

临床数据传输与SDTM标准

Clinical Data Transition & SDTM Standard

研师：曹德宏

目录

概念

01

02

流程

格式

03

04

CDISC

SDTM

05



06

Q&A

- 临床数据传输

- Final Data Delivery、Data Transition、Final Data Transfer
 - 传送给申办方
 - 归档、递交
- 在临床试验结束阶段(Closeout)
 - DBL之后
 - EDC下线(Decommission/Shutdown)之前

- 临床数据传输流程

- 约定、确认数据传输时间节点、内容
 - 归档、递交
- 按约定格式、内容准备临床数据
 - XPT、SAS数据集、CSV。。。 
 - EDC数据、外部数据。。。 
- 完成数据传输
 - 电子传输
 - 电子邮件 (加密)
 - SFTP
 - 物理介质传输
 - CD
 - 优盘
 - 移动硬盘
- 双方签字确认

关于补交“临床试验数据库”资料的通知

发布日期：20170601

为加快解决审评积压，提高审评效率，更好地为注册申请人服务，避免技术审评时因缺少“临床试验数据库”资料而延误审评时间，经总局同意，现就相关注册申请人补交“临床试验数据库”资料事宜，通知如下：

一、补交“临床试验数据库”资料的品种范围
完成临床试验后申报、尚在中心审评且未提交“临床试验数据库”资料的注册申请。

二、补交时间和方式
自本通知发布之日起2个月内，申请人可以将自行已锁定的“临床试验数据库”资料直接寄送至药品审评中心业务管理处资料组。逾期未提交的，中心在审评过程将以补充资料的形式通知申请人提交。

三、补交“临床试验数据库”资料的内容

- 1、受理通知书(复印件)。
- 2、已锁定的“临床试验数据库”(光盘)中应包括原始数据库、分析数据库及相应的变量说明文件，且数据库应为SAS XPORT传输格式(即xpt格式)。

Use	Data Exchange Standard	Exchange Format	Standards Development Organization (SDO)	Supported Version
Clinical and Non-Clinical study data sets - Transport	SAS (XPORT)	XPT	SAS	5
Clinical and Non-Clinical study data sets - Transport	SAS XPORT	XPT	SAS	5

用途	规格	バージョン
データセット - 提出	SAS Transport (XPORT)	5

- 临床数据传输XPT格式

- SAS文件传输格式 (Transport File Format)是SAS开发的支持系统间数据传输的开放标准
 - SAS V5 Transport Format
 - New version Transport Format (V8、V9)

区别项目	SAS V5文件传输格式	SAS V8、V9文件传输格式
变量名长度	<=8位	<=32位
字符变量长度	<=200字节	<=32767字节
变量命名	仅允许字母、数字和下划线 ('_')	除了null ('00' x) 之外的任意字符
变量标签长度	<=40字符	<=256字符

- SAS XPORT数据引擎只支持V5文件传输格式

- 临床数据传输XPT文件生成

- SAS提供两种方式生成XPT传输格式文件
 - XPORT数据引擎, PROC COPY方式
 - 生成SAS V5传输格式文件
 - PROC CPORT/CIMPORT方式
 - 生成仅可以在SAS不同版本之间传输数据的传输格式
 - 经过压缩, 文件小

```
1) XPORT engine:  HEADER RECORD*****LIBRARY HEADER RECORD!!!!!!00000000000000
2) PROC CPORT:   **COMPRESSED** **COMPRESSED** **COMPRESSED** **COMPRESSED**
```

```
//XPORT JOB ...
// EXEC SAS
//OUTAPE DD DSN=filedesc.V612.XPT UNIT=TAPE,
// LABEL=(1,NL),VOL=SER=vvv,DISP=(OLD,KEEP),
// DCB=(RECFM=FB,LRECL=,BLKSIZE=8000)
//SASD DD DSN=projname.filedesc.SASLIB,DISP=SHR
//SYSIN DD *
LIBNAME outape XPORT;
TITLE "SEND datname1, datname2 FOR projname";
PROC COPY IN=SASD OUT=OUTAPE;
    SELECT datname1 datname2;
```

```
LIBNAME library 'c:\sasdata';
FILENAME fmtrans 'a:\fmtrans.xpt';

PROC CPORT CATALOG=library.fmtrans
    FILE=fmtrans;
RUN;
```

```
LIBNAME library 'd:\sasdata';
FILENAME fmtrans 'a:\fmtrans.xpt';

PROC CIMPORT CATALOG=library.fmtrans
    FILE=fmtrans;
RUN;
```

CDISC -- Clinical Data Interchange standard Consortium

起源于1997成立的志愿者组织

2000转为国际非盈利性组织

目前超过400家会员单位

- Pharmaceutical & biotech companies; CROs/service providers; technology providers, academic institutions, government agencies, non-profit organizations

CDISC 分支机构

- 美国、日本、欧洲、中国、韩国
- 印度、澳大利亚以及世界其他地区都有筹建中

基于共识的标准开发模式

- 公示;
- 完全开放

CDISC Standards in the Clinical Research Process

- Foundational Standard
- Therapeutic Area
- Data Exchange
- Controlled Terminology

