SDTM Pilot
A CRO Perspective

Jeff Abolafia, Rho, Inc.
Frank DiIorio, CodeCrafters, Inc.
Laura Brewington, Rho, Inc.
Brooke Millman, Rho, Inc.

Introduction

- Rationale
- Goals of Pilot Study
- Methods
- Demo of 2 Domains
- Trial Design Datasets
- Lessons Learned
- CRO Challenges

Rationale

- Client is considering submitting in CDISC format and asks:
  - Can you do SDTM
  - How much will it cost
  - How long will it take
  - Is it worth the investment during Phase I
- Initial response
  - Of course we can create convert to SDTM
  - We have no idea of the cost
  - We also aren't sure how long it will take
  - So, let's do a pilot on 1 or 2 of your studies

Goals of Pilot Study

- How to convert CRF database to SDTM
- SAS transport or XML files
- Define.pdf or Define.xml
- Develop an overall process
- Estimate timelines
- Estimate resources needed
  - Cost
  - Personnel
- Estimate types of personnel needed
- Estimate reduction in time and resources from study 1 to study 2

Goals of Pilot Study

- What tools need to be developed
- What existing tools can be used or expanded
- What processes can be automated
- What are the challenges
- What are potential roadblocks
- Lessons learned

Methods

- Choose 1 Phase III study and 1 Phase II study and if possible also pick a Phase I study
- Choose studies from different clients
- Choose studies with diverse protocols
- Within study choose at least 1 domain from each observation class
- Choose at least 1 domain where data must be transposed
- Create trial design datasets
- Create a new domain
- Create Suppqual for at least 1 domain

Methods

- Metadata driven system
  - Document the conversion of CRF data to SDTM format
  - Provide specifications to SC programmers
  - Used as a code generator
  - Input for generating define file
  - Facilitate the conversion to SDTM
- Use Hybrid Method
  - Step 1: create CRF variables
  - Step 2: add in derived variables from analysis datasets
  - Step 3: create final SDTM datasets

Background Materials

- Protocol
- SAP (if exists)
- Annotated CRF
- Re-annotated CRF documenting SDTM variables
- Format Library
- SDTM Guide
- SDTM Implementation Guide
- SDTM Metadata stored in a database (we used Microsoft ACCESS)
  - Get from SDTM Implementation Guide
  - Store in machine readable format

SDTM AE METADATA

6.2 EVENTS

6.2.1 Adverse Events — AE

ae.xpt, Adverse Events — Events, Version 3.1.1, August 26, 2005. One record per adverse event per subject, Tabulation

Variable Name	Variable Label	Type	Controlled Terms or Format	Origin	Role	CDISC Notes	Core	References
STUDYID	Study Identifier	Char		CRF	Identifier	Unique identifier for a study within the submission.	Req	SDTM 2.2.4
DOMAIN	Domain Abbreviation	Char	**AE	Derived	Identifier	Two-character abbreviation for the domain most relevant to the observation.	Req	SDTM 2.2.4
USUBJID	Unique Subject Identifier	Char		Sponsor Defined	Identifier	Unique subject identifier within the submission.	Req	SDTM 2.2.4
AESEQ	Sequence Number	Num		CRF or Derived	Identifier	Sequence number given to ensure uniqueness within a dataset for a subject. Can be used to join related records.	Req	SDTM 2.2.4
AEGRPID	Group ID	Char		Sponsor Defined	Identifier	Used to tie together a block of related records in a single domain to support relationships within the domain and between domains.	Perm	SDTM 2.2.4
AEREFID	Reference ID	Char		Sponsor Defined	Identifier	Optional internal or external identifier such as a serial number on an SAE reporting form	Perm	SDTM 2.2.4
AESPID	Sponsor-Defined Identifier	Char		Sponsor Defined	Identifier	Optional Sponsor-defined reference number. Perhaps pre-printed on the CRF as an explicit line identifier or defined in the sponsor's operational database. Example: Line number on a Adverse Events page.	Perm	SDTM 2.2.4
AETERM	Reported Term for the Adverse Event	Char		CRF	Topic	Verbatim name of the event.	Req	SDTMIG 4.1.3.6
AEMODIFY	Modified Reported Term	Char		Sponsor Defined	Synonym Qualifier	If AETERM is modified, then AEMODIFY will contain the modified text.	Perm	SDTMIG 4.1.3.6
AEDECOD	Dictionary-Derived Term	Char	**	Derived	Synonym Qualifier	Dictionary-derived text description of AETERM or AEMODIFY. Equivalent to the Preferred Term (PT in MedDRA). The sponsor should specify the dictionary name and version in the Sponsor Comments column of the Define document.	Req	SDTMIG 4.1.3.5 SDTMIG 4.1.3.6
AECAT	Category for Adverse Event	Char	*	Sponsor Defined	Grouping Qualifier	Used to define a category of related records. Example: BLEEDING, HYPOGLYCEMIA.	Perm	SDTMIG 4.1.2.6

SDTM AE METADATA

Variable Name	Variable Label	Type	Controlled Terms or Format	Origin	Role	CDISC Notes	Core	References
AESCAT	Subcategory for Adverse Event	Char	*	Sponsor Defined	Grouping Qualifier	A further categorization of adverse event. Example: NEUROLOGIC.	Perm	SDTMIG 4.1.2.6
AEOCCUR	Adverse Event Occurrence	Char	**Y, N or Null	CRF or Sponsor Defined	Record Qualifier	Used when the occurrence of specific adverse events is solicited to indicate whether an adverse event occurred or not. Values are null for spontaneously reported events. *Note: Use of this variable is under evaluation by the SDS Team.*	Perm	
AEBODSYS	Body System or Organ Class	Char	**	CRF or Derived	Record Qualifier	Body system or organ class (Primary SOC) that is involved in an event or measurement from the standard hierarchy (e.g., MedDRA).	Exp	SDTMIG 4.1.3.5
AELOC	Location of the Reaction	Char	*	CRF or Derived	Record Qualifier	Describes anatomical location relevant for the event. (e.g., LEFT ARM for skin rash).	Perm	
AESEV	Severity/Intensity	Char	*	CRF	Record Qualifier	The severity or intensity of the event. Examples: MILD, MODERATE, SEVERE.	Perm	
AESER	Serious Event	Char	**Y, N	CRF or Derived	Record Qualifier	Is this is a serious event?	Exp	
AEACN	Action Taken with Study Treatment	Char	*	CRF	Record Qualifier	Describes changes to the study treatment as a result of the event. Examples include ICH E2B values: DRUG WITHDRAWN, DOSE REDUCED,DOSE INCREASED, DOSE NOT CHANGED, UNKNOWN or NOT APPLICABLE	Exp	
AEACNOTH	Other Action Taken	Char		CRF	Record Qualifier	Describes other actions taken as a result of the event. Usually reported as free text. Example: "Treatment unblinded. Primary care physician notified."	Perm	
AEREL	Causality	Char	*	CRF	Record Qualifier	Records the investigator's opinion as to the causality of the event to the treatment. Examples: NOT RELATED,UNLIKELY RELATED, POSSIBLY RELATED, RELATED	Exp	
AERELNST	Relationship to Non-Study Treatment	Char		CRF	Record Qualifier	Records the investigator's opinion as to whether the event may have been to due to a treatment other than study drug. Reported as free text. Example: "More likely related to aspirin use".	Perm	
AEPATT	Pattern of Adverse Event	Char	*	CRF	Record Qualifier	Used to indicate the pattern of the event over time. Examples: INTERMITTENT, CONTINUOUS, SINGLE EVENT.	Perm	

