



2023

CHINA

INTERCHANGE

BEIJING | 25-26 AUGUST



中国CDISC实施专家共识框架 v0.1

吴崇胜(迪时咨询), 王军(康迪亚), 王海波(普瑞盛), 赖认真(信达生物), 魏志军(诺华)等



Meet the Speaker

Haibo Wang, 王海波

Title: Director, DM & Statistics

Organization: GCP ClinPlus Co., Ltd.

- 超过12年行业经验
- 参与了多个领域共150余项海外及国内临床试验项目SAS编程工作，其中多个产品成功获得FDA及EMA上市批准
- 目前担任生物统计、统计程序、数据管理、工具研发四个部门超过250人团队管理工作
- 中山大学预防医学毕业



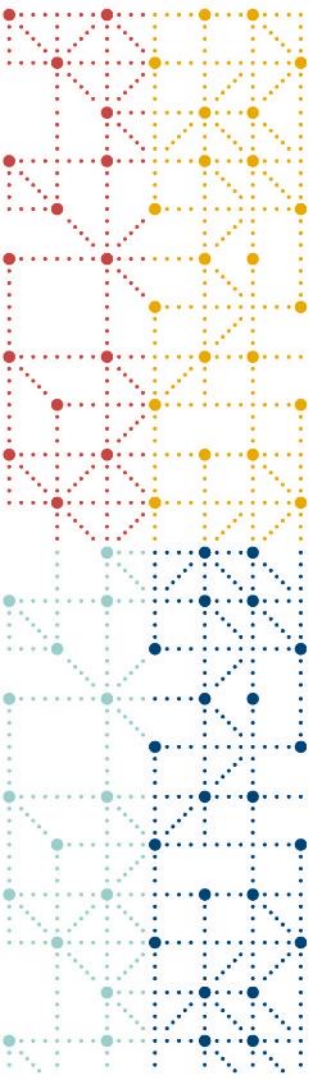
Disclaimer and Disclosures

- *The views and opinions expressed in this presentation are those of the author(s) and do not necessarily reflect the official policy or position of CDISC.*
- *The author(s) have no real or apparent conflicts of interest to report.*



