



2023

CHINA  
INTERCHANGE

BEIJING | 25-26 AUGUST



## CDISC China Update and Metadata Driven Automation

Victor WU

CDISC China Coordinating Committee(C3C)

北京迪时咨询有限公司



## Meet the Speaker

Victor Wu

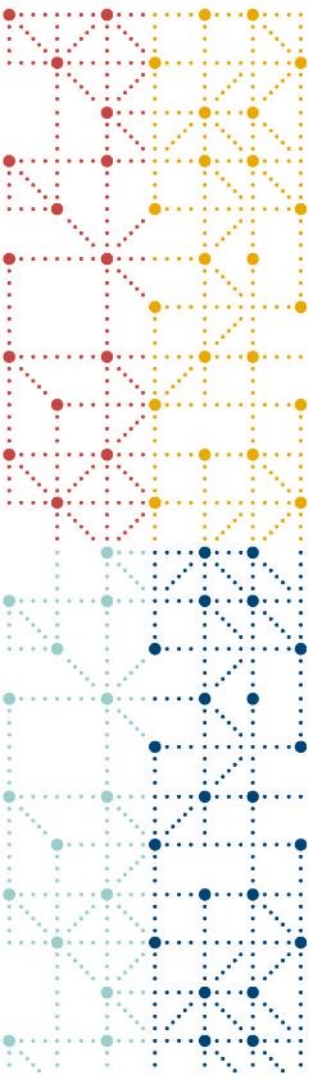
**Title:** EVP, Data Science

**Organization:** 北京迪时咨询有限公司

Victor WU(吴崇胜), Ph.D., CDISC中国协调委员会(C3C)主席, CDISC官方认证SDTM及ADaM讲师; 北京迪时咨询有限公司共同创立者; 临床数据标准化顾问组主席、北京生物医学统计与数据管理研究会(BBA)常务理事。在数据标准化翻译与推广、流程规范化、自动化工具开发方面有丰富的经验, 为行业贡献良多。

# 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.*
- *{Please disclose any financial relationship or conflict of interest relevant to this presentation here OR}*
- *The author(s) have no real or apparent conflicts of interest to report.*