Clinical

Plan

Collect

Organize

Analyze



Model for Planning



Model for Data Collection



Analysis Data Model

标准	描述
研究数据列表模型 (SDTM)	有关临床研究病例报告表数据标准，用于向监管部门递交的内容标准。
分析数据模型 (ADaM)	有关分析数据集及元数据的基本原则和标准，用于向监管部门递交的内容标准。
XML技术 (ODM、Define-XML与Dataset-XML)	操作数据模型 (ODM) 是基于XML概要描述如何遵循监管要求获取、交换和归档临床数据和元数据。Define-XML是基于ODM的描述研究数据集的元数据标准。Dataset-XML是基于ODM的描述研究数据集的XML Schema说明。
受控术语集 (CT)	支持CDISC模型/标准所涉及的标准词汇和编码集。
临床数据获取的协调标准 (CDASH)	用于病例报告中基础数据收集字段的内容标准。
实验室数据模型 (LAB)	描述临床实验室和研究申办者/CRO间关于临床实验室数据的获取与交换的内容标准说明细则。
非临床数据交换标准 (SEND)	描述临床前研究数据的内容标准。
方案呈现模型 (PR)	基于BRIDG模型来描述临床研究方案元素和关系的工具。
治疗领域数据标准 (TA)	为目标治疗领域确定的一套有关概念和研究终点等的标准，以提高语义的理解，支持数据共享、便于全球注册递交。如阿尔茨海默病、心血管病、糖尿病等。

CDISC标准化组织

FDA

- 2016年12月强制要求

PMDA

- 2016年10月强制要求

CFDA

- 推荐使用

Use	Data Exchange Standard	Exchange Format	Standards Development Organization (SDO)	Supported Version	Implementation Guide Version	FDA Center(s)	Date Support Begins (MM/DD/YYYY)	US/10 Support Ends (MM/DD/YYYY)	Date Requirement Begins (MM/DD/YYYY)	Date Requirement Ends	Regulatory Reference and Information Sources
Regulatory Applications (IND, NDA, ANDA, BLA master files)	Electronic Common Technical Document (eCTD)	Extensible Markup Language (XML)	International Council for Harmonisation (ICH)	3.2.2	M2 eCTD, Electronic Common Technical Document Specifications	CDER, CBER, CBER	06/01/2008		05/05/2017 [5] 05/05/2018 [6]		Electronic Submissions- Electronic Common Technical Document (eCTD)
Clinical study datasets	SDTM	XPT	CDISC	1.3	3.1.3	CDER, CBER, CBER	12/01/2012		12/17/2016 [1] 12/17/2017 [2]		CDISC.org - SDTM
Clinical study datasets	SDTM	XPT	CDISC	1.2	Version 3.1.2 Amendment 1	CDER, CBER, CBER	08/07/2013		12/17/2016 [1] 12/17/2017 [2]		CDISC.org - SDTM
Clinical study datasets	SDTM	XPT	CDISC	1.2	3.1.2	CDER, CBER, CBER	10/30/2009		12/17/2016 [1] 12/17/2017 [2]		CDISC.org - SDTM
Clinical study datasets	SDTM	XPT	CDISC	1.1	3.1.1	CDER, CBER, CBER	Ongoing	01/28/2015			CDISC.org - SDTM
Clinical study datasets	Analysis Data Model (ADaM)	XPT	CDISC	2.1	1.0	CDER, CBER, CBER	Ongoing		12/17/2016 [1] 12/17/2017 [2]		CDISC.org - ADaM
Animal study datasets	Standard for Exchange of Nonclinical Data (SEND)	XPT	CDISC	1.2	3.0	CDER	06/13/2011		12/17/2016 [1] 12/17/2017 [2]		CDISC.org - SEND

[The FDA Data Standards Catalog](#)

为了提高临床试验数据质量以及统计分析的质量和效率，方便数据的交流与汇总分析，在新药上市注册申请时，建议采用 CDISC标准递交原始数据库和分析数据库。

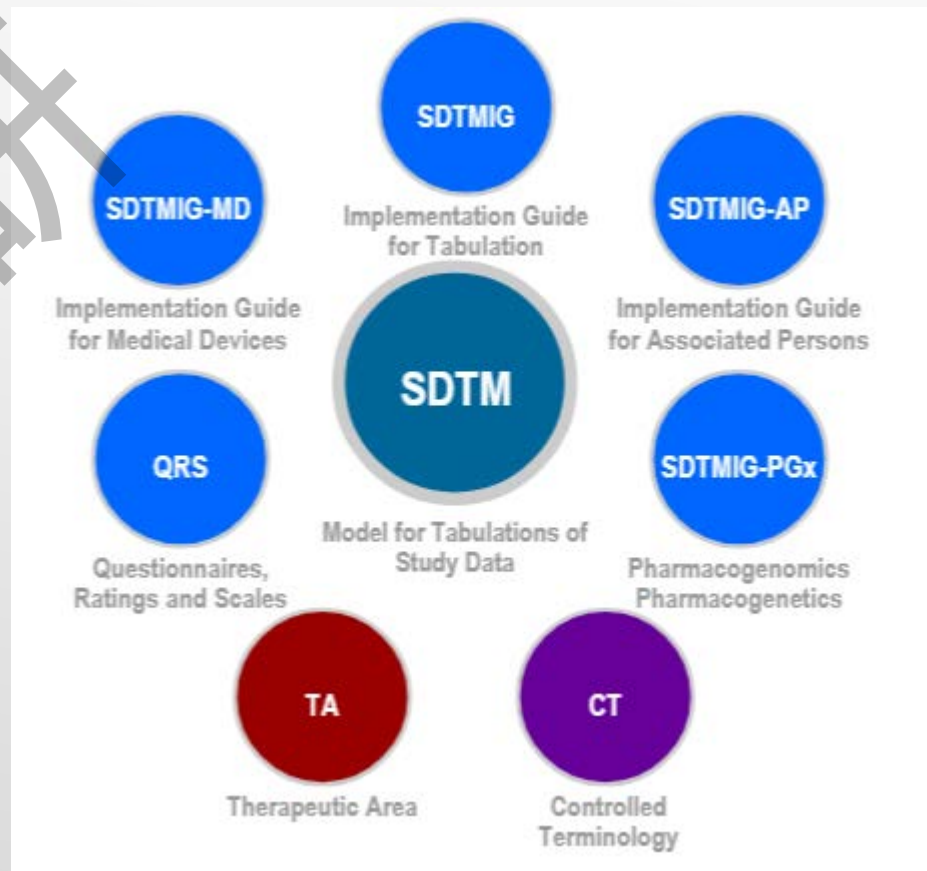
[关于发布临床试验数据管理工作技术指南的通告（2016年第112号）](#)

