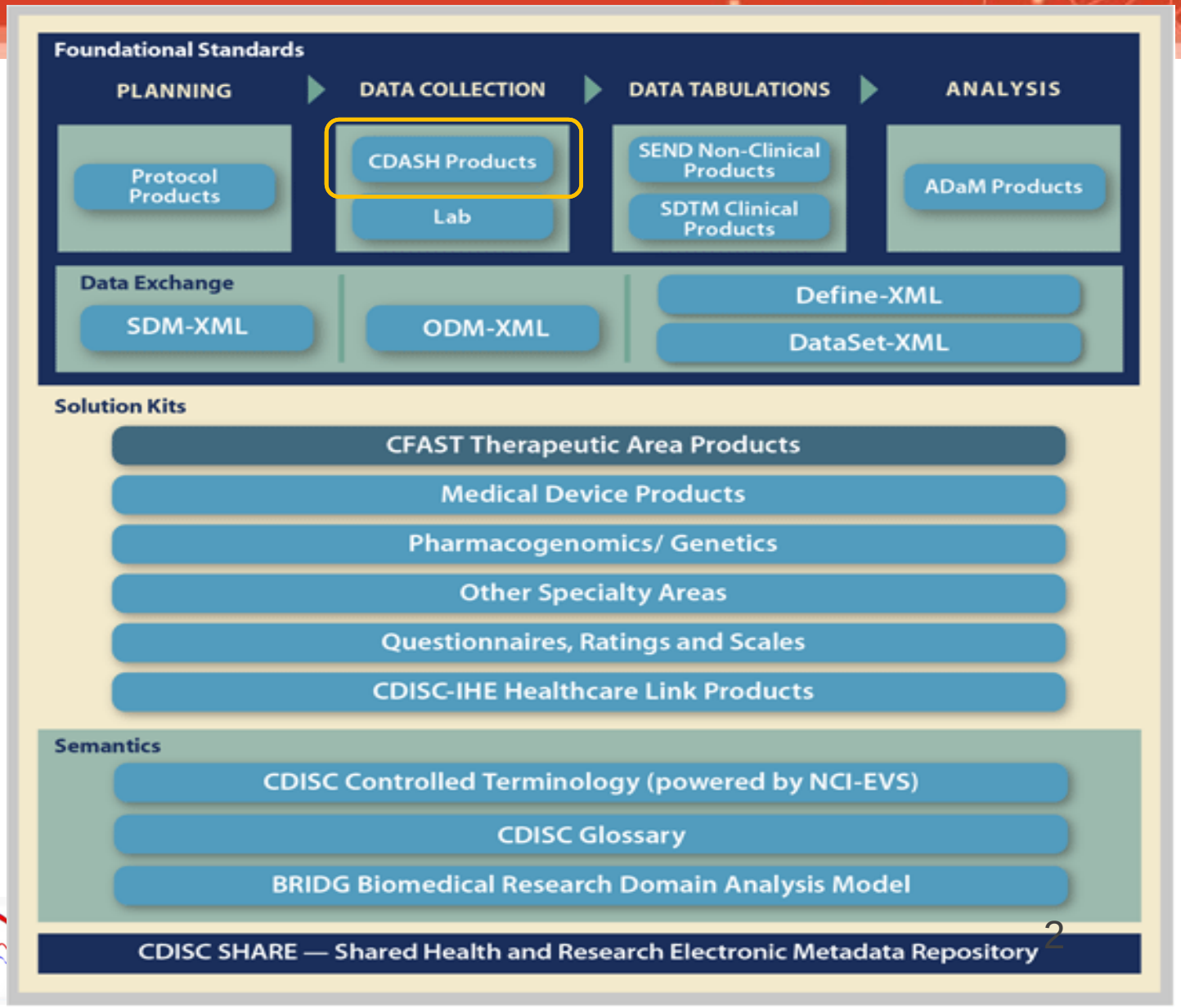
Best Practice in CRF design – CDASH

Victor Wu
CDISC Authorized SDTM Instructor
Vice Chair, China CDISC Coordinating Committee
Chief Consultant
Executive Director, Data Science Department,
Beijing Data Science Express Consulting Co., ltd

吴崇胜
CDISC 认证SDTM讲师
CDISC中国协调委员会副主席
首席咨询师
数据科学中心 总监
北京迪时咨询有限公司

www.datascie.com
victor2050wu@hotmail.com

Relationship between CDASH and other CDISC standards



Topics

- Overview of CDASH
  - Purpose and basic concepts of CDASH
  - Relationship between CDASH and other CDISC standards
  - Collecting data in normalized and de-normalized structure
- CDASH conformant CRF examples
  - Demographics
  - Vital Signs
- CDASH Best Practice recommendations for data collection