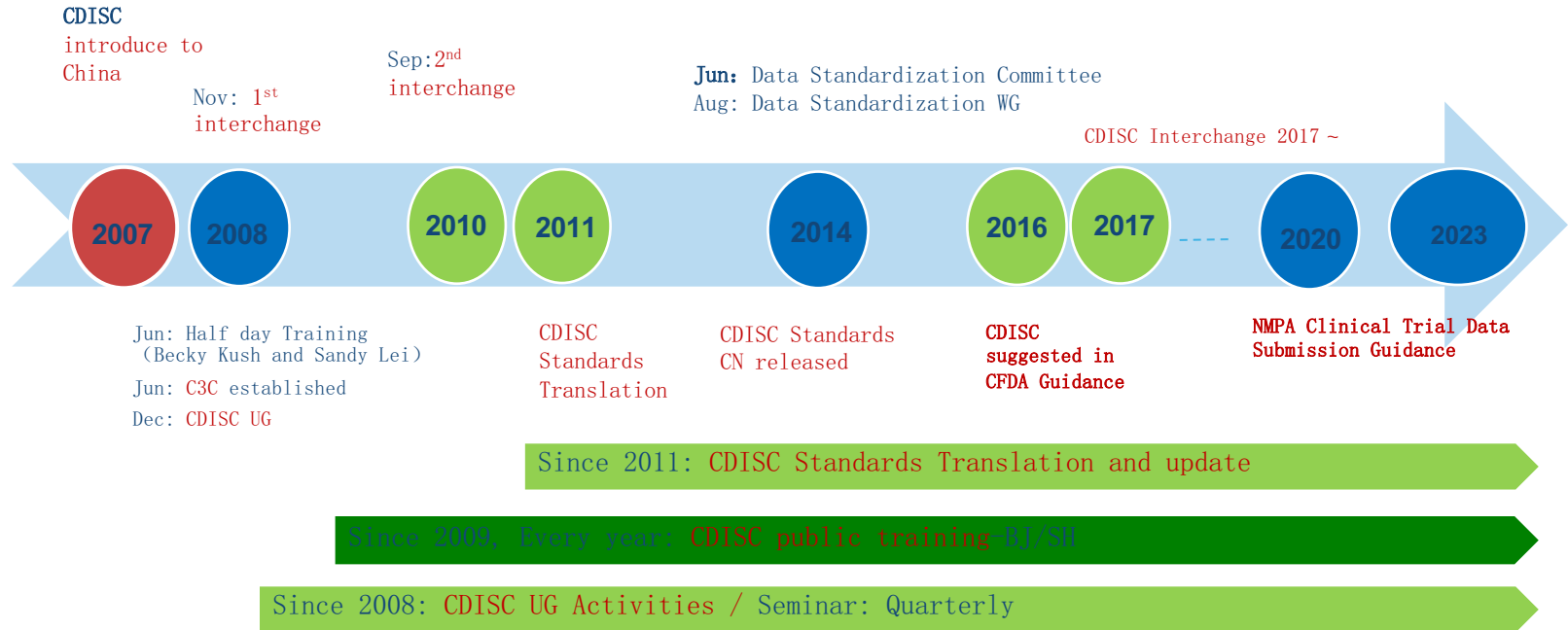
# Agenda

1. CDISC in China
2. Metadata Driven Automation
3. How to Study CDISC

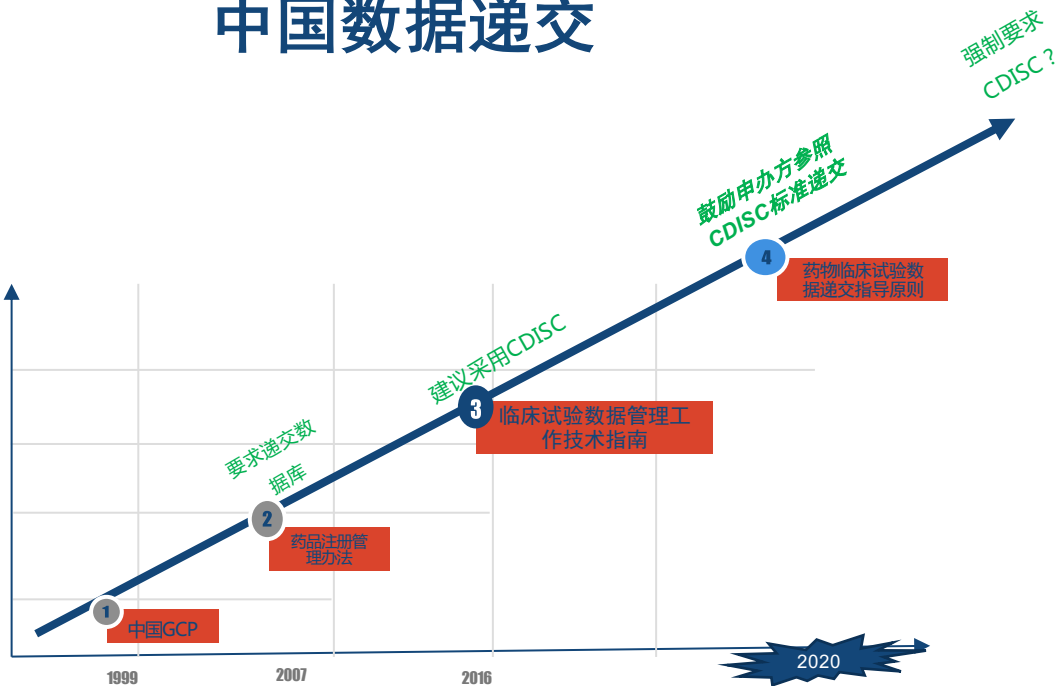


# 1. CDISC in China Update

# CDISC in China (2007-2021)



# 中国数据递交



- 2007: Provisions for Drug Registration 《药品注册管理办法》
- database submission was required in NDA

- 2016: Technical Guidance on Data Management 《临床试验数据管理工作技术指南》
- ...Suggested to adopt **CDISC** standards to submit raw database and analysis database.

○ 2020 Guideline on the Submission of Clinical Trial Data 《药物临床试验数据递交指导原则(试行)》

- **encouraged** to submit clinical trial data following **CDISC** standards

# 《药物临床试验数据递交指导原则（试行）》（NMPA 2020）

国家药监局药审中心关于发布《药物临床试验数据递交指导原则（试行）》的通告（2020年第16号）

发布日期：20200720

为规范药品注册申请人递交药物临床试验数据及相关资料，配合新修订的药品注册申报资料要求，提高药品审评效率，药审中心组织制定了《药物临床试验数据递交指导原则（试行）》（见附件）。根据《国家药监局综合司关于印发药品技术指导原则发布程序的通知》（药监综药管〔2020〕9号）要求，经国家药品监督管理局审核同意，现予发布。

化学药品、生物制品自2020年10月1日起实施。中药实施日期按国家药监局发布中药注册分类及申报资料要求的通告中相关规定执行。

特此通告。

国家药品监督管理局药品审评中心  
2020年7月20日

<http://www.cde.org.cn/news.do?method=viewInfoCommon&id=7a43c3abfde95950>



# Why Standardization from Regulatory Reviewer's View?

Standardized collection, organization, analysis, and presentation of clinical trial data play an important role in improving the efficiency and quality of clinical researches and development, **shortening review timelines**.