Study Data Tabulation Model研究数据列表模型

- 最初设计用来支持电子化信息系统和软件
- 不断优化成为监管机构的数据递交标准

一种标准化的方法来组织、格式化研究数据

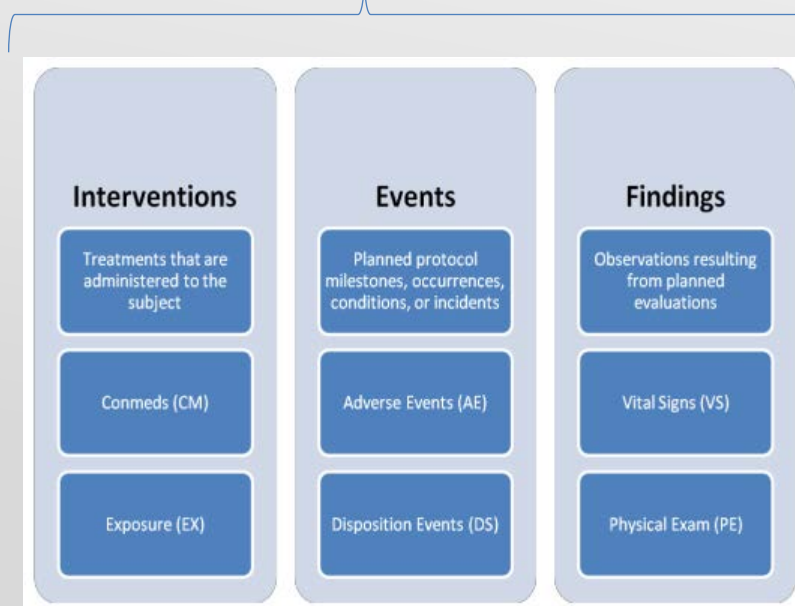
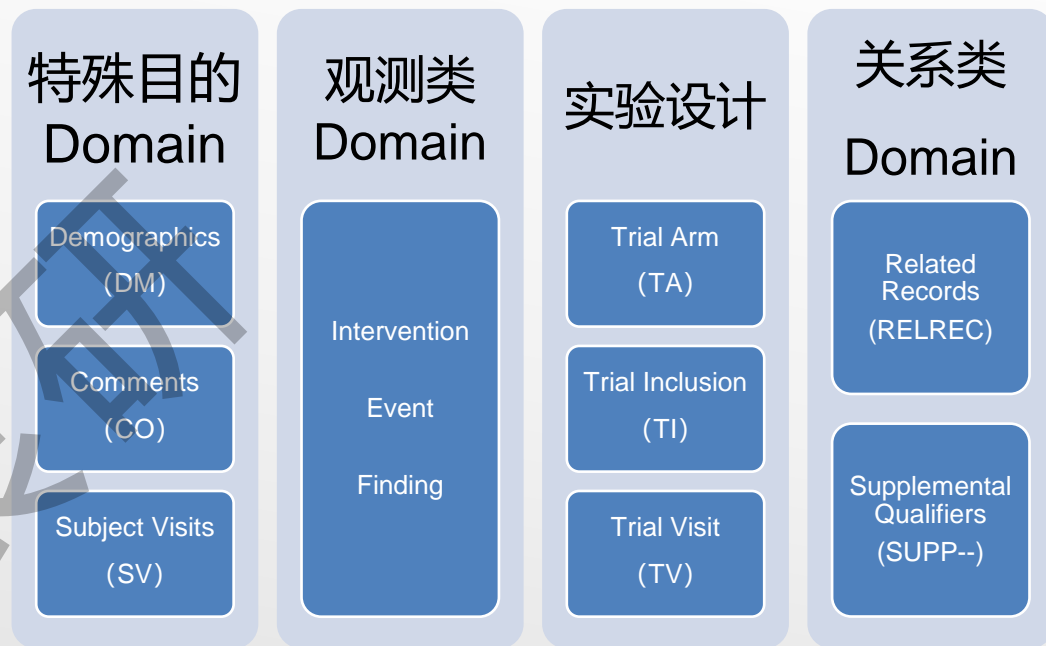
- 协助数据采集、处理、分析和报告
- 支持数据的集成和仓库存储
- 数据的充分、重复利用
- 数据的传输、交换、合并
- 极大的提高监管审阅、审评的效率
- 机器可读、识别、处理、自动化