What is CDASH

- CDASH - **C**linical **D**ata **A**cquisition **S**tandards **H**armonization

- Describe CDISC-recommended basic standards for the collection of clinical trial data

Purpose of CDASH

- Purpose
  - To describe recommended basic standards for the collection of clinical trial data
  - and be used by those functions involved in the *planning*, *collection*, *management* and *analysis* of clinical trials and clinical data

Basic Concepts of CDASH

- Develop CRF content standards for a basic set of global industry-wide CRF fields to support clinical research
  - ✓ Initial scope limited to most commonly collected data
  - ✓ These CRF standards apply across most therapeutic areas and phases of clinical development (I-IV)
- Maximize re-use of data, CRFs, programming, etc.
- Increase transparency and traceability in the data
- Support data repository, and data sharing
- Support integrating research into clinical care workflow
  - ✓ CDASH is used as a content standard to harvest data from electronic health records

Basic Concepts of CDASH – CDASH Traceability

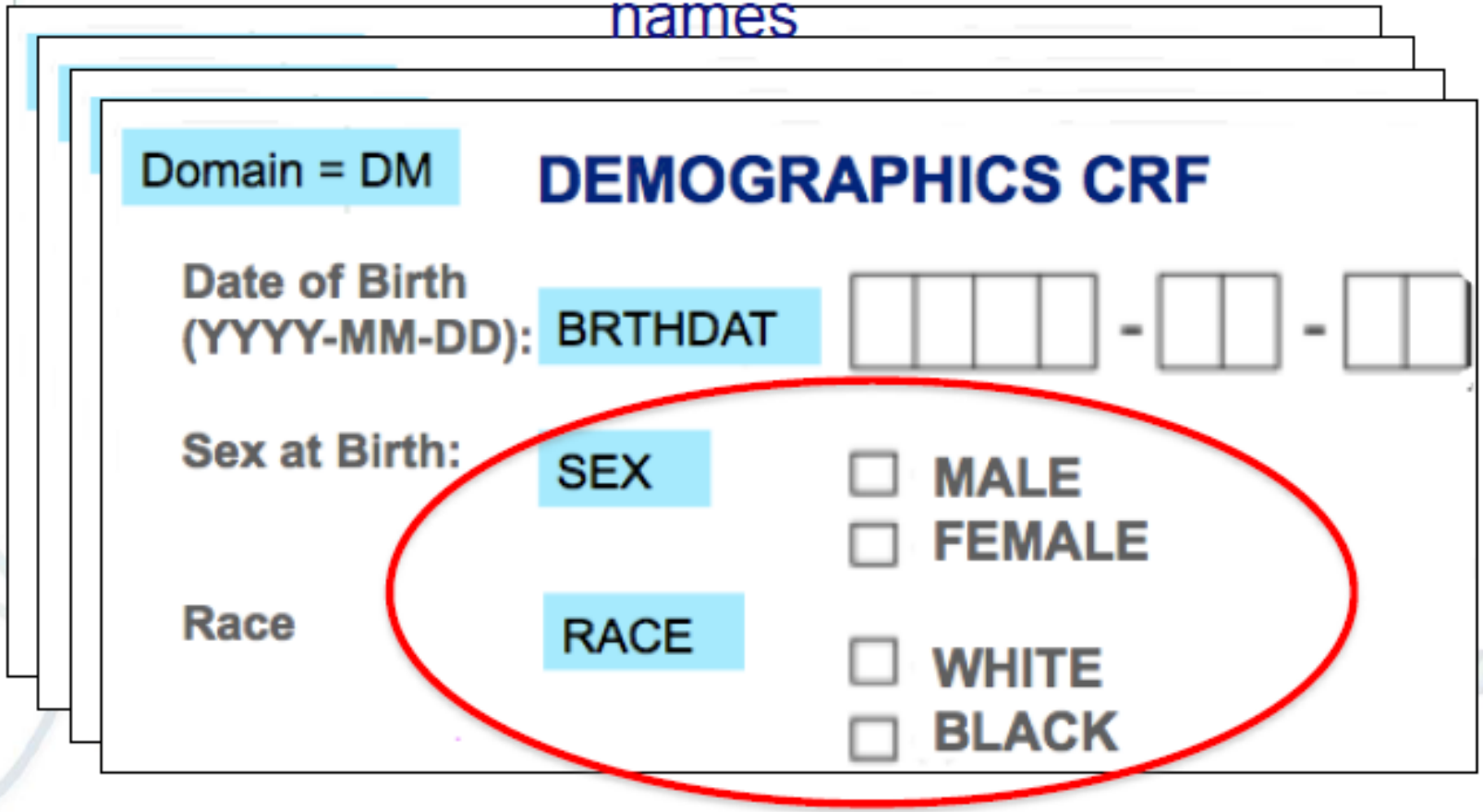
- CDASH uses the same data groupings (domains) as SDTM