SDTM AE METADATA

Variable Name	Variable Label	Type	Controlled Terms or Format	Origin	Role	CDISC Notes	Core	References
AEOUT	Outcome of Adverse Event	Char	*	CRF	Record Qualifier	Description of the outcome of an event. Examples include ICH E2B values: RECOVERED/ RESOLVED, RECOVERING/ RESOLVING, NOT RECOVERED/ NOT RESOLVED, RECOVERED/ RESOLVED WITH SEQUELAE, FATAL or UNKNOWN.	Perm	
AESCAN	Involves Cancer	Char	**Y, N or Null	CRF	Record Qualifier	Was the event associated with the development of cancer?	Perm	
AESCONG	Congenital Anomaly or Birth Defect	Char	**Y, N or Null	CRF	Record Qualifier	Was the event associated with congenital anomaly or birth defect?	Perm	
AESDISAB	Persist or Signif Disability/Incapacity	Char	**Y, N or Null	CRF	Record Qualifier	Did the event result in persistent or significant disability/incapacity?	Perm	
AESDTH	Results in Death	Char	**Y, N or Null	CRF or Derived	Record Qualifier	Did the event result in death?	Perm	
AESHOSP	Requires or Prolongs Hospitalization	Char	*Y, N or Null	CRF	Record Qualifier	Did the event require or prolong hospitalization?	Perm	
AESLIFE	Is Life Threatening	Char	**Y, N or Null	CRF	Record Qualifier	Was the event life threatening?	Perm	
AESOD	Occurred with Overdose	Char	**Y, N or Null	CRF	Record Qualifier	Did the event occur with an overdose?	Perm	
AESMIE	Other Medically Important Serious Event	Char	**Y, N or Null	CRF	Record Qualifier	Do additional categories for seriousness apply?	Perm	
AECONTRT	Concomitant or Additional Trtmnt Given	Char	**Y, N	CRF	Record Qualifier	Was another treatment given because of the occurrence of the event?	Perm	
AETOXGR	Standard Toxicity Grade	Char	**	CRF or Derived	Record Qualifier	Toxicity grade according to a standard toxicity scale such as Common Terminology Criteria for Adverse Events v3.0 (CTCAE). Sponsor should specify scale and version used in the Comments column of the Define data definition document. Value should contain valid numbers only; datatype will be changed to numeric in future version of SDTM.	Perm	
AESTDTC	Start Date/Time of Adverse Event	Char	ISO 8601	CRF or Derived	Timing		Exp	SDTMIG 4.1.4.1; SDTMIG 4.1.4.2
AEENDTC	End Date/Time of Adverse Event	Char	ISO 8601	CRF or Derived	Timing		Exp	SDTMIG 4.1.4.1; SDTMIG 4.1.4.2

SDTM AE METADATA

Variable Name	Variable Label	Type	Controlled Terms or Format	Origin	Role	CDISC Notes	Core	References
AESTDY	Study Day of Start of Adverse Event	Num		Derived	Timing	Study day of start of adverse event relative to the sponsor-defined RFSTDTC.	Perm	SDTMIG 4.1.4.4
AEENDY	Study Day of End of Adverse Event	Num		Derived	Timing	Study day of end of event relative to the sponsor-defined RFSTDTC.	Perm	SDTMIG 4.1.4.4
AEDUR	Duration of Adverse Event	Char	ISO 8601	CRF	Timing	Collected duration and unit of an adverse event. Used only if collected on the CRF and not derived from start and end date/times. Example: P1DT2H (for 1 day, 2 hours).	Perm	SDTMIG 4.1.4.3
AEENRF	End Relative to Reference Period	Char	** BEFORE, DURING, AFTER, DURING/ AFTER, U	Derived	Timing	Identifies the end of the event as being BEFORE, DURING, DURING/AFTER or AFTER the sponsor-defined reference period. The sponsor-defined reference period is a continuous period of time defined by a discrete starting point and a discrete ending point. Typically, this period is defined by the start (RFSTDTC) and end (RFENDTC) of the trial. Sponsors should define the reference period in the study metadata. Events that are ongoing at the end of the reference period should have a value of AFTER for this variable. If information such as "PRIOR", "ONGOING", or "CONTINUING" was collected, this information should be translated into AEENRF.	Perm	SDTMIG 4.1.4.7

Rho's AE Metadata

CDName	CDLabel	type	CT	Origin	CDISCNotes	Core	Domain	Length	FDADef	ProgDef
STUDYID	Study Identifier	Char		CRF	Unique identifier for a study within the	Req	AE	10	= 'xxx-CL-xx5'	CONSTANT | 'xxx-CLxx5'
DOMAIN	Domain Abbreviation	Char	**AE	Derived	Two-character abbreviation for the domain most relevant	Req	AE	2	= 'AE'	CONSTANT | 'AE'
USUBJID	Unique Subject Identifier	Char		Sponsor Defined	Unique subject identifier within the submission.	Req	AE	6	CRF Page: Adverse Events Log, Pg. 23	= DOMAIN.AEXPCODE.ID
AESEQ	Sequence Number	Num		CRF	Sequence number given to ensure uniqueness	Req	AE	8		SEQ
AEGRPID	Group ID	Char		Sponsor Defined	Used to tie together a block of related records	Perm	AE			NC
AEREFID	Reference ID	Char		Sponsor Defined	Optional internal or external identifier such	Perm	AE			NC
AESPID	Sponsor ID	Char		Sponsor Defined	Optional Sponsor-defined reference	Perm	AE			NC
AETERM	Reported Term for the Adverse Event	Char		CRF	Verbatim name of the event.	Req	AE	140	CRF Page: Adverse Events Log, Pg. 23	= DOMAIN.AEXPCODE.AETEST
AEMODIFY	Modified Reported Term	Char		Sponsor Defined	If AETERM is modified, then AEMODIFY will contain the	Perm	AE			NC
AEDECOD	Dictionary-Derived Term	Char	MedDRA Version 5.1	Derived	Dictionary-derived text description of AETERM or	Req	AE	100	MedDRA Version 5.1	= DOMAIN.AEXPCODE.PREFERREDNAME
AECAT	Category of Adverse Event	Char	*	Sponsor Defined	Used to define a category of related records.	Perm	AE			NC
AESCAT	Subcategory for Adverse Event	Char	*	Sponsor Defined	A further categorization of adverse event.	Perm	AE			NC
AEOCCUR	Adverse Event Occurrence	Char	**Y, N or Null	CRF or Sponsor Defined	Used when the occurrence of	Perm	AE			NC