Agenda

- 本共识框架目的
- 背景介绍
- CDISC标准介绍
- 实施CDISC总的策略
- 核验
- 参考文献

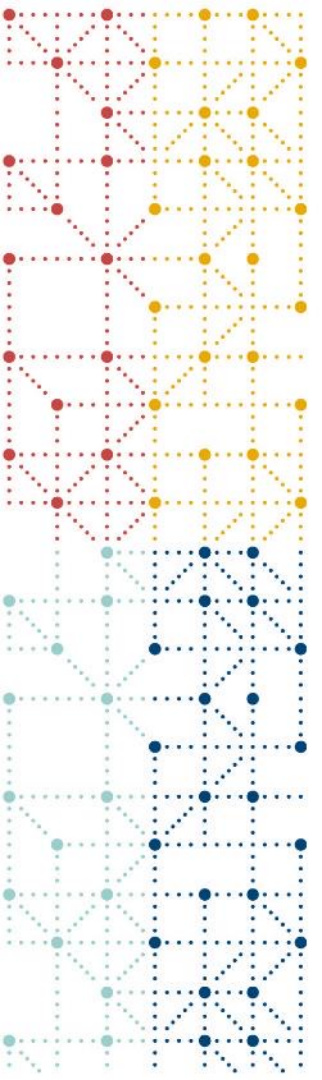


一、本共识框架目的



本共识框架目的

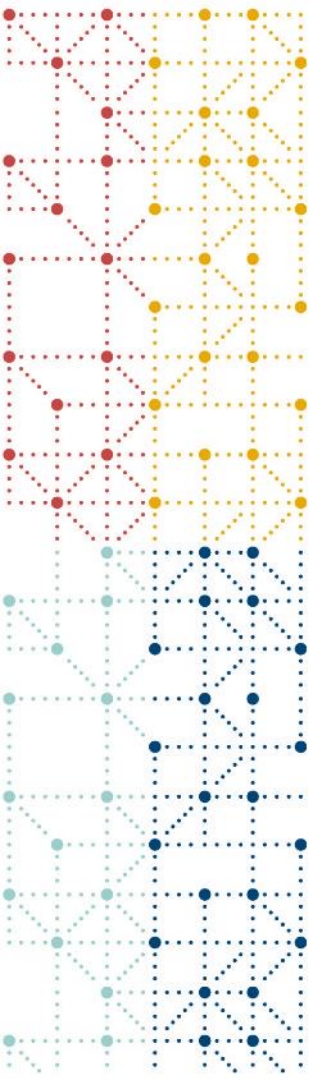
本共识框架的目的：指导CDISC标准在中国临床试验中的落地实施。



二、背景介绍

背景介绍

- 美国FDA强制要求2016年12月17日之后开始的研究数据递交需满足CDISC标准。
- 日本PMDA强制要求2020年03月31日起数据递交满足CDISC标准。
- 2022年，欧洲EMA启动数据递交pilot study，参与pilot递交要求遵循CDISC的SDTM和ADaM标准，define.xml为可选。
- 2020年7月国家药监局药审中心发布《药物临床试验数据递交指导原则（试行）》，对数据递交的提出了具体要求，并明确“鼓励申办方参照临床数据交换标准协会（CDISC）标准递交临床试验数据及相关的申报资料”。



三、CDISC标准介绍

CDISC标准介绍

- 数据标准化(内容标准), 是从数据集水平(命名, 描述, 结构, 允许的变量)、变量水平(命名, 标签, 类型, 格式)及值水平(受控术语与外部词典)进行了约定, 以标准的方式组建与展示数据。
- CDISC是致力于数据标准开发的非盈利组织, CDISC标准涵盖了从研究设计、采集、转换、分析、报告的各个环节, 应用最为广泛的CDISC标准如CDASH(临床数据采集标准)、SDTM(原始数据标准模型)、ADaM(分析数据标准模型)、define.xml(数据定义标准)。

SDTM基本原理

- SDTM(Study Data Tabulation Model , **原始数据标准模型**)。是标准化的组织和显示研究数据的模型。
- SDTM是针对**原始数据**(即采集的数据，无填补、很少量的衍生)的标准化。
- 通常情况下，所有研究受试者的观测数据都会收录到一系列不同的域中。**域 (Domain)**是一组具有共同主题并且逻辑相关的观测结果的集合。
- 大多数在试验过程中采集到的受试者级别的观测数据，都可以被归为干预类(Interventions)、事件类(Events)和发现类(Findings)三大通用观测**类**中的某一类。SDTM还包含3类特殊的数据集：特殊目的数据集(人口学[DM]，备注[CO]，受试者元素[SE]和受试者访视[SV])，试验设计模型 (TDM) 数据集，关联数据集(如RELREC)。
- **变量**可被归纳为以下5种主要角色：标识符 (Identifier)变量、主题 (Topic)变量、时间 (Timing)变量、修饰语 (Qualifier)变量、规则 (Rule)变量。

ADaM基本原理 – ADaM原则

➤ 分析数据集及相关元数据必须：

- 促进清晰无歧义的交流
- 提供分析数据和它的源数据之间的可溯源性
- 能很容易地在常用软件工具上使用

➤ 分析数据集必须：

- 伴有元数据
- 即可分析

➤ 实施考虑：

- 必须包括ADSL
- 数据集命名遵循ADxxxxxx规范
- 尽量使用ADaM标准变量,遵循命名规范
- 同名、同义、同值(SDTM变量拷贝)
- 多个研究间的一致性

□ **可溯源**：从元数据可溯源(Metadata traceability)与数据点可溯源(Data point traceability)二个途径来实现：（1）在metadata中交代清楚来源、衍生/填补规则（2）保留来源变量/记录。

□ **即可分析**：需要或很少量编程来处理数据，即可对分析进行再现；可以有冗余变量。如subset, sort是可以的，需要transpose则不符合要求。但不是指一个表只能用一个分析数据集，同一统计表格，可以用多个数据集、多个统计过程步。