- CDASH uses SDTM variables when the data collected are exactly the same as the data submitted
- CDASH uses a standard CDASH variable when the data collected have to be transformed prior to putting the data into an SDTM variable
- CDASH uses a standard CDASH variable when the data collected are not included in the submission

- CDASH uses SDTM controlled terminology so that what is collected in the CRF is the same as the value or variable in SDTM

Basic Concepts of CDASH – Example

CDASH standardizes using the same domain name and variable names



Name for RACE variable is consistent between CDASH and SDTM

Study #1 – DM domain

SEX	RACE
M	WHITE
F	BLACK
F	WHITE
M	ASIAN
F	BLACK

Study #2 – DM domain

SEX	RACE
M	ASIAN
M	WHITE
F	WHITE
F	BLACK
M	WHITE

Study #3 – DM domain

SEX	RACE
M	ASIAN
F	WHITE
F	WHITE
M	BLACK
F	WHITE

Study #4 – DM domain

SEX	RACE
M	BLACK
M	ASIAN
F	BLAC
F	W
M	W

Name for SEX variable is consistent between CDASH and SDTM

Basic Concepts of CDASH – CDASH Domain

General Observation Class

- Events
  - **Adverse Events (**AE)
  - **Medical History (**MH)
  - **Disposition (**DS)
  - **Protocol Deviations (**DV)
- Findings
  - **Laboratory Test Results (**LB)
  - **Drug Accountability (**DA)
  - **ECG Test Results (**EG)
  - **Inclusion and Exclusion Criteria (**IE)
  - **Physical Examination (**PE)
  - **Subject Characteristics (**SC)
  - **Vital Signs (**VS)
- Intervention
  - **Exposure (**EX)
  - ***Prior and* Concomitant Medications (**CM)
  - **Substance Use (**SU)

Special purpose

- Comments (CO)
- Demographic (DM)

Others

- Common identifier
- Common Timing variables

Basic Concepts of CDASH – CDASH Domain Tables

- Explanation of Table Headers
  - Question Text
  - Prompt
  - SDTM or CDASH Variable Name (CDASH variable name shaded)
  - BRIDG
  - Definition: *CRF text examples*; {*code list name*}
  - CRF Completion Instructions
  - Information for Sponsors
  - Core

1	2	3	4	5	6	7	8
Question Text	Prompt	SDTM or CDASH Variable Name	BRIDG	Definition	CRF Completion Instructions	Information for Sponsors	Core

CDASH Domain Tables – Core

- Core Designations for Basic Data Collection Fields
  - ✓ Highly Recommended
  - ✓ Recommended/Conditional
  - ✓ Optional

CDASH Alignment with SDTM – Purpose

- SDTM/SDTMIG - Data submission
- CDASH - Data collection
- The CDASH data collection fields (or variables) can be mapped to the SDTM structure, but there are instances where the variables do not exactly match due to their different purposes

CDASH Alignment with SDTM - Scope

- Derived Data - The SDTM standard contains some derived data whereas CDASH data collection fields are not derived at the data acquisition stage

- Data Collection Fields not Included in the SDTM - Intended to assist in the cleaning of data and in confirming that no data are missing (e.g., AEYN, CMYN, CMONGO). Variable Name will be shaded to indicate it is not to be submitted in SDTM