Rho's AE Metadata

AE : Table

CDName	CDLabel	type	CT	Origin	CDISCNotes	Core	Length	FDADef	ProgDef
AEDECOD	Dictionary-Derived Term	Char	MedDRA Version 5.1	Derived	Dictionary-derived text description of AETERM or	Req	100	MedDRA Version 5.1	= DOMAIN.AEXPCODE.PREFERREDNAME
AECAT	Category of Adverse Event	Char	*	Sponsor Defined	Used to define a category of related records.	Perm			NC
AESCAT	Subcategory for Adverse Event	Char	*	Sponsor Defined	A further categorization of adverse event.	Perm			NC
AEOCCUR	Adverse Event Occurrence	Char	**Y, N or Null	CRF or Sponsor Defined	Used when the occurrence of specific adverse events	Perm			NC
AEBODSYS	Body System or Organ Class	Char	MedDRA Version 5.1	CRF of Derived	Body system or organ class (Primary SOC) that is involved	Exp		MedDRA Version 5.1	DOMAIN.AEXPCODE.SYSTEMORGANCLASSNAME
AELOC	Location of the Reaction	Char	*	CRF of Derived	Describes anatomical location relevant for the	Perm			NC
AESEV	Severity/Intensity	Char	*Mild, Moderate, Severe	CRF	The severity or intensity of the event. Examples:	Perm	8	CRF Page: Adverse Events Log, Pg. 23	RECODE | DOMAIN.AEXPCODE.INTENSE | $INTENSE.
AESER	Serious Event	Char	**Y, N	CRF or Derived	Is this a serious event?	Exp	1	CRF Page: Adverse Events Log, Pg. 23	RECODE | DOMAIN.AEXPCODE.NOTSAE | $AESAE.
AEACN	Action Taken with Study Treatment	Char	*	CRF	Describes changes to the study treatment as a	Exp			DERIVED

Rho's AE Metadata

AE : Table

CDName	CDLabel	type	CT	Origin	CDISCNotes	Core	Length	FDADef	ProgDef
AEACNOTH	Other Action Taken	Char		CRF	Describes other actions taken as a result of the	Perm			DERIVED
AEREL	Causality	Char	*Definite, Probable, Unlikely	CRF	Records the investigator's opinion as to the causality	Exp	8	CRF Page: Adverse Events Log, Pg. 23	RECODE | DOMAIN.AEXPCODE.CAUSE | $CAUSE.
AERELNST	Relationship to Non-Study Treatment	Char		CRF	Records the investigator's opinion as to whether the	Perm			NC
AEPATT	Pattern of Event	Char	*	CRF	Used to indicate the pattern of the event over	Perm			NC
AEOUT	Outcome of Adverse Event	Char	*Recovered/resolved,Recovered	CRF	Description of the outcome of an event. Examples	Perm	32	CRF Page: Adverse Events Log, Pg. 23	RECODE | DOMAIN.AEXPCODE.OUTCOME | $OUTCOME.
AESCAN	Involves Cancer	Char	**Y, N or Null	CRF	Was the event associated with the development of	Perm			NC
AESCONG	Congenital Anomaly or Birth Defect	Char	**Y, N or Null	CRF	Was the event associated with congenital anomaly or	Perm			NC
AESDISAB	Persist or Signif Disability/Incapacity	Char	**Y, N or Null	CRF	Did the event result in persistent or significant	Perm	1	CRF Page: Adverse Events Log, Pg. 23	RECODE | DOMAIN.AEXPCODE.DISAB | $AEDISAB.
AESDTH	Results in Death	Char	**Y, N or Null	CRF or Derived	Did the event result in death?	Perm	1	CRF Page: Adverse Events Log, Pg. 23	RECODE | DOMAIN.AEXPCODE.DEATH | $AESDTH.

Rho's AE Metadata

AE : Table

CDName	CDLabel	type	CT	Origin	CDISCNotes	Core	Length	FDADef	ProgDef
AESHOSP	Requires or Prolongs Hospitalization	Char	**Y, N or Null	CRF	Did the event require or prolong hospitalization?	Perm	1	CRF Page: Adverse Events Log, Pg. 23	RECODE | DOMAIN.AEXPCODE.HOSP | $AESHOSP.
AESLIFE	Is Life Threatening	Char	**Y, N or Null	CRF	Was the event life threatening?	Perm	1	CRF Page: Adverse Events Log, Pg. 23	RECODE | DOMAIN.AEXPCODE.LIFETH | $AESLIFE.
AESOD	Occurred with Overdose	Char	**Y, N or Null	CRF	Did the event occur with an overdose?	Perm			NC
AESMIE	Other Medically Important Serious Event	Char	**Y, N or Null	CRF	Do additional categories for seriousness apply?	Perm	1	CRF Page: Adverse Events Log, Pg. 23	RECODE | DOMAIN.AEXPCODE.IMPMED | $AESMIE.
AECONTRT	Concomitant or Additional Trtmnt Given	Char	**Y, N	CRF	Was another treatment given because of the	Perm	1	CRF Page: Adverse Events Log, Pg. 23	RECODE | DOMAIN.AEXPCODE.TREAT1 | $AECOTRT.
AETOXGR	Standard Toxicity Grade	Char	*	CRF or Derived	Toxicity grade according to a standard toxicity scale	Perm			NC
AESTDTC	Start Date/Time of Adverse Event	Char	ISO8601	CRF of Derived		Exp	19	CRF Page: Adverse Events Log, Pg. 23	DTPART | DOMAIN.AEXPCODE.AESTYR | AESTMO | AESTDA | AESTHR | AESTMI
AEENDTC	End Date/Time of Adverse Event	Char	ISO 8601	CRF of Derived		Exp	19	CRF Page: Adverse Events Log, Pg. 23	DTPART | DOMAIN.AEXPCODE.AESPYR | AESPMO | AESPDA | AESPHR | AESPMI
AESTDY	Study Day of Start of Event	Num		Derived	Study day of start of adverse event relative to the	Perm			DERIVED