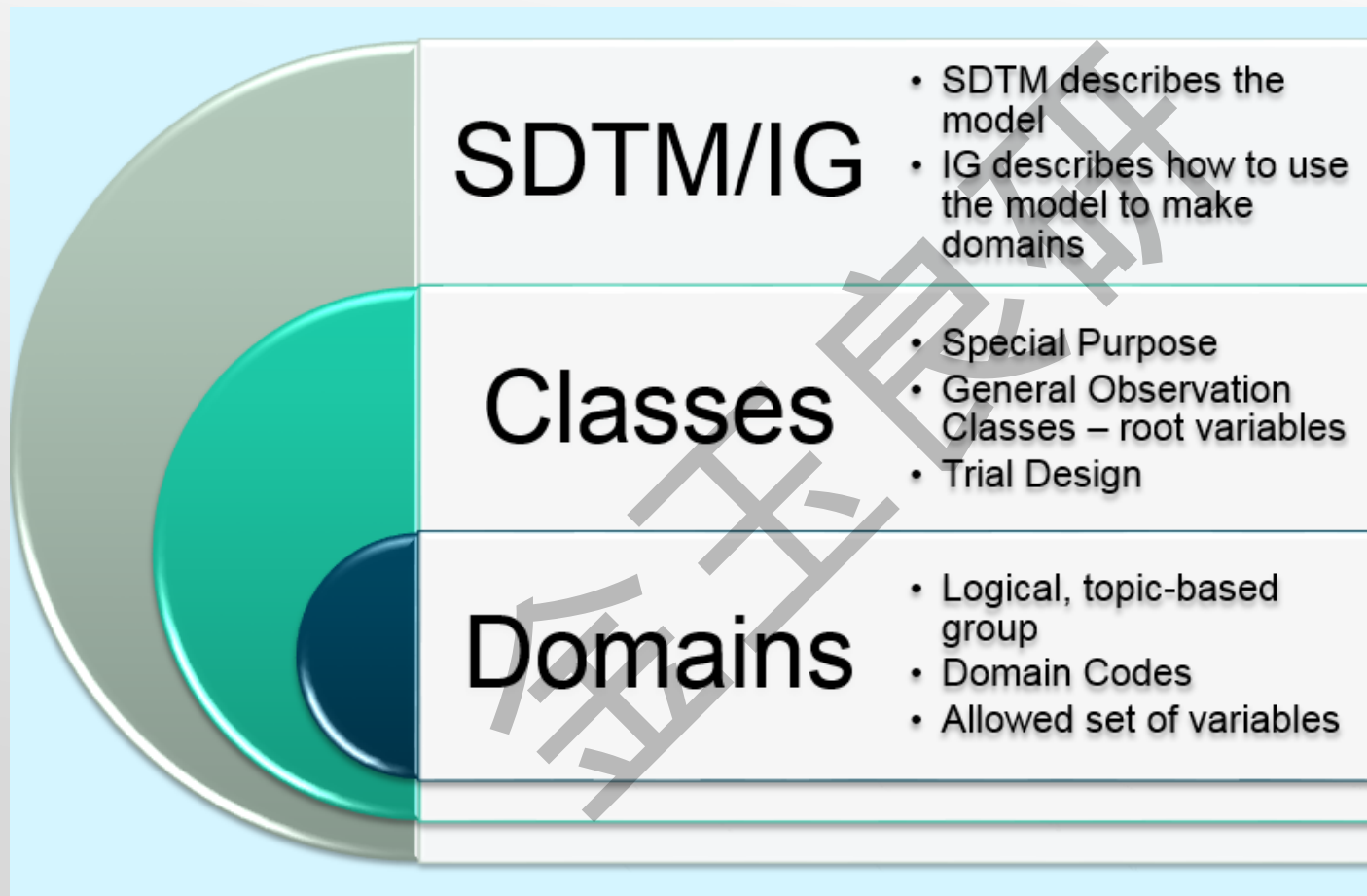


Date	SDTM版本更新、大事记
2004年6月	FDA接受CDSIC SDTM作为电子递交的可接受数据格式（SDTM v1.0模型，SDTMIG v3.1实施指南）
2005年4月 2005年8月	SDTM v1.1发布 SDTMIG v3.1.1发布
2006年12月	FDA宣布计划要求在临床数据电子递交中采用SDTM
2007年7月	SDTM v1.2, SDTMIG v3.1.2公开征求意见
2008年11月	CDISC发布SDTMIG v3.1.2
2009年11月	FDA宣布开始接受基于SDTMIG v3.1.2的递交
2011年11月	CDISC发布SDTMIG v3.1.2第一版修订
2012年7月	CDISC发布SDTMIG v3.1.3
2013年11月 2013年12月	SDTMIG v3.2发布 SDTM v1.4发布

SDTM内容组织

- 基于域Domain的数据组织结构
 - 域是根据数据描述核心点的逻辑划分
 - 每个域具有相对固定的可用变量集合
 - 域变量具有标准的命名、标签
 - 域变量具有标准的数据类型、格式以及其他属性
 - 可用受控术语 (Control Term)
 - 很多标准化的问题, 比如Test name检测项目名称
 - 根据域Domain性质分为5类
 - 目前已有超过100个不同的域结构标准化
- 标准受控术语CT (Control Term)





域变量按既定顺序排列!

Identifier	Topic	Qualifier	Timing
<ul style="list-style-type: none">Identifies the study, the subject involved in the study, the domain, and the record sequence number	<ul style="list-style-type: none">Contains the focus of the observation“What is this record about?”Domains are logical groupings of data based on the Topic	<ul style="list-style-type: none">Provides more information about the topicGroupingResultSynonymRecordVariable	<ul style="list-style-type: none">Describes the timing of the observationVisitsDate/TimesRelative times