CDASH Alignment with SDTM – Example

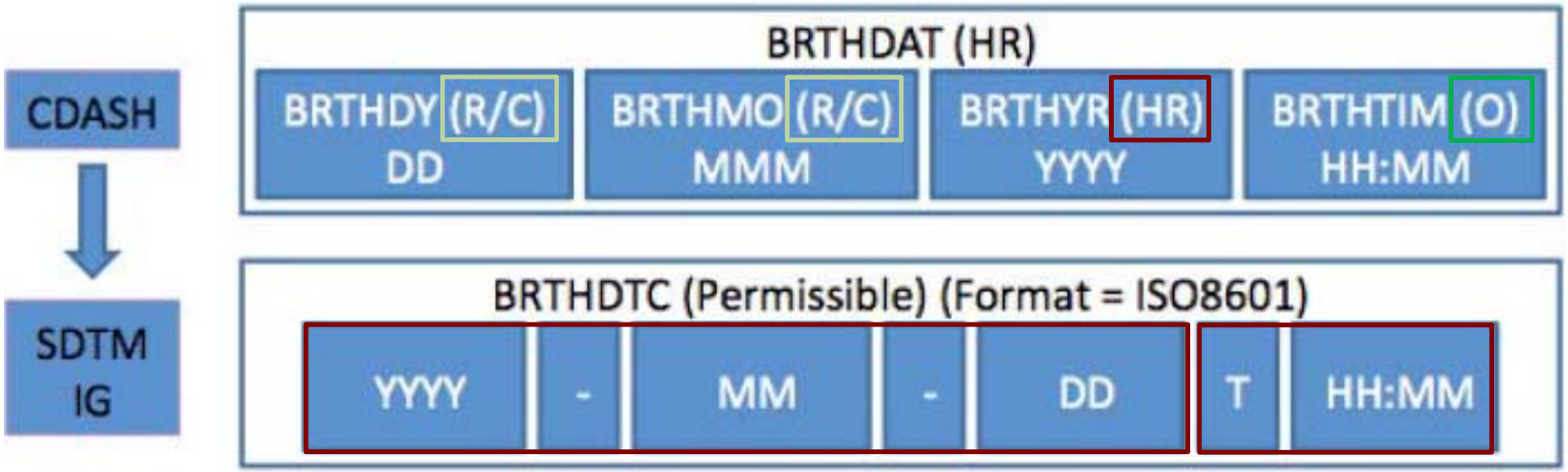
5.3 Adverse Event – AE (Events)

These recommendations are for *non-solicited* or *pre-specified* adverse events. As with all the data collection variables recommended in CDASH Standard Version 1.0, it is assumed that sponsors will add other data variables as needed to meet protocol-specific and other data collection requirements (e.g., TA-specific data elements and others as required per protocol, business practice and operating procedures). Sponsors should define the appropriate collection period for adverse events.

	Question Text	Prompt	SDTM or CDASH Variable Name	BRIDG	Definition	CRF Completion Instructions	Information for Sponsors	Core
1	Were any adverse events experienced?	Any AEs?	AEYN	PerformedObservation Result.value	General prompt question regarding whether or not any AEs were experienced during the study. This provides verification that all other fields on the CRF were deliberately left blank. {NY} (See Section 2.2.)	Indicate if the subject experienced any adverse events. If yes, include the appropriate details where indicated on the CRF.	The intent/purpose of collecting this field is to help with data cleaning and monitoring. It provides verification that all other fields on the CRF were deliberately left blank. This field does not map directly to an SDTM variable.	O
2	AE Identifier	<line number> or <AE number>	AESPID	Implementation specific	A sponsor-defined identifier that can be used for pre-printed numbers on the CRF.	Example instruction: Record unique identifier for each adverse event for this subject. Number sequence for all following pages should not duplicate existing numbers for the subject.	It can be beneficial to use an identifier in a data query to communicate clearly to the site the specific record in question or to reconcile concomitant medications and/or medical history records with AEs. If CMAENO is used, this is the identifier to which CMAENO refers.	O

CDASH Alignment with SDTM

-- Date format from CDASH to SDTM



CDASH Alignment with SDTM
-- Date format from CDASH to SDTM (example)

Date and Time as Originally Recorded (CDASH format)	Precision	ISO 8601 Date/Time (SDTM format)
15 Dec 2003 13:14	Complete date and time – unknown seconds	2003-12-15T13:14
15 Dec 2003	Complete date – unknown time	2003-12-15
Dec 2003	Unknown day – unknown time	2003-12
2003	Unknown month, day and time	2003

CDASH Alignment with CDISC Controlled Terminology

- Under development by the CDISC Terminology Team
- Published by the National Cancer Institute's Enterprise Vocabulary Services (NCI EVS) and can be accessed via the following link: http://www.cancer.gov/cancertopics/terminologyresources/CDISC