Rho's AE Metadata

AE : Table

CDName	CDLabel	type	CT	Origin	CDISCNotes	Core	Length	FDADef	ProgDef
AEENDY	Study Day of End of Event	Num		Derived	Study day of end of event relative to the sponsor-	Perm			DERIVED
AEDUR	Duration of Event	Char	ISO 8601	CRF	Collected duration and unit of an adverse event.	Perm			NC
AEENRF	End Relative to Reference Period	Char	**	Derived	Identifies the end of the event as being BEFORE.	Perm			DERIVED

Rho's VS Metadata

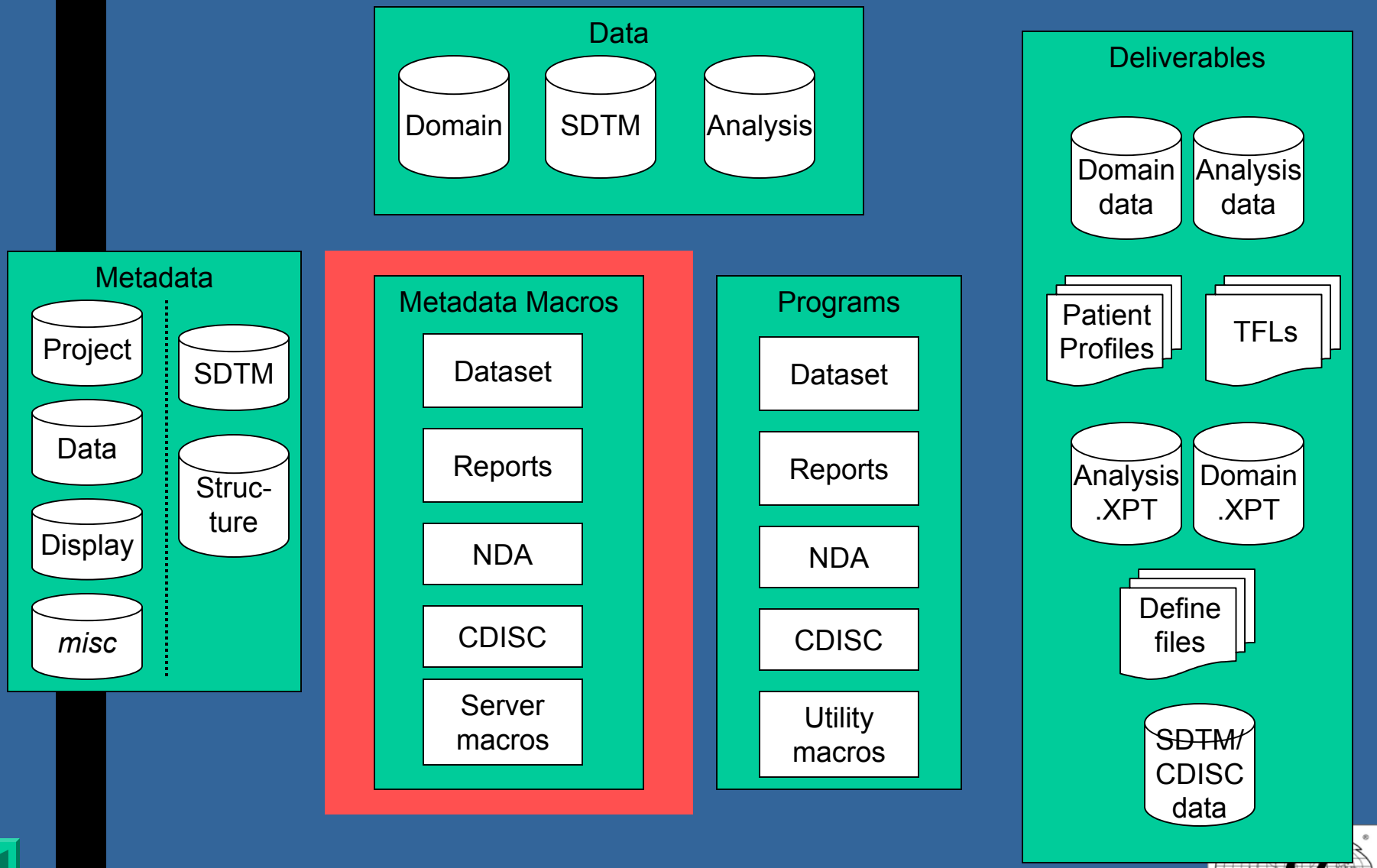
CDLabel	type	CT	Origin	Role	CDISCNotes	Core	Length	FDADef	ProgDef
Study Identifier	Char		CRF	Identifier	Unique identifier for a study within	Req	10	= 'xxx-xxxxx'	CONSTANT | 'xxx-xxxxxx'
Domain Abbreviation	Char	**VS	Derived	Identifier	Two-character abbreviation for	Req	2	= 'VS'	CONSTANT | 'VS'
Unique Subject Identifier	Char		Sponsor Defined	Identifier	Unique subject identifier within	Req	6	CRF Page: Cover Page	= WORK.VS_TEMP.ID~= clinical.aexpcode.id~=work.vs_map.usubjid
Sequence Number	Num		CRF or Derived	Identifier	Sequence number given to	Req	8		DERIVED
Group ID	Char		Sponsor Defined	Identifier	Used to tie together a block	Perm	8		NC
Sponsor ID	Char		Sponsor Defined	Identifier	Optional sponsor-defined reference	Perm			NC
Vital Signs Test Short Name	Char	**	CRF or Derived	Topic	Short name of the	Req	8		VSMAP
Vital Signs Test Name	Char	**	CRF	Synonym Qualifier	Verbatim name of the test or	Req	25		RECODE | WORK.VS_MAP.VSTESTCD | $VSTEST.
Category for Vital Signs	Char	*	Sponsor Defined	Grouping Qualifier	Used to define a category of	Perm			NC
Subcategory for Vital Signs	Char	*	Sponsor Defined	Grouping Qualifier	A further categorization of	Perm	9	FCD test var	assign | clinical.aexpcode | put(aestdt, date9.)
Vital Signs Position of Subject	Char	*	CRF or Derived	Record Qualifier	Position of the subject during a	Exp	8		VSMAP
Result or Finding in Original Units	Char		CRF or Derived	Result Qualifier	Result of the vital signs	Exp	6		VSMAP
Original Units	Char	*	CRF or Derived	Variable Qualifier	Original units in which the data	Exp	10		VSMAP
Character Result/Finding in Std Format	Char		Derived	Result Qualifier	Contains the result value for all	Exp	6		VSMAP
Numeric Result/Finding in Standard Units	Num		Derived	Result Qualifier	Used for continuous or	Exp	8		VSMAP
Standard Units	Char	*	CRF or Derived	Variable Qualifier	Standardized unit used for	Exp	10		VSMAP
Vitals Status	Char	**NOT DONE	CRF or Derived	Result Qualifier	Used to indicate that a vital sign	Perm	8		RECODE | WORK.VS_MAP.VSORRES | $VSSTAT.
Reason Not Performed	Char		CRF or Derived	Record Qualifier	Describes why a measurement or	Perm			NC
LOINC Code	Char	**	Derived	Synonym Qualifier	1. LOINC Code for	Perm			NC
Location of Vital Signs Measurement	Char	*	CRF or Derived	Record Qualifier	Location relevant to the collection	Perm	10		VSMAP
Baseline Flag	Char	**Y or Null	CRF or Derived	Record Qualifier	Indicator used to identify a	ExpS	1		DERIVED
Derived Flag	Char	**Y or Null	Derived	Record Qualifier	Used to indicate a derived record.	Perm	1		CONSTANT | ' '
Visit Name	Char		CRF or Derived	Timing	1. Protocol-defined	Perm	10		RECODE | WORK.VS_TEMP.PHASE | $VISIT.
Visit Number	Num		CRF or Derived	Timing	1. Clinical	Req	8		ASSIGN | WORK.VS_TEMP | INPUT(PHASE,4.)~assign |