规则变量“rule” variables – 只用作实验设计域中

CDISC SDTM Controlled Terminology

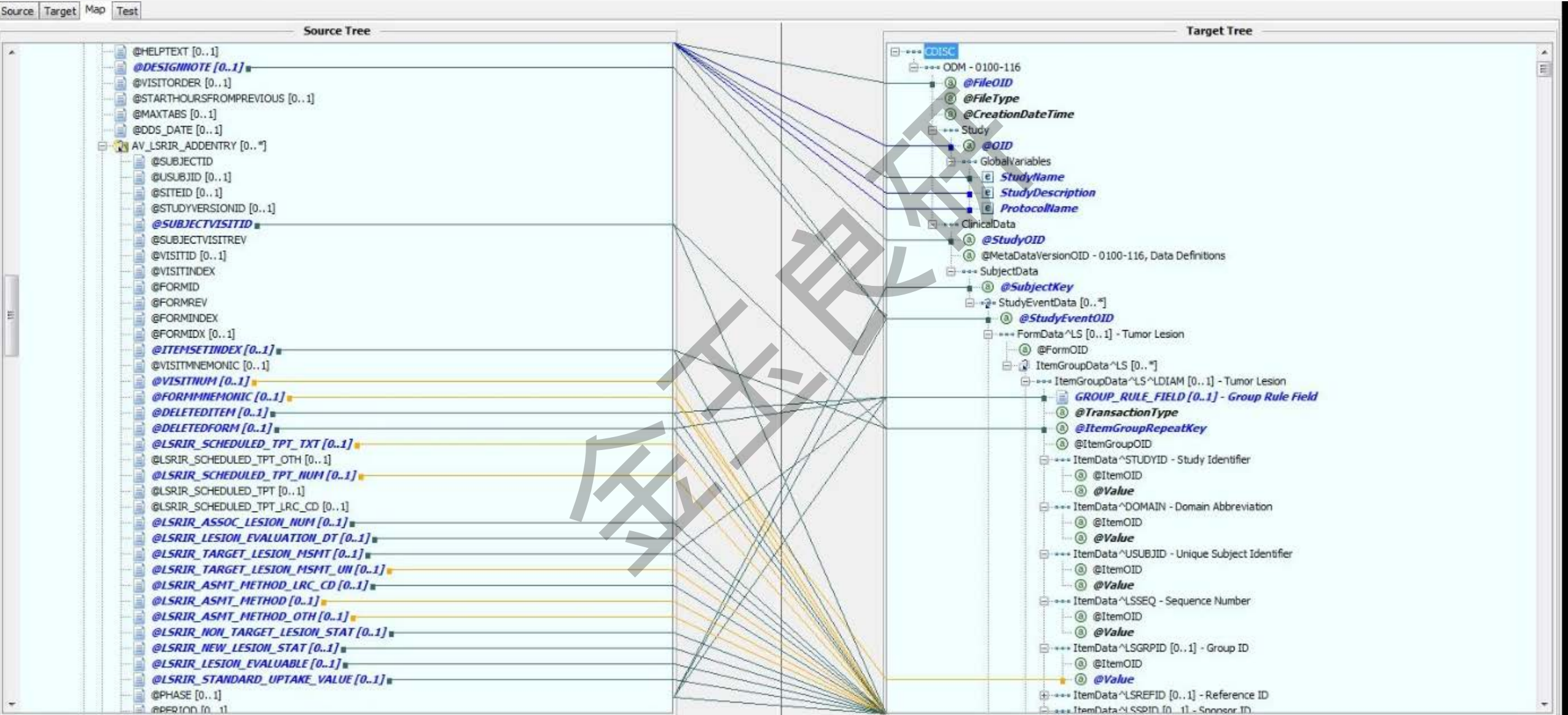
CDISC SDTM Controlled Terminology, 2016-06-24

Legend:
 CodeList
 CodeListItem
 NCI attributes

Codelist Definitions

C115388	6 Minute Walk Functional Test Test Code
C115387	6 Minute Walk Functional Test Test Name
C101805	Abnormal Involuntary Movement Scale Questionnaire Test Code
C101806	Abnormal Involuntary Movement Scale Questionnaire Test Name
C66767	Action Taken with Study Treatment
C101865	Acute Coronary Syndrome Presentation Category
C120985	Acute Physiology and Chronic Health Evaluation II Clinical Classification Test Code
C120984	Acute Physiology and Chronic Health Evaluation II Clinical Classification Test Name
C101807	ADNI Auditory Verbal Learning Functional Test Test Code
C101808	ADNI Auditory Verbal Learning Functional Test Test Name
C66781	Age Unit
C112435	Airway Questionnaire 20 Questionnaire Test Code
C112434	Airway Questionnaire 20 Questionnaire Test Name
C119062	Alcohol Use Disorder Identification Test-Consumption Questions Questionnaire Test Code
C119061	Alcohol Use Disorder Identification Test-Consumption Questions Questionnaire Test Name
C119060	Alcohol Use Disorder Identification Test- Self-Report Version Questionnaire Test Code
C119059	Alcohol Use Disorder Identification Test- Self-Report Version Questionnaire Test Name