CDASH Alignment with CDISC Controlled Terminology
-- Example

	A	B	C	D	E	F	G	H
1	Code	Codelist Code	Codelist Extensible (Yes/No)	Codelist Name	CDISC Submission Value	CDISC Preferred Term	CDISC Synonym(s)	CDISC Definition
2	C66767		No	Action Taken with Study Treatment	ACN	ACN	Action Taken with Study Treatment	The reason that changes were made to the therapy under study. (NCI)
3	C49503	C66767		Action Taken with Study Treatment	DOSE INCREASED	DOSE INCREASED		An indication that a medication schedule was modified by addition; either by changing the frequency, strength or amount. (NCI)
4	C49504	C66767		Action Taken with Study Treatment	DOSE NOT CHANGED	DOSE NOT CHANGED		An indication that a medication schedule was maintained. (NCI)
5	C49505	C66767		Action Taken with Study Treatment	DOSE REDUCED	DOSE REDUCED		An indication that a medication schedule was modified by subtraction, either by changing the frequency, strength or amount. (NCI)
6	C49501	C66767		Action Taken with Study Treatment	DRUG INTERRUPTED	DRUG INTERRUPTED		An indication that a medication schedule was modified by temporarily terminating a prescribed regimen of medication. (NCI)
7	C49502	C66767		Action Taken with Study Treatment	DRUG WITHDRAWN	DRUG WITHDRAWN		An indication that a medication schedule was modified through termination of a prescribed regimen of medication. (NCI)
8	C48660	C66767		Action Taken with Study Treatment	NOT APPLICABLE	NOT APPLICABLE	NA	Determination of a value is not relevant in the current context. (NCI)
9	C17998	C66767		Action Taken with Study Treatment	UNKNOWN	UNKNOWN	U; Unknown	Not known, not observed, not recorded, or refused. (NCI)
10	C66768		No	Outcome of Event	OUT	OUT	Outcome of Event	A condition or event that is attributed to the adverse event and is the result or conclusion of the adverse event. (NCI)
	C48275	C66768		Outcome of Event	FATAL	FATAL	Grade 5; 5	The termination of life as a result of an adverse event. (NCI)

CDISC Controlled Terminology – Example

	Question Text	Prompt	SDTM or CDASH Variable Name	BRIDG	Definition	CRF Completion Instructions	Information for Sponsors	Core
10a	What was the severity of the adverse event?	Severity	AESEV	AdverseEvent.severity Code*	Description of the severity of the adverse event. {AESEV} (See Section 2.2.)	The reporting physician/healthcare professional will assess the severity of the event using the sponsor-defined categories. This assessment is subjective and the reporting physician/ healthcare professional should use medical judgment to compare the reported Adverse Event to similar type events observed in clinical practice. Severity is not equivalent to seriousness.	Either AESEV or AETOXGR **must** appear on the CRF. Some studies may mandate the collection of both. Refer to ICH E3 guidelines for CSR Section 12.2.4. * See the BRIDG model for complete path.	R/C
10b	What is the toxicity grade of the adverse event?	Toxicity Grade	AETOXGR	AdverseEvent.grade Code*	Description of the toxicity grade of the adverse event. {TOXGR} (See Section 2.2.)	Severity CTCAE Grade The reporting physician/healthcare professional will assess the severity of the adverse event using the toxicity grades.	Either AESEV or AETOXGR **must** appear on the CRF. Some studies may mandate the collection of both. Refer to ICH E3 guidelines for CSR Section 12.2.4. CTCAE grade is commonly used in oncology studies although it can also be used elsewhere. * See the BRIDG model for complete path.	R/C
11	Is the adverse event serious?	Serious	AESER	AdverseEventSeriousness .code	Indicates whether or not the adverse event is determined to be "serious" based on what is defined in the protocol. {NY} (See Section 2.2.)	Assess if an adverse event should be classified as serious based on the "serious" criteria defined in the protocol.	This field is related to the individual serious adverse event type fields (reference 11a to 11f), which may or may not be collected on the CRF. Either AESER or all the AESxxx fields (12a-12f) must be present on the CRF.	R/C

Conformance Rules for CDASH Implementation

- Tier 1 conformance – describe the minimum implementation in CDASH standards V1.1
  - All Highly Recommended and applicable Recommended/Conditional Common Identifying and Timing Variables are present in the CRF or available from the operational database.
  - All code lists displayed in the CRF use or map to current published CDISC Controlled Terminology when it is available. Subsets of published Controlled Terminology can be used.
  - The implementation of the CRF follows the Best Practice recommendations
  - CDASH Question Text or Prompt is used
  - Proper use of Questionnaires

Conformance Rules for CDASH Implementation

- Tier 2 conformance – focus on the operational implementation of CDASH in CDASH UG (User Guide)
  - All Level 1 conformances are met
  - All data collection fields are defined using CDASH naming conventions in the operational database unless an equivalent SDTMIG variable can be used for data collection in a user-friendly manner (e.g., using a recommended input format for data collection)
  - All non-CDASH Variable Names in CRFs follow CDASH recommendations for Creating Fields That Do Not Exist in CDASH (Section 2.4.3).
  - All Best Practice recommendations in Section 3 of CDASH V1.1 are followed.