ADaM基本原理 – ADaM数据结构

- ADaM发布的数据结构目前有三种：（1）ADSL(受试者水平分析数据集)；（2）BDS (基本数据结构); (3) OCCDS(发生类数据结构)。
- ADSL：每个受试者一条记录。
- BDS: 每个受试者每分析参数每个分析时间点一条记录。如生命体征分析数据集(advs)、实验室检查分析数据集(adlb)。
- OCCDS：每个受试者每个事件/干预方式一条记录。如不良事件分析数据集(adae)、病史分析数据集(admh)、既往与合并用药分析数据集(adcm)。



四、实施CDISC总的策略

1) 《药物临床试验数据递交指导原则(试行)》概要

指导原则对数据递交包要求总结如下：

➤ 相关文件齐全、格式正确

- ✓ 数据说明文件：xml: define.xml(推荐)、pdf: define.pdf (不推荐)
- ✓ aCRF.pdf: 不是EDC数据集的aCRF, 而是递交的xpt原始数据集对应的aCRF
- ✓ 数据审阅说明: pdf (原始数据库/分析数据库), 说明数据集所用编码(encoding): 如utf-8, euc-cn
- ✓ 数据文件格式：XPT V5或更高版本；XPT v8 (推荐)。注意：SAS CPORT做的xpt是不符合要求的

➤ 必备数据集

- ✓ 原始数据库：人口学数据集 – dm.xpt
 - 命名为dm; 每人一条记录；必备变量：STUDYID/USUBJID/SUBJID
- ✓ 分析数据库：受试者水平分析数据集 – adsl.xpt
 - 命名为adsl; 每人一条记录；包括人口学、重要的基线特征/分层因素、治疗组、预后因素、重要日期、分析人群划分等信息

➤ 其他数据集

- ✓ 必须变量：STUDYID/USUBJID
- ✓ 原始数据库：数据集命名参照CDISC：ae, cm, ds, lb...；VISITNUM、VISIT 1:1对应
- ✓ 分析数据库：通常以“adxxxxxx”命名; 分析数据集的命名应尽量与原始数据集保持对应，如：adcm、adae、adlb等

2)中国实施CDISC标准策略

结合中国临床数据标准推广、相关人才储备等现状，并参照FDA和PMDA相关数据递交标准要求，本共识框架对中国临床数据标准实施策略推荐分步骤按以下3个等级进行，并逐步提高标准化程度：

- **最低标准**：满足目前NMPA法规/指南对数据递交的要求
- **推荐标准**：满足最低标准要求；数据集与变量满足CDISC标准对数据集(结构、标签)和变量(命名、标签)要求；常用术语遵循CDISC CT。
- **最佳实践**：满足推荐标准要求，满足CDISC标准对受控术语要求
 - 以上所提到的CDISC标准包括SDTM(原始数据标准模型)和ADaM(分析数据标准模型)。在数据采集阶段，为了方便后续的数据标准化，建议遵循CDASH标准。
 - 常用 CT：对于常用受控术语，C3C遴选了最常用的受控术语进行中文翻译。常用CT中文翻译版，可通过CDISC官网下载，具体地址如下：<https://www.cdisc.org/translations/chinese>

2)中国实施CDISC标准策略

- 原始数据库常用变量举例如下, 详细的请参考SDTM IG V3.3:

变量名	变量标签(EN)	变量标签(CN)	Comments
STUDYID	Study Identifier	研究标识符	指南要求
USUBJID	Study Identifier	研究标识符	指南要求
SUBJID	Subject Identifier for the Study	受试者标识符	人口学(dm), 指南要求
SITEID	Study Site Identifier	研究中心标识符	人口学(dm)
SEX	Sex	性别	人口学(dm)
AGE	Age	年龄	人口学(dm)
AGEU	Age Units	年龄单位	人口学(dm)
ARM	Description of Planned Arm	计划分组描述	人口学(dm)
ACTARM	Description of Actual Arm	实际分组描述	人口学(dm)
VISITNUM	Visit Number	访视编号	指南提及
VISIT	Visit Name	访视名称	指南提及
--DTC	Date/Time of Measurement/Collection	采集/检查日期/时间	发现类、事件类、干预类
--STDTC	Start Date/Time	开始日期/时间	事件类、干预类