OID	Name (CDISC Submission Value)	Data Type Extensible	NCI Code	CDISC Synonym	CDISC Definition	
CL.C115388.SIXMW1TC	6 Minute Walk Functional Test Code (SIXMW1TC)	text Extensible: No	C115388	6 Minute Walk Functional Test Test Code	6 Minute Walk Test test code.	
	SIXMW101		C115800	SIXMW1-Distance at 1 Minute	6 Minute Walk Test - Distance at 1 minute.	
	Stenosis Within a Graft		AVL01-List B Recall Word 6	C102003	AVL01-List B Recall Word 6	Auditory Verbal Learning Test Version A - List B, Word 6.
	Marital Status		AVL01-List B Recall Word 7	C102004	AVL01-List B Recall Word 7	Auditory Verbal Learning Test Version A - List B, Word 7.
	Mayo-Portland Adaptability Inventory-4 Questionnaire Test Code		AVL01-List B Recall Word 8	C102005	AVL01-List B Recall Word 8	Auditory Verbal Learning Test Version A - List B, Word 8.
	Mayo-Portland Adaptability Inventory-4 Questionnaire Test Name		AVL01-List B Recall Word 9	C102006	AVL01-List B Recall Word 9	Auditory Verbal Learning Test Version A - List B, Word 9.
Back to top	Mayo-Portland Adaptability Inventory-4 Questionnaire Test Name					
CL.C11	Mechanical Circulatory Support Placement Timing					
	Medical Evaluator					
	Medical History Event Date Type					
	Method					
Back to top	Michigan Neuropathy Screening Instrument Questionnaire Test Code					
	Michigan Neuropathy Screening Instrument Questionnaire Test Name					
	Microbial Identification Certainty					
Back to top	CL.C66781.AGEU	Age Unit (AGEU)	text Extensible: No	C66781	Age Unit	Those units of time that are routinely used to express the age of a subj
	DAYS		C25301	Day	The time for Earth to make a complete rotation on its axis; ordinarily div refers to a specific day. (NCI)	
	HOURS		C25529	Hour,hr	A unit measure of time equal to 3,600 seconds or 60 minutes. It is appr	
	MONTHS		C29846	Month	One of the 12 divisions of a year as determined by a calendar. It corres approximately to one cycle of the moon's phases, about 30 days or 4 w	
	WEEKS		C29844	Week	Any period of seven consecutive days. (NCI)	
	YEARS		C29848	Year	The period of time that it takes for Earth to make a complete revolution days; a specific one year period. (NCI)	
Back to top	CL.C112435.AQ01TC	Airway Questionnaire 20 Questionnaire Test Code (AQ01TC)	text Extensible: No	C112435	Airway Questionnaire 20 Questionnaire Test Code	Airway Questionnaire 20 test code.
	AQ0101		C112529	AQ01-Cough Often During Day	Airway Questionnaire 20 - Do you cough often during the day?	
Back to top	CL.C66731.SEX	Sex (SEX)	text Extensible: No	C66731	Sex	The assemblage of physical properties or qualities by which male is distinguished fr difference between male and female ; the distinguishing peculiarity of male or female .
	F		C16576	Female	A person who belongs to the sex that normally produces ova. The term is used to inc distinctions, or cultural gender role distinctions, or both. (NCI)	
	M		C20197	Male	A person who belongs to the sex that normally produces sperm. The term is used to distinctions, cultural gender role distinctions, or both. (NCI)	
	U		C17998	U;Unknown	Not known, not observed, not recorded, or refused. (NCI)	
	UNDIFFERENTIATED		C45908		A person (one of unisexual specimens) who is born with genitalia and/or secondary s indeterminate sex, or which combine features of both sexes. (NCI)	



SDTM在临床研究中的实践

VISIT 1 VISITNUM	
Not done <input type="checkbox"/> VSSTAT	
VITAL SIGNS	
Weight	<input type="text"/> <input type="text"/> <input type="text"/> <input type="text"/> kg
Height	<input type="text"/> <input type="text"/> <input type="text"/> <input type="text"/> cm
Heart Rate	<input type="text"/> <input type="text"/> <input type="text"/> <input type="text"/> beats/min
Blood Pressure	<input type="text"/> <input type="text"/> <input type="text"/> <input type="text"/> / <input type="text"/> <input type="text"/> <input type="text"/> <input type="text"/> mmHg
	Systolic SYSBP Diastolic DIABP

VISIT 1 VISITNUM		
Not done <input type="checkbox"/> VSSTAT		
VITAL SIGNS		
VSTESTCD when VSTESTCD="WEIGHT" Weight	VSORRES when VSTESTCD="WEIGHT" <input type="text"/> <input type="text"/> <input type="text"/> <input type="text"/> kg	VSORRESU when VSTESTCD="WEIGHT"
VSTESTCD when VSTESTCD="HEIGHT" Height	VSORRES when VSTESTCD="HEIGHT" <input type="text"/> <input type="text"/> <input type="text"/> <input type="text"/> cm	VSORRESU when VSTESTCD="HEIGHT"
VSTESTCD when VSTESTCD="HR" Heart Rate	VSORRES when VSTESTCD="HR" <input type="text"/> <input type="text"/> <input type="text"/> <input type="text"/> beats/min	VSORRESU when VSTESTCD="HR"
VSORRES when VSTESTCD="SYSBP" Blood Pressure	VSORRES when VSTESTCD="DIABP" <input type="text"/> <input type="text"/> <input type="text"/> <input type="text"/> / <input type="text"/> <input type="text"/> <input type="text"/> <input type="text"/> mmHg systolic diastolic	VSORRESU when VSTESTCD="SYSBP" OR "DIABP"
VSTESTCD when VSTESTCD="SYSBP"	VSTESTCD when VSTESTCD="DIABP"	

SDTM在临床研究中的实践

VITAL SIGNS

WEIGHT

Weight | 8 | 5 | kg

HEIGHT

Height | 1 | 8 | 0 | cm

HR

Heart Rate | 7 | 0 | beats/min

Systolic
SYSBP

Diastolic
DIABP

Blood Pressure | 8 | 0 | / | 1 | 2 | 0 | mmHg

PATIENT	WEIGHT (kg)	HEIGHT (cm)	HR (bpm)	DBP (mmHg)	SBP (mmHg)
001	85	180	70	80	120

USUBJID	VSTESTCD	VSTEST	VSORRES	VSORRESU
CH-001	WEIGHT	Weight	85	kg
CH-001	HEIGHT	Height	180	cm
CH-001	HR	Heart Rate	70	bpm
CH-001	DIABP	Diastolic Blood Pressure	80	mmHg
CH-001	SYSBP	Systolic Blood Pressure	120	mmHg