Conformance Rules for CDASH Implementation
-- Creating Fields That Do Not Exist in CDASH

- Evaluate Variable Category
  1. Data cleaning purpose only, not submitted in SDTM
  2. Can be mapped to SDTM directly
  3. Will be mapped to SDTM via SDTM mapping rule, e.g.: date variables
  4. New variable, to be submitted in SDTM
- Generate Variable Name
  - Create variable following CDASH naming convention as needed, if Category 1
  - Use SDTM variable, if Category 2
  - Create variable following CDASH naming convention while referring to SDTM variable, if Category 3
  - Follow list of reserved domain codes, and variable naming fragment in SDTMIG, if Category 4

Collecting data in normalized and de-normalized structures

- CDASH Findings domain tables (e.g., DA, EG, IE, LB, SU, and VS) are presented in a structure that is similar to the SDTM submission model, which is to list the variable names and some examples of the tests in a normalized structure

- However, some Data Management systems collect the data in a de-normalized structure

Collecting data in normalized and de-normalized structures

De-normalized Structure	Normalized Structure
Horizontal view, short and wide	Vertical view, long and narrow
One record per subject visit	One test per record
Sometimes optimal and simplest for entering data	May complicate entry of data in some cases
Not optimal for data analysis and reporting	Optimal for data mining, adopted by the CDISC SDS team
Additional programming is needed to restructure data into the for data analysis/submission.	No additional mapping required for data analysis or submission
Conformant to CDASH, as long as naming convention rules are followed and controlled terminology are used	Conformant to CDASH, as long as naming convention rules are followed and controlled terminology are used

Collecting data in normalized and de-normalized structures

- De-normalized Structure

Subject	Timepoint	Draw_date	Sodium	Potassium	Chloride	Glucose	Calcium	Urea Nitrogen	Creatinine	Bilirubin	AST	ALT	Protein	Cholesterol	Triglycerides	Platelet Count
1	Screening	7-Aug-03	143	4.8	98	4.8	2.48	5.4	72	7.1	26	18	80	6.36	1.62	228

- Normalized Structure

SUBJID	LBSPID	LBTESTCD	LBTEST	LBCAT	LBORRES
ABC-123-01-01-001	1	ALT	ALANINE AMINOTRANSFERASE	CHEMISTRY	18
ABC-123-01-01-001	2	AST	ASPARTATE AMINOTRANSFERASE	CHEMISTRY	26
ABC-123-01-01-001	3	BILI	BILIRUBIN	CHEMISTRY	7.1
ABC-123-01-01-001	4	BUN	BLOOD UREA NITROGEN	CHEMISTRY	5.4
ABC-123-01-01-001	5	CA	CALCIUM	CHEMISTRY	2.48
ABC-123-01-01-001	6	CL	CHLORIDE	CHEMISTRY	98
ABC-123-01-01-001	7	CHOL	CHOLESTEROL	CHEMISTRY	6.36
ABC-123-01-01-001	8	CREAT	CREATININE	CHEMISTRY	72
ABC-123-01-01-001	9	GLUC	GLUCOSE	CHEMISTRY	4.8
ABC-123-01-01-001	10	PLAT	PLATELET	HEMATOLOGY	228
ABC-123-01-01-001	11	K	POTASSIUM	CHEMISTRY	4.8

CDASH conformant CRF examples

StudyDesign: Demographics (DM_1) [DM_UseCase1]

Demographics [DM_UseCase1]

1.* Birth Date
[Birth Date]

[BRTHDAT]
[BRTHYR] **[BRTHMO]**
Birth Year Birth Month

2.* Sex
[Sex]

3.* Ethnicity
[Ethnicity]

4.* Race
[Race]

BRTHDAT (a logical "container" for the BRTHYR and BRTHMO variables)

Row	SUBJID	BRTHYR	BRTHMO	SEX	ETHNIC	RACE
1	100008	1930	Aug	M	NOT HISPANIC OR LATINO	ASIAN
2	100014	1936	Nov	F	HISPANIC OR LATINO	AMERICAN INDIAN OR ALASKA NATIVE
3	200001	1923	Sep	M	HISPANIC OR LATINO	WHITE
4	200002	1933	Jul	F	NOT HISPANIC OR LATINO	BLACK OR AFRICAN AMERICAN
5	200005	1937	Feb	M	NOT HISPANIC OR LATINO	WHITE