2)中国实施CDISC标准策略

- 原始数据库常用变量举例如下, 详细的请参考SDTM IG V3.3:

变量名	变量标签(EN)	变量标签(CN)	Comments
--ENDTC	End Date/Time	结束日期/时间	事件类、干预类
--TERM	Reported Term	报告名称	事件类
--DECOD	Dictionary-Derived Term	标准化名称	事件类
--BODSYS	Body System or Organ Class	系统器官分类	事件类
--TRT	Reported Name of Drug, Med, or Therapy	治疗报告名称	干预类
--DECOD	Standardized Medication Name	标准化名称	干预类
--CLAS	Medication Class	归类	干预类
--TESTCD	--Test Short Name	检查项简称	发现类
--TEST	--Test Name	检查项名称	发现类
--ORRES	Result or Finding in Original Units	原始结果	发现类
--ORRESU	Original Units	原始单位	发现类
--STRESC	Character Result/Finding in Std Format	标准化结果	发现类
--STRESN	Numeric Result/Finding in Standard Units	标准化结果(N)	发现类
--STRESU	Standard Units	标准化单位	发现类

2)中国实施CDISC标准策略

分析数据库常用变量举例如下, 详细的请参考ADaM IG V1.2/1.3

➤ ADSL :

- ✓ 标识符: STUDYID , USUBJID , SUBJID, SITEID,
- ✓ 分析人群: FASFL, SAFFL, ITTFL, PPROTFL, COMPLFL, RANDFL, ENRFL
- ✓ 人口学: AGE, AGEU, SEX, RACE
- ✓ 治疗组 : ARM, ACTARM, TRTxxP(N), TRTxxA(N)

.....

➤ BDS结构 :

- ✓ PARAM, PARAMCD, AVAL, AVALC, ABLFL, BASE, CHG, PCHG, AVISIT(N), ADT, ATM
- ✓ Time-to-Event还包括 : STARTDT, CNSR(N), EVNTDESC, CNSDTDSC
- ✓ 标帜变量: ANLxxFL

.....

➤ OCCDS结构, 来源数据集的基础上, 根据分析需要 : 增加

- ✓ 字符型变量对应的数值型变量
- ✓ 时间变量(ASTDT, AENDT, ADURN/ADURU),
- ✓ 标帜变量(TRTEMFL)

3)外文数据递交包翻译

《药物临床试验数据递交指导原则(试行)》四.(三)节“外文数据库”中，对外文数据库翻译提出了具体要求。

- ✓ 翻译的**一致性**：不同文件之间的中文表述应保持一致，如分析数据集与CSR中的不良事件名称（作者注：指编码后的标准术语）
- ✓ **数据库**中至少以下内容为中文：数据集标签、变量标签；CSR中出现的不良事件名称、合并用药名称、病史名称（作者注：指编码后的标准术语）。
- ✓ **数据说明文件**中至少以下内容为中文：数据库中各数据集的描述/标签和说明；数据集中各变量的描述/标签和衍生过程；涉及疗效指标的取值或编码列表。
- ✓ **注释病例报告表**中至少以下内容应为中文：为了收集数据所设计的问题描述；涉及疗效指标问题的取值或编码。
- ✓ 数据审阅说明应为中文

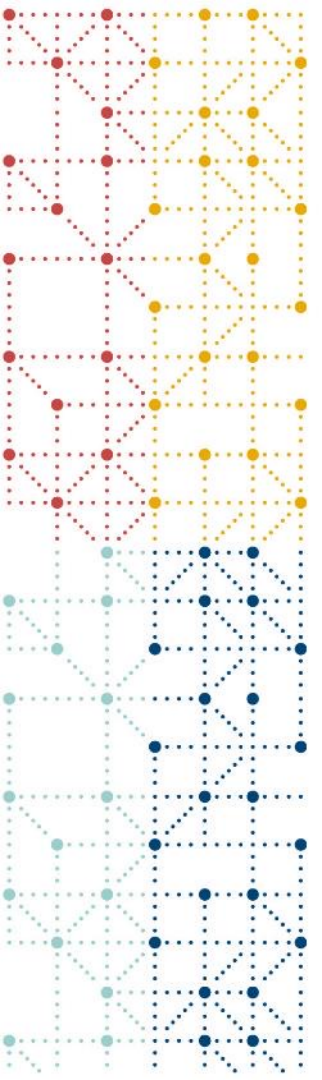
作者注：原则上，预期审阅人员要阅读、复核的内容，均应翻译为中文，并在原始数据库、分析数据库及相关文档中保持翻译的准确性、一致性。