Not done **VSSSTAT**

VITAL SIGNS

VSTESTCD when **VSTESTCD="WEIGHT"** **VSORRES** when **VSTESTCD="WEIGHT"** **VSORRESU** when **VSTESTCD="WEIGHT"**

Weight | |__| 8 |__| 5 | kg

VSTESTCD when **VSTESTCD="HEIGHT"** **VSORRES** when **VSTESTCD="HEIGHT"** **VSORRESU** when **VSTESTCD="HEIGHT"**

Height | | 1 |__| 8 |__| 0 | cm

VSTESTCD when **VSTESTCD="HR"** **VSORRES** when **VSTESTCD="HR"** **VSORRESU** when **VSTESTCD="HR"**

Heart Rate | |__| 7 |__| 0 | beats/min

VSORRES when **VSTESTCD="SYSBP"** **VSORRES** when **VSTESTCD="DIABP"** **VSORRESU** when **VSTESTCD="SYSBP" OR "DIABP"**

Blood Pressure | |__| 8 |__| 0 |__| / | | 1 |__| 2 |__| 0 |__| mmHg

systolic diastolic

VSTESTCD when **VSTESTCD="SYSBP"** **VSTESTCD** when **VSTESTCD="DIABP"**

USUBJID	VSTESTCD	VSTEST	VSORRES	VSORRESU
CH-001	WEIGHT	Weight	85	kg
CH-001	HEIGHT	Height	180	cm
CH-001	HR	Heart Rate	70	bpm
CH-001	DIABP	Diastolic Blood Pressure	80	mmHg
CH-001	SYSBP	Systolic Blood Pressure	120	mmHg

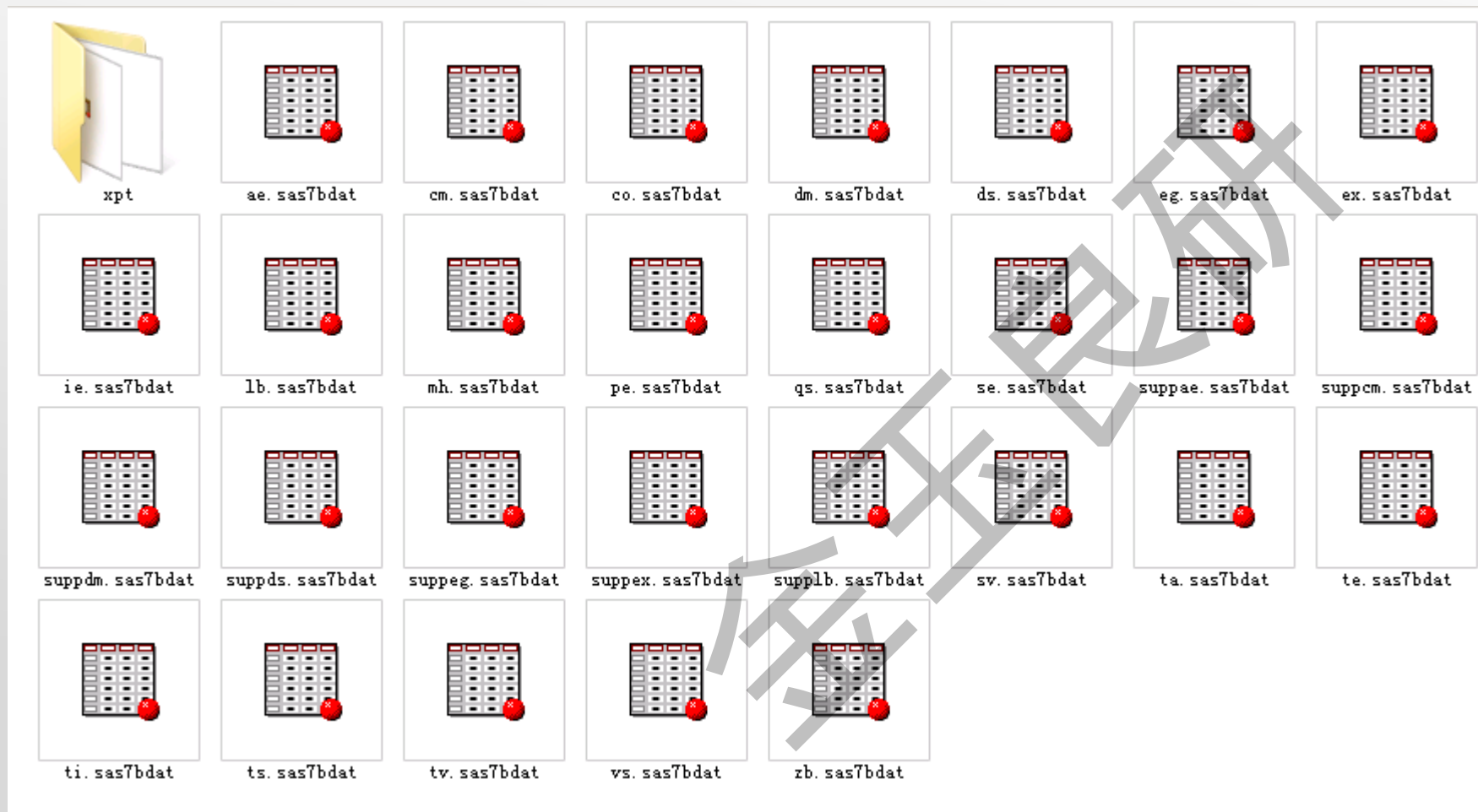
- [-] Annotated Case Report Form
- [-] Datasets
- [+] Value Level Metadata
- [-] Computational Algorithms
- [+] Controlled Terminology