ALASKA NATIVE] Alaska Native
[A:ASIAN] Asian
[A:BLACK OR AFRICAN AMERICAN] Black or African American
[A:NATIVE HAWAIIAN OR OTHER PACIFIC ISLANDER] Native Hawaiian or Other Pacific Islander
[A:WHITE] White

Key: [*] = Item is required

CDASH conformant CRF examples

DOMAIN='VS'

Vital Signs

VSORRES, VSSTRESN

Blood Pressure: SYSBP / DIABP mmHg VSORRESU VSSTRESU

VSTESTCD='SYSBP', 'DIABP'

Systolic Diastolic

Pulse: PULSE Beats/minute

Height HEIGHT cm

Weight WEIGHT . kg

Body Mass Index: BMI . Kg/m^2

VSTESTCD='PULSE','HEIGHT','WEIGHT','BMI'

Best Practice Recommendations

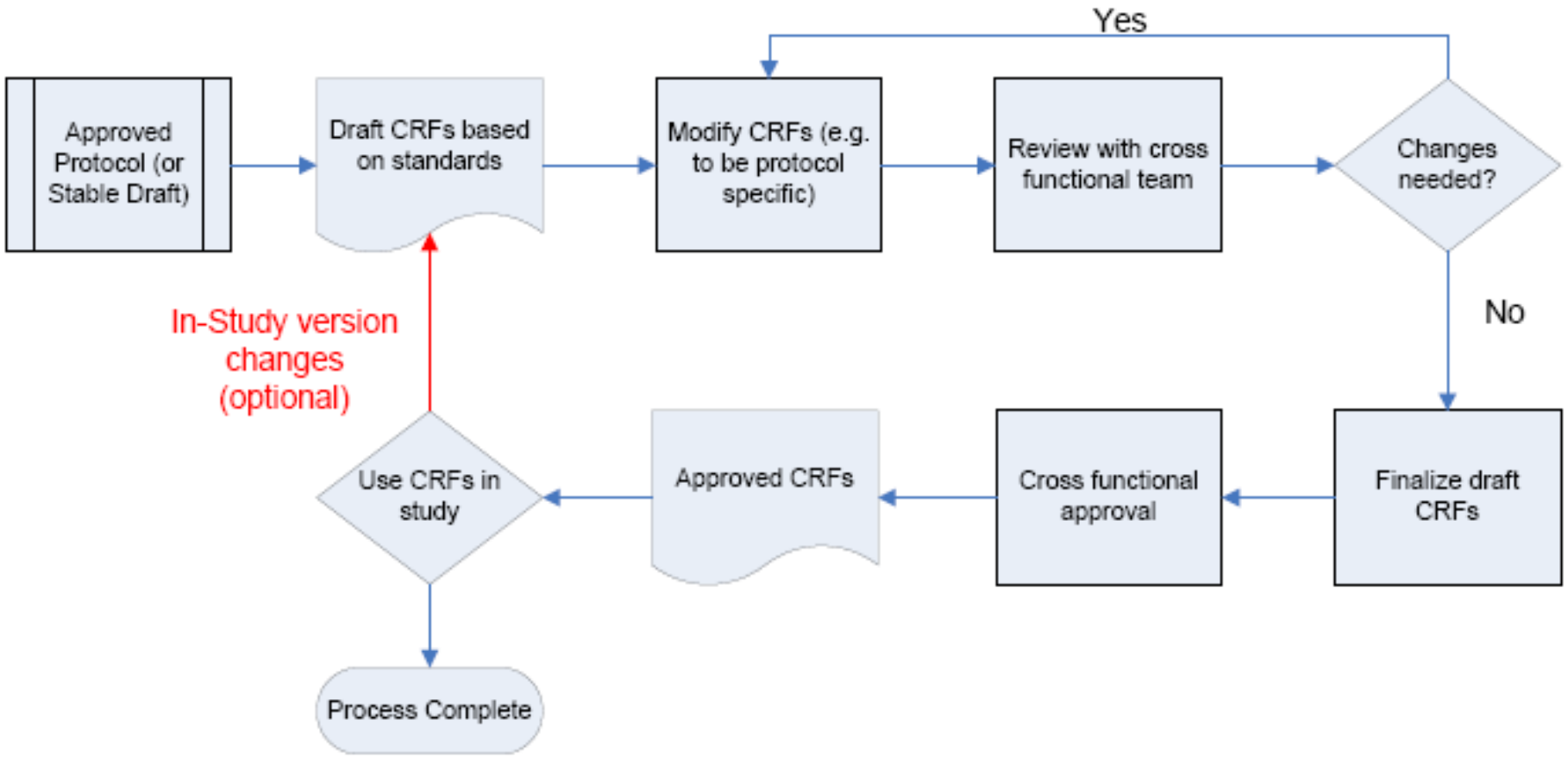
- Necessary Data Only
- Clarity
- CRF Completion Guidelines
- Site Workflow