cord: 1 of 31

Rho's VS Metadata

VSTESTCD	VSPOS	VSORRESU	VSLOC	VSSTRESU	VSDTC	VSORRES	VSSTRESN
ARTPRES	SITTING		constant | ' '		dtpart | work.VS_TEMP.infuyr | infumo | infuda | vitlhr | vitlmi	ASSIGN | WORK.VS_TEMP | put(sysbp, 3.0)	constant | .
DIAST	SITTING	mm Hg	constant | ' '	mm Hg	dtpart | work.VS_TEMP.infuyr | infumo | infuda | vitlhr | vitlmi	ASSIGN | WORK.VS_TEMP | put(DIABP, 3.0)	constant | .
HR	SITTING	beats/min	constant | ' '	beats/min	dtpart | work.VS_TEMP.infuyr | infumo | infuda | vitlhr | vitlmi	ASSIGN | WORK.VS_TEMP | put(HEARTRT, 3.0)	constant | .
OSAT	SITTING	%	constant | ' '	%	dtpart | work.VS_TEMP.infuyr | infumo | infuda | vitlhr | vitlmi	ASSIGN | WORK.VS_TEMP | put(OSAT, 3.0)	constant | .
RESP	SITTING	resp/min	constant | ' '	resp/min	dtpart | work.VS_TEMP.infuyr | infumo | infuda | vitlhr | vitlmi	ASSIGN | WORK.VS_TEMP | put(RESP, 3.0)	constant | .
SYST	SITTING	mm Hg	constant | ' '	mm Hg	dtpart | work.VS_TEMP.infuyr | infumo | infuda | vitlhr | vitlmi	ASSIGN | WORK.VS_TEMP | put(SYSBP, 3.0)	constant | .
TEMP	SITTING	Celcius	constant | ' '	Celcius	dtpart | work.VS_TEMP.infuyr | infumo | infuda | vitlhr | vitlmi	ASSIGN | WORK.VS_TEMP | put(TEMP,4.1)	constant | .
WEIGHT	STANDING	gms	constant | ' '	gms	dtpart | work.VS_TEMP.infuyr | infumo | infuda	ASSIGN | WORK.VS_TEMP | put(WEIGHT, 4.0)	constant | .

Organization: Using Metadata



Tools Developed

- Metadata System
  - Document conversion process
    - Input format differs by sponsor/project
    - Creating SDTM datasets directly from DMS is problematic
    - Metadata describes how to convert CRF data to SDTM format
  - Provide Specifications to the programmer
  - Input for documentation to the FDA
  - Enhance ability to use automated tools
  - Manage and organize entire project
  - Facilitates uniformity across projects

Tools Developed

- Identify highly repetitive processes
- Macros
  - Create date/times in ISO format
  - Create times in ISO format
  - Transpose data
  - Create define files
  - Create transport or XML files
- Code generator for CRF variables
  - AE example
  - VS example
  - VSMAP example

Processes that can be Automated

- CRF variables
  - Direct copy
  - Rename
  - Recode / reformat
  - Data type conversion (num → Char)
- Datetime / Date / Time variables created from CRF variables
- Code for about 75% of variables can be generated from the metadata
- Most other variables will be created from analysis data sets

The Trial Design Domain (TD)

- A Brand New Tool – Describes the design of a clinical trial in a standardized way
- Datasets created based on study documents
  - Protocol – identify study criteria, timepoints
  - Statistical Analysis Plan – define analysis objectives
- Purpose
  - Quickly grasp the trial design
  - Compare scheduled vs. actual visits, assessments and procedures

Trial Design Datasets

- There are 6 Trial Design Datasets in the TD Domain – As of Version 3.1 of the SDTM Implementation Guide
- 4 are Trial-level, 2 are Subject-level

Dataset	Level	Components	Structure
TE	Trial	Planned Trial Elements	1 record per Element
TA	Trial	Planned Trial Arms	1 record per Arm
TV	Trial	Planned Trial Visits	1 record per Visit
TI	Trial	Inclusion/Exclusion Criteria	1 record per Inclusion/Exclusion Criteria
SE	Subject	Actual Elements, per subject	1 record per Element, per subject
SV	Subject	Actual Visits, per subject	1 record per Visit, per subject

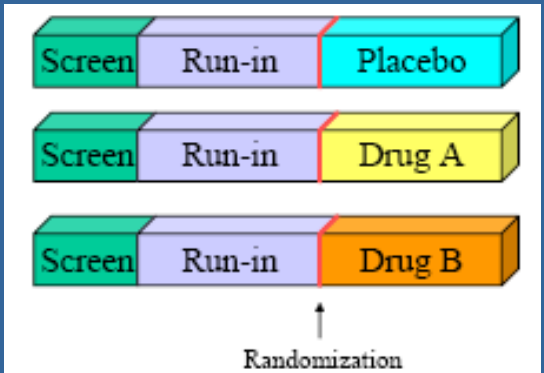
Trial Design Dataset Components

- 4 Components make up the Trial Design model
  - Elements
  - Arms
  - Visits
  - I/E Criteria

Components	Description	Examples
Elements	Consecutive intervals of time within a trial	Screening, TrtA, Washout, Follow-up
Arms	Assigned treatment arm	Placebo, TrtA, TrtB
Visits	Clinical encounters, including assessments and procedures	Baseline, Dosing, Follow-up
I/E Criteria	Protocol-defined Inclusion/Exclusion Criteria	"Subject must be at least 18 years of age."