Datasets for Study CDISC01

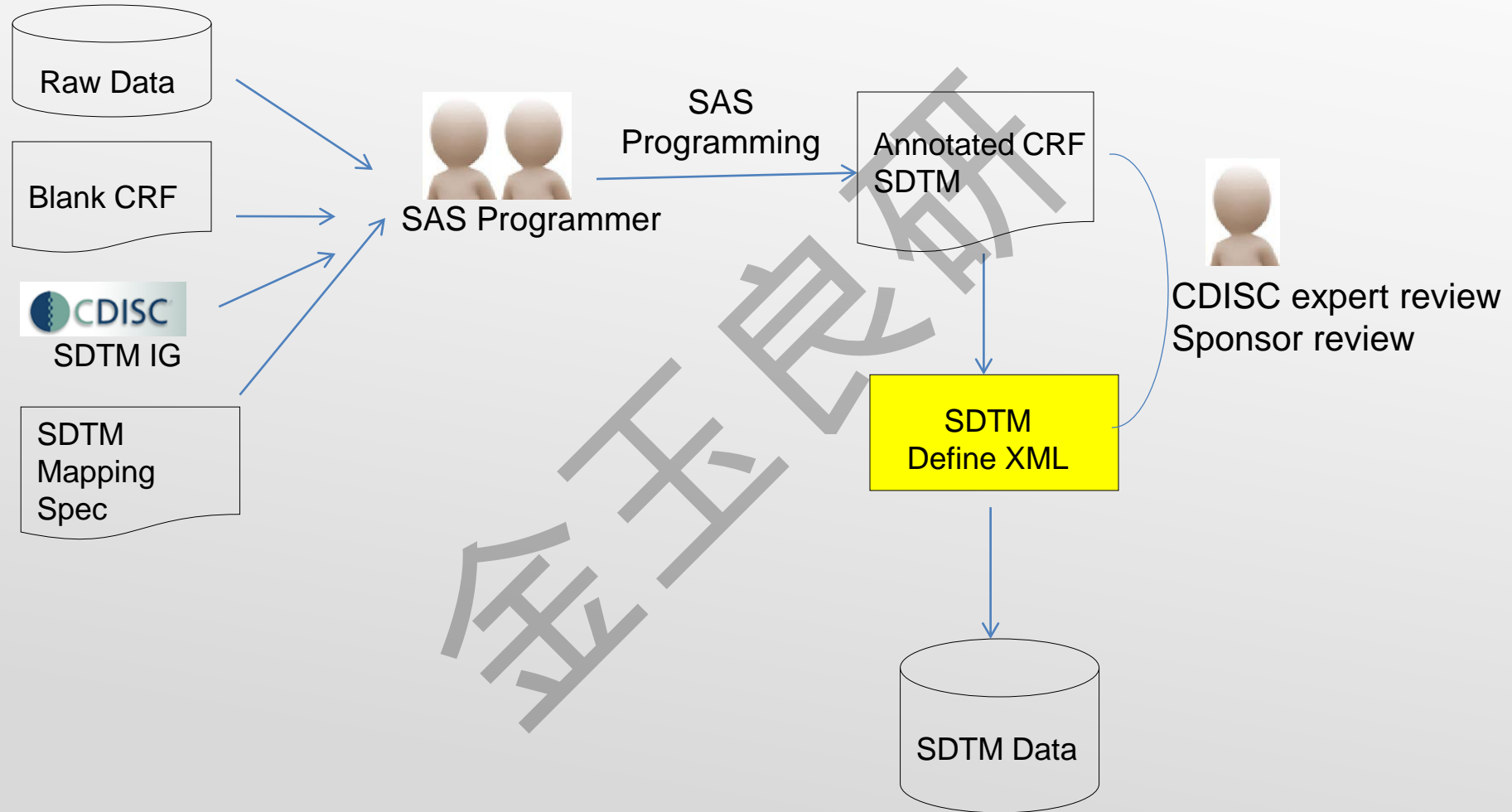
Dataset	Description	Class	Structure	Purpose	Keys	Location
SV	Subject Visits	TRIAL DESIGN	One record per subject per actual visit	Tabulation	STUDYID, USUBJID, SVSTDTC, SVENDTC, VISITNUM, SVUPDES	sv.xpt
TA	Trial Arms	TRIAL DESIGN	One record per element per arm	Tabulation	STUDYID, ARMCD, TAETORD	ta.xpt
TE	Trial Elements	TRIAL DESIGN	One record per element	Tabulation	STUDYID, ETCD	te.xpt
TI	Trial Inclusion	TRIAL DESIGN	One record per I/E criterion	Tabulation	STUDYID, IETESTCD	ti.xpt
TS	Trial Summary	TRIAL DESIGN	One record per subject	Tabulation	STUDYID, SUBJID, TSTESTCD, TSTESTCD	ts.xpt

Adverse Events Dataset (AE) [ae.xpt](#)

Variable	Label	Type	Controlled Terminology	Origin	Role	Comment
STUDYID	Study Identifier	text		CRF	IDENTIFIER	
DOMAIN	Domain Abbreviation	text	DOMAIN	Assigned	IDENTIFIER	
USUBJID	Unique Subject Identifier	text		Derived	IDENTIFIER	Concatenation of STUDYID.SUBJID
AESEQ	Sequence Number	integer		Derived	IDENTIFIER	Sequential number identifying AEs within each USUBJID and generated using the key sequence of the AE domain
AESPID	Sponsor-Defined Identifier	text		Assigned	IDENTIFIER	ID of original SAS dataset
AETERM	Reported Term for the Adverse Event	text		CRF	TOPIC	



SDTM在临床研究中的实践





Q&A

金玉良研



Thank You !