4) 中医药数据递交

目前可以参考的CDISC中文翻译文件有：

- ◆ 中医针刺治疗领域标准用户指南 (版本1.0) ，已发布
- ◆ 中医冠心病心绞痛治疗领域数据标准用户指南1.0版，已发布

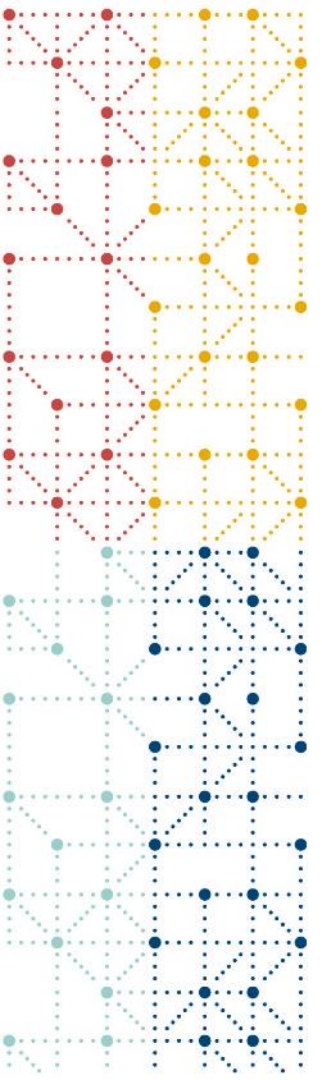
可以在CDISC官网， <https://www.cdisc.org/translations/chinese> 下载到以上翻译文档。



五、核验

核验

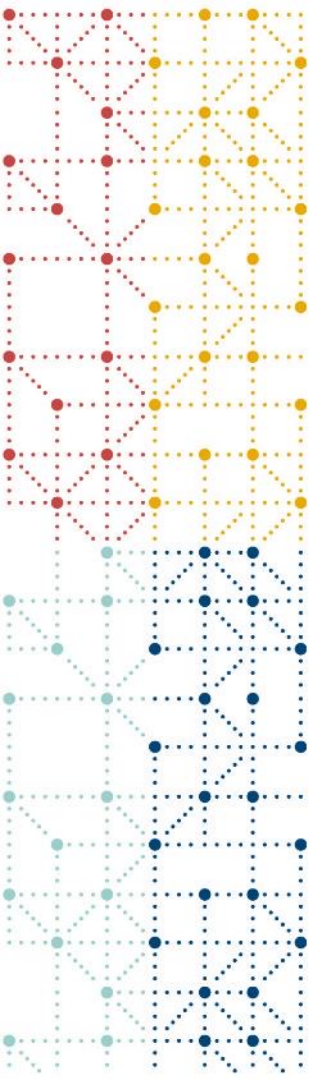
- 临床数据标准化顾问组对《药物临床试验数据递交指导原则（试行）》（国家药监局药审中心于 2020年7月20日发布）进行解析，开发了《**NMPA数据递交核验规则**》，以便于业界递交数据前可以进行核验，见单独的文档。
- 可用市售验证工具或自己单位开发的工具进行数据递交验证，以确定是否符合指导原则递交要求。对于不符合核验规则/指导原则要求的，应在《数据审阅说明》中进行解释。



六、参考文献

参考文献

- 国家药监局药审中心. 《药物临床试验数据递交指导原则（试行）》. 2020年7月20日发布
- CDISC. CDISC Standards. www.cdisc.org/standards



Thank You!

cdisc