- Control
- Adequate Review
- Employ Standards
- Translations

- Data Cleaning Prompts
- What to database

Best Practice Recommendations
-- CRF development workflow

3.3 Suggested CRF Development Workflow



Best Practice Recommendations
-- FAQs

- Should ‘Yes/No’ question be preferred over ‘Check all that apply’ questions?
- Should there be a standard order for ‘Yes/No’ response boxes and other standardized lists?
- What date format should be used for subject and site completed CRF data? DD-MMM-YYYY
- What time format should be used for subject and site completed CRF data? HH:MM:SS
- Should manually-calculated data items be recorded on the CRF?

Best Practice Recommendations
-- FAQs

- Should ‘Was assessment X performed?’ be collected and/or databased? Prefer which, ‘Yes/No’ or ‘Check if ND’?
- Should “Yes/No” exam/assessment completed be preferred over “Check if not done” questions?
- Should free text be an option for a response to a specific question?
- Should data be pre-populated in the CRF?
- Should location of measurement and position of subject be collected for each assessment?
- Should sites be given guidance on how to record verbatim terms for AEs, CMs or medical history in the CRF?
- What should implementers consider when designing the flow order for CRFs?

Case 1 Necessary Data Only

Please record all previous medications used to treat hypertension

Medication Name	Total daily dose and frequency	Start Date	Stop Date	Indication

Case 2 **Should calculated data items be recorded on the CRF?**

阿尔茨海默病评价量表—认知分表（ADAS-Cog）（第 1 套）

指导语：1-7 项请研究者向患者提问，8-12 项请研究者根据对患者之前的表现作评定。

评测项目	评测内容	记分	得分
1. 单词回忆	回忆常用的单词	0 －10	|__|__| 分
2. 命名	给提交的 12 个物体和一只手上的手指命名	0 － 5	|__| 分
3. 指令	理解和完成 1～5 步命令	0 － 5	|__| 分
12. 注意力	患者测量过程中注意力	0 － 5	|__| 分
总分			|__|__| 分

- It depends

Case 3

Should “Yes/No” questions be preferred over “Tick/Check all that apply” questions?

- Ethnicity:
  - ☐ Chinese
  - ☐ Korean
  - ☐ Spanish

Case 5

Should there be a standard order for YES/NO response boxes and other standardized lists?

纳入标准	否	是
1. 年龄≥40 岁	☐1	☐2
2. 诊断为原发性高血压，且危险度分层为高危组或很高危组 (高危组及很高危组定义参照《中国高血压防治指南》2005 年修订版，依据病史判定；高血压的分级依据既往最高血压)	☐1	☐2
高危组：高血压水平属 1 级或 2 级，兼有 3 种或更多危险因素，或兼患糖尿病或靶器官损害；或高血压水平属 3 级但无其他危险因素。	☐1 是	
很高危组：高血压 3 级同时有 1 种以上危险因素或兼患糖尿病或靶器官损害；或高血压 1-3 级并有临床相关疾病。	☐1 是	

排除标准	是	否
1. 妊娠、哺乳或产后 6 个月内	☐2	☐1
2. 诊断为冠心病、脑卒中/TIA、动脉瘤、心衰，或有外周动脉疾病症状(间歇性跛行)	☐2	☐1

Summary

- Use CDASH standards in (e)CRF design as much as possible
- If no CDASH forms/variables are available, take reference to SDTM domains/variables and TA standards
- When referencing SDTM/TA standards, evaluate the variable first, instead of 100% copying
- The more closer your domains/variables are to CDASH/SDTM, the less work your programmer will need to spend time on mapping

References

- CDISC Website:
  - http://www.cdisc.org/
- CDASH Standards:
  - http://www.cdisc.org/cdash
- Controlled Terminology
  - http://www.cancer.gov/cancertopics/terminologyresources/CDISC

Ask