Trial Design Dataset Components, cont.

- The SDTM guide provides examples of different study structures to help the Statistician choose trial Elements, Arms and Visits
  - A 3-Arm parallel trial, with 5 Elements



Arms

Elements

Trial Design Dataset Components, cont.

- Sample SDTM Trial Arms (TA) table
- Easy to see 3 Elements per Arm, with a rule for moving from one Element to the next

Arms Elements Rules

ARM	TAETORD	ELEMENT	TABRANCH
Placebo	1	Screen	
Placebo	2	Run-In	Randomized to Placebo
Placebo	3	Placebo	
A	1	Screen	
A	2	Run-In	Randomized to Drug A
A	3	Drug A	
B	1	Screen	
B	2	Run-In	Randomized to Drug B
B	3	Drug B	

Trial Design Example - Elements

- Focus on Elements – *what are they*?
  - The basic building blocks of a clinical trial
  - Both Consecutive intervals of time *and* activities within a trial
  - Can be defined multiple ways based on design/analysis objective
    - e.g. In a dose-titration study, the goal may be to analyze effect at each specific dose (=multiple Elements, 1 per dose) or after a certain level has been achieved (=one variable-length Element)

Trial Design Example - Elements

- Process for creating Trial Elements (TE) and Subject Elements (SE) tables
  - Read protocol and analysis plan
  - Define back-to-back Elements
  - Create metadata
  - Program datasets manually

Rho's TE Metadata

Variable Name	Variable Label	Type	Controlled Terms or Format	Origin	Role	CDISC Notes	ProgDefinition
STUDYID	Study Identifier	Char		CRF	Identifier	Unique identifier for a study within the submission.	= 'STUDY01'
DOMAIN	Domain Abbreviation	Char	**TE	Derived	Identifier	Two-character abbreviation for the domain which must be TE.	='TE'
ETCD	Element Code	Char	*	Sponsor Defined	Record Qualifier	Short 8-character version of ELEMENT, used for programming.	= 1 record for each of the following eleme 'SCREEN', 'PLACEBO', 'TRT1', 'TRT2', 'TR 'FOLLOWUP'
ELEMENT	Description of Element	Char	*	Sponsor Defined / Protocol	Topic	The name of the Element.	='Screening' if ETCD='SCREEN'; ~ = 'Place ETCD='PLACEBO'; ~ = 'Treat 1' if ETCD=' 'Treat 2' if ETCD='TRT2'; ~ = 'Treat 3' if ETCD='TRT3'; ~ = 'Follow-up' if ETCD='FOLLOWUP'
TESTRL	Rule for Start of Element	Char		Sponsor Defined / Protocol	Rule	Expresses rule for beginning the Element.	='Informed consent' if ETCD='SCREEN'; ~ dose of placebo' if ETCD='PLACEBO'; ~ = dose of Treat 1' if ETCD='TRT1'; ~ = 'Firs Treat 2' if ETCD='TRT2'; ~ = 'First dose of ETCD='TRT3'; ~ = 'Patch removal' if ETCD='FOLLOWUP
TEENRL	Rule for End of Element	Char		Sponsor Defined / Protocol	Rule	Expresses rule for ending the Element.	='24 hours after start' if ETCD='SCREEN'; hours after start' if ETCD in('PLACEBO' 'T 'TRT2' 'TRT3') ~ = '6 hours after start' if ETCD='FOLLOWUP'
TEDUR	Planned Duration of Element	Char		Sponsor Defined / Protocol	Timing	Planned Duration of Element in ISO 8601 format. For use when the rule for ending the Element is to end after a fixed duration.	='1D' if ETCD='SCREEN'; ~ = '3D' if ETCD in('PLACEBO' 'TRT1' 'TRT2' 'TRT3') ~ = '6 ETCD='FOLLOWUP'

Sample TE Table

etcd	studyid	domain	element	testrl	teenrl	tedur
SCREEN	STUDY01	TE	Screening	Informed Consent	24 hours after start	1D
PLACEBO	STUDY01	TE	Placebo	Application of placebo patch	72 hours after start	3D
TRT1	STUDY01	TE	Treat 1	Application of Treatment 1	72 hours after start	3D
TRT2	STUDY01	TE	Treat 2	Application of Treatment 2	72 hours after start	3D
TRT3	STUDY01	TE	Treat 3	Application of Treatment 3	72 hours after start	3D
FOLLOWUP	STUDY01	TE	Follow-up	Treatment removal	6 hours after start	6H

- 1 record per planned Element
- Start and End rules, plus planned Duration

Rho's SE Metadata

CDName	CDLabel	type	CT	Origin	CDISCNotes	FDADef
STUDYID	Study Identifier	Char		CRF	Unique identifier for a study within the submission.	='STUDY01'
DOMAIN	Domain Abbreviation	Char	**SE	Derived	Two-character abbreviation for the domain which must be SE.	='SE'
USUBJID	Unique Subject Identifier	Char		Sponsor Defined	Unique identifier within the submission.	Concatenate DM.STUDYID, DM.INVID,DM.SUBJID with a dash in between (ex. XYZ100-0100-1234)
ETCD	Subject Element Code	Char	*	Sponsor Defined	Short 8-character version of ELEMENT, used for programming. If an encountered Element differs from the planned ELEMENT to the point that it is considered a new ELEMENT, then use UNPLAN as the value for ETCD to represent this Element.	=TE.ETCD unless the element refers t treatment sequence; then match the s treatment (compress(A_DISPOS.TRT to the correct value of ETCD and delet other treatment elements.
ELEMENT	Description of Subject Element	Char	*	Sponsor Defined/ Protocol	The name of the Element. If an encountered Element differs from the planned ELEMENT to the point that it is considered a new ELEMENT, then ELEMENT should be null.	=TE.ELEMENT, corresponding to ETC
SESTDTC	Start Date/Time of Element	Char	ISO 8601	CRF or Derived	Start date/time for an Element for each subject, represented in ISO 8601 character format.	For start of screening, derive date/time DOM.DEMOG.INTERVDT and format according to ISO standards.
SEENDTC	End Date/Time of Element	Char	ISO 8601	CRF or Derived	End date/time of an Element for each subject, represented in ISO 8601 character format.	For end of screening, derive date/time DOM.DEMOG.INTERVDT and format according to ISO standards.

