of an element defines the end of the previous element
- If trial is blinded, the start rule for a treatment element needs to differentiate one blinded treatment from another
    - ❑ *A subject is always in a trial element throughout their study participation*

Trial Design – Trial Elements (2)

DOMAIN	ETCD	ELEMENT	TESTRL	TEENRL	TEDUR
TE	FU	Follow-Up	End of study	Disease Progression	
TE	SCRN	Screening	Informed consent obtained	Eligibility confirmed	
TE	TRT	Treatment	Eligibility confirmed	End of study or withdrawal	

In the above TE table, there is only a single treatment element; The start rule for a treatment element should point to a record in EX, and not something like “Eligibility Confirmed”. Does this correspond to a data point?

DOMAIN	ETCD	ELEMENT	TESTRL	TEENRL	TEDUR
TE	FU	Follow-Up	Disease progression or unacceptable toxicity	End of Study	
TE	SCRN	Screening	Informed consent obtained	First dose of study medication	
TE	DOCET	Docetaxel	First dose of study medication when subject is randomized to receive Docetaxel	Disease progression or unacceptable toxicity	
TE	DRGADOC	Drug A plus Docetaxel	First dose of study medication when subject is randomized to receive Drug A plus Docetaxel	Disease progression or unacceptable toxicity	
TE	DRGBDOC	Drug B plus Docetaxel	First dose of study medication when subject is randomized to receive Drug B plus Docetaxel	Disease progression or unacceptable toxicity	

The End