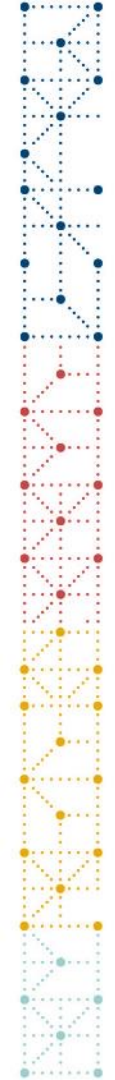
【**规范地**收集、整理、分析和呈现临床试验数据对于提高药物临床研发的效率和质量、**缩短审评时间**具有重要的作用，并且有利于药品全生命周期管理，促进研发或监管的信息互通或共享。】

If the clinical trial data submitted by sponsor does not follow certain specifications, it will take significant resources for reviewers to get familiarized with and understand the data structure and content.

【申办方递交的临床试验数据若不遵循一定的规范，熟悉和理解数据结构及内容将**占用大量的审评资源**。】

Sponsor is encouraged to submit clinical trial data and the associated materials according to CDISC standards

【**鼓励**申办方以**CDISC标准**递交临床试验数据及相关的申报资料。】



If the clinical trial data submitted by sponsor does not follow certain specifications, it will take significant resources for reviewers to get familiarized with and understand the data structure and content.

【申办方递交的临床试验数据若不遵循一定的规范，熟悉和理解数据结构及内容将占用大量的审评资源。】

【某些情况下，申办方或监管机构可能需要针对多来源的临床试验数据进行**汇总分析**，数据没有规范化，**整合利用**也几乎难以实现。】

In some cases, sponsor or regulatory agency may need to conduct pooled analyses using multi-sources of clinical trial data. Non-standardized data will make this task almost impossible.

# NMPA递交准备Tips

- 相关文件齐全、格式正确
  - 数据说明
    - 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)
- 数据文件
  - 文件格式: XPT V5或更高版本
    - XPT v5: SAS (proc copy)
    - XPT v8 (推荐): <http://support.sas.com/kb/46/944.html>
    - ⊗ 注意: SAS CPORT做的xpt是不符合要求的
  - Encoding: 需要在《数据审阅说明》中说明
    - UTF8: 推荐
    - EUC-CN

- 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
    - 必须变量: 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
    - VISITNUM、VISIT——对应
    - 数据集命名参照CDISC: ae, cm, ds, lb, vs, qs...
  - 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

申办方应【在数据审阅说明中】说明所用编码（Encoding, 如utf-8、euc-cn等），以避免所递交的数据集出现乱码的情形

# 外文数据库 Foreign language database

## • What to translate 翻译内容

- ✓ 数据库：CSR报表中涉及的内容 Values referred in CSR:
  - AE/MH/CM 名称: *coded values*
  - 疗效指标
  - ...
- ✓ 数据说明文件meta: 数据集标签、变量标签；衍生过程说明；涉及疗效指标的取值或编码列表
- ✓ 数据审阅说明
- ✓ aCRF：问题描述；涉及疗效指标问题的取值或编码

C3C解读: 相关文件、数据集/变量标签、CSR中图表中涉及的数据内容应为中文.

# CDISC Standards Translation

目前可以参考的CDISC中文翻译文件有：

- SDTM metadata (IG V3.2/3.3), 待发布
- ADaM metadata (IG V1.2/1.3), 待发布
- 常用受控术语文件, 已发布 (更新版准备中, call for volunteers)
- SDTMIG v3.2, 已发布
- ADaMIG v1.1、ADaM OCCDS v1.0, 待发布
- 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。



# 临床数据标准化顾问组

- Volunteers/Representatives from CRO/Pharma/Institution
- 《中国CDISC实施专家共识框架》 draft 2023-08
- 《NMPA数据递交核验规则》 draft 2023-08
- 临床数据科学SOP/Templates

# CDISC China Activities 2023

- CDISC Public Training, 北京22-24Aug2023
- CDISC China Interchange, 北京25-26Aug2023
- CDISC User Group 活动: Oct ~ Dec 2023 planning

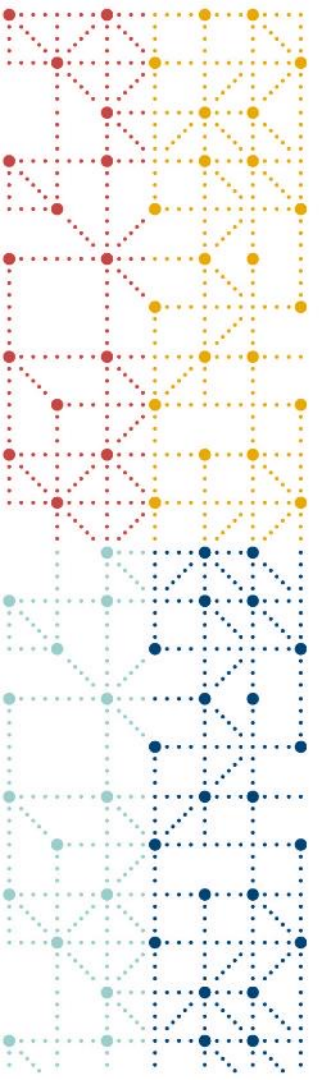


- CDISC webinar

<https://www.cdisc.org>



A screenshot of the CDISC website. The top navigation bar includes a search box and menu items: 'New to CDISC', 'Standards', 'Education', 'Resources', 'Events' (highlighted with a red box), 'Membership', and 'Members Only'. Below the navigation bar is a large banner image showing a man presenting to a group of people seated at a table. The banner text reads 'Greater Clarity Starts Here' and 'CDISC Education'. To the right of the banner is a sidebar menu with sections: 'Conferences' (listing 2023 China Interchange, 2023 CDISC TMF Interchange, 2023 US Interchange, 2023 Korea Interchange, and 2024 Europe Interchange), 'Webinars' (highlighted with a red box, listing 'Upcoming Webinars', 'Public Webinars Archive', and 'Members Only Webinar Archive'), and 'All Events' (listing 'Calendar', 'COSA', and 'Partner Events').

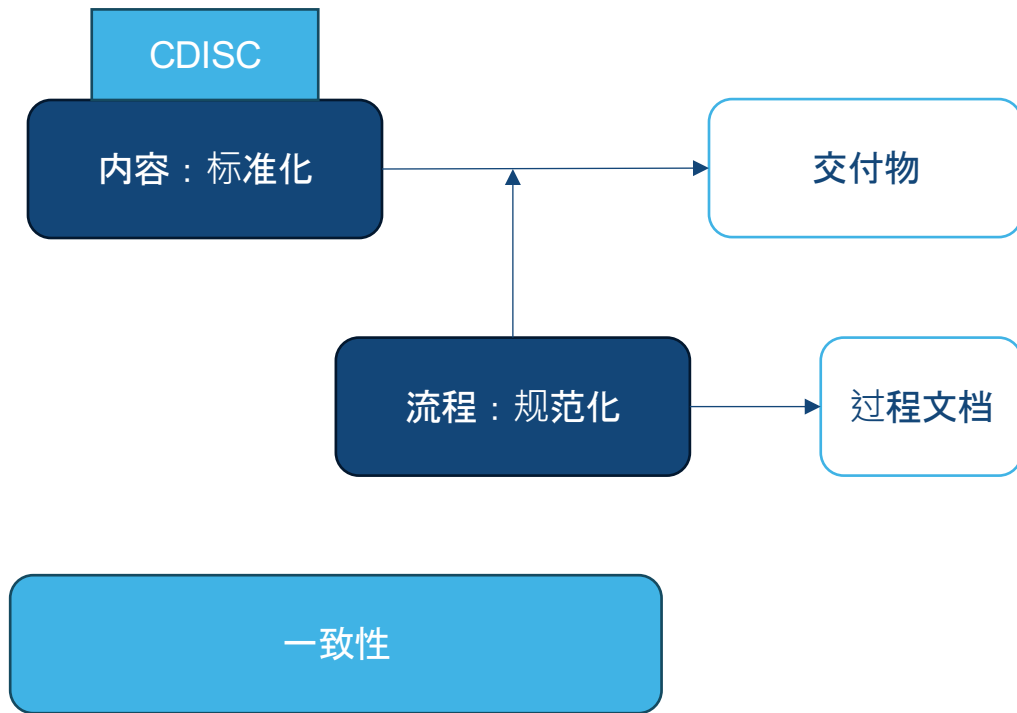


## 2. Metadata Driven Automation





# 标准化



# Standards-based/Metadata-driven Automation



Standardized Contents

Standardized Process



Automation