Sample SE Table

USUBJID	etcd	studyid	element	domain	sestdtc	seendtc
STUDY01-1215-00001	SCREEN	STUDY01	Screening	SE	11/4/1996	11/4/1996
STUDY01-1215-00001	TRT3	STUDY01	Treat 3	SE	1996-11-06T08:13:00	1996-11-09T08:10:00
STUDY01-1215-00001	FOLLOWUP	STUDY01	Follow-up	SE	1996-11-09T08:10:00	1996-11-09T13:50:00
STUDY01-1215-00002	SCREEN	STUDY01	Screening	SE	11/15/1996	11/15/1996
STUDY01-1215-00002	TRT3	STUDY01	Treat 3	SE	1996-11-20T07:56:00	1996-11-23T08:09:00
STUDY01-1215-00002	FOLLOWUP	STUDY01	Follow-up	SE	1996-11-23T08:09:00	1996-11-23T14:15:00
STUDY01-1215-00003	SCREEN	STUDY01	Screening	SE	11/22/1996	11/22/1996
STUDY01-1215-00003	TRT2	STUDY01	Treat 2	SE	1996-12-16T08:03:00	1996-12-21T08:05:00
STUDY01-1215-00003	FOLLOWUP	STUDY01	Follow-up	SE	1996-12-21T08:05:00	1996-12-21T14:00:00
STUDY01-1215-00004	SCREEN	STUDY01	Screening	SE	12/17/1996	12/17/1996
STUDY01-1215-00004	TRT3	STUDY01	Treat 3	SE	1996-12-21T08:04:00	1996-12-21T20:40:00
STUDY01-1215-00004	FOLLOWUP	STUDY01	Follow-up	SE	1996-12-21T20:40:00	1996-12-22T02:04:00
STUDY01-1215-00005	SCREEN	STUDY01	Screening	SE	1/4/1997	1/4/1997
STUDY01-1215-00005	TRT2	STUDY01	Treat 2	SE	1997-01-06T08:23:00	1997-01-09T08:16:00
STUDY01-1215-00005	FOLLOWUP	STUDY01	Follow-up	SE	1997-01-09T08:16:00	1997-01-09T14:20:00
STUDY01-1215-00006	SCREEN	STUDY01	Screening	SE	1/5/1997	1/5/1997
STUDY01-1215-00006	TRT1	STUDY01	TRT1	SE	1997-01-06T08:48:00	1997-01-09T08:05:00
STUDY01-1215-00006	FOLLOWUP	STUDY01	Follow-up	SE	1997-01-09T08:05:00	1997-01-09T14:38:00

- 1 record per subject per Element
- Actual Start and End date/times

Finalizing Trial Design Domains

- The Next Steps
  - Define Trial Arms and Visits and build TA, TV and SV
  - Build TI
  - Beware updates to analysis plans!
- The Good News
  - Trial Design datasets can be built early, as soon as study documents are delivered

Trial Design Challenges

- Creation requires several steps
  - Statistician reviews protocol
  - Specs written for SDTM datasets about study-specific design elements
  - Datasets created manually from metadata
- Statistician required
  - Must have intimate familiarity with study documents

Trial Design Challenges, cont.

- No automatic programming
  - Dataset setup is time-consuming and must be specified and programmed by a statistician
  - During the course of the trial, any variable definition changes must be hard-coded
- And many updates to the Trial Design model are pending…

Lessons Learned

- Metadata based system
- Develop tools to use metadata efficiently
- Document, document, document! (m'data, tools, processes, …)
- Expect, and welcome change. New standards, new client requirements, and the like require adaptability.
- Get into the process by identifying a highly-repetitive or labor-intensive task

Lessons Learned

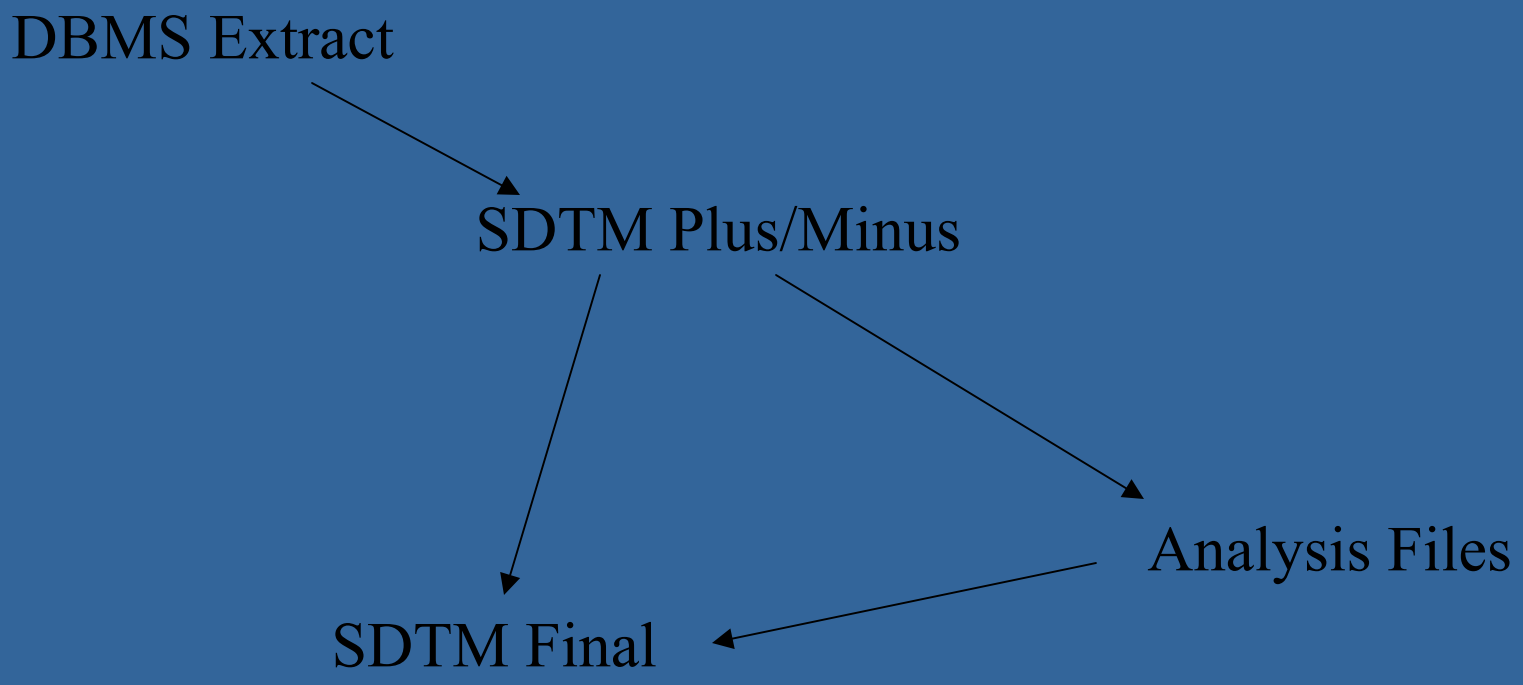
- Just like creating analysis datasets
  - CRF variables copied directly
  - CRF variables recoded
  - Derived Variables
  - Variables from multiple data sources
  - Data often restructured
- Fewer variables than average analysis dataset
- In general, variables less complex than analysis dataset
- Can utilize analysis dataset tools
- More documentation needed for define file

Lessons Learned
When to create SDTM Files

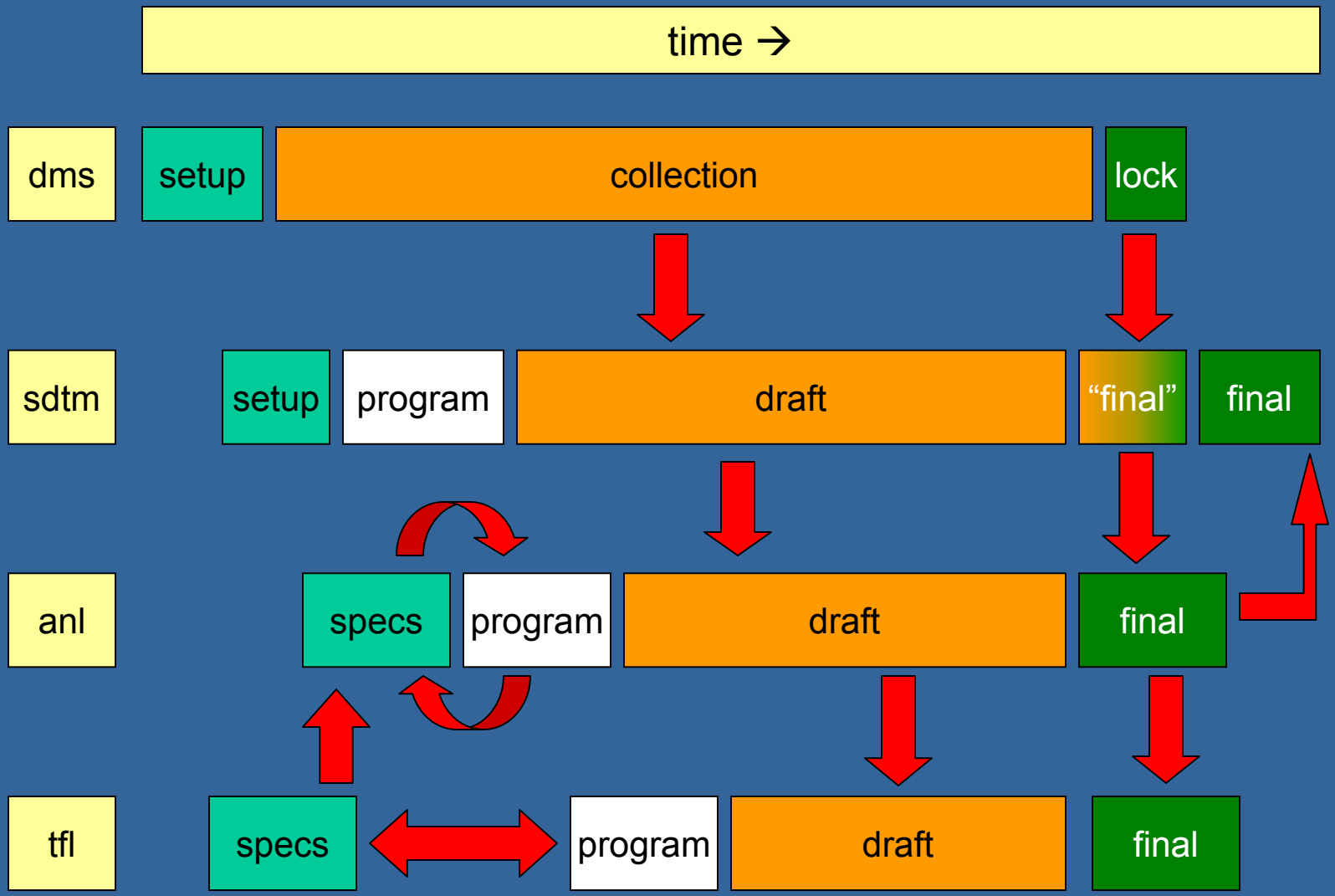
- Alternatives outlined by Susan Kenny (2005) and Jack Shostak (2005)
  - Parallel method
  - Retrospective method
  - Linear method
  - Hybrid Method

Hybrid Method

Source: Susan Kenny (2005)



When to Create SDTM Files



Lessons Learned
When to create SDTM Files

- Hybrid Method Used
- Step 1 – DBMS setup
- SDTM specs started immediately (before patient data arrived)
  - CRFs reannotated
  - Specs written for SDTM datasets variables that do not require statistical input
  - Programs written for SDTM datasets
- Specs/programs/draft datasets developed before 50% of clinical data collected for all domains needed as input for analysis data sets

Lessons Learned
When to create SDTM Files

- Draft SDTM datasets used as input for draft analysis datasets
- After data lock
  - SDTM "plus/minus" datasets finalized
  - Final analysis datasets created
  - SDTM trial design datasets created
  - Final SDTM datasets created
- Additional team dedicated to SDTM is needed
- START SDTM EARLY!!!

Challenges

- CROs must deal with diverse data sources
  - In house DMS
  - CRF data from various clients with various standards and formats
  - Input is not uniform
- Timing
  - SDTM datasets must be developed in parallel with DMS datasets
  - SDTM datasets used as input for analysis datasets
- Resources – more concurrent resources needed due to timing issues

Challenges

- Findings class datasets (VS)
  - Restructured
  - More difficult to automate
  - Need value level metadata
  - Transposed to original structure to create analysis datasets
- How to create define.xml

References

- "Implementation of the CDISC SDTM at the Duke Clinical Research Institute ", J. Shostak . 2005 PharmaSUG Users Group Conference, Phoenix, AZ, May 2005.
- "Strategies for Implementing SDTM and ADaM Standards ", S.J. Kenny . 2005 PharmaSUG Users Group Conference, Phoenix, AZ, May 2005.
- "The Design and Use of Metadata: Part Fine Art, Part Black Art", F. DiIorio and J. Abolafia. Paper to be presented at SUGI 31 available at http://www.codecraftersinc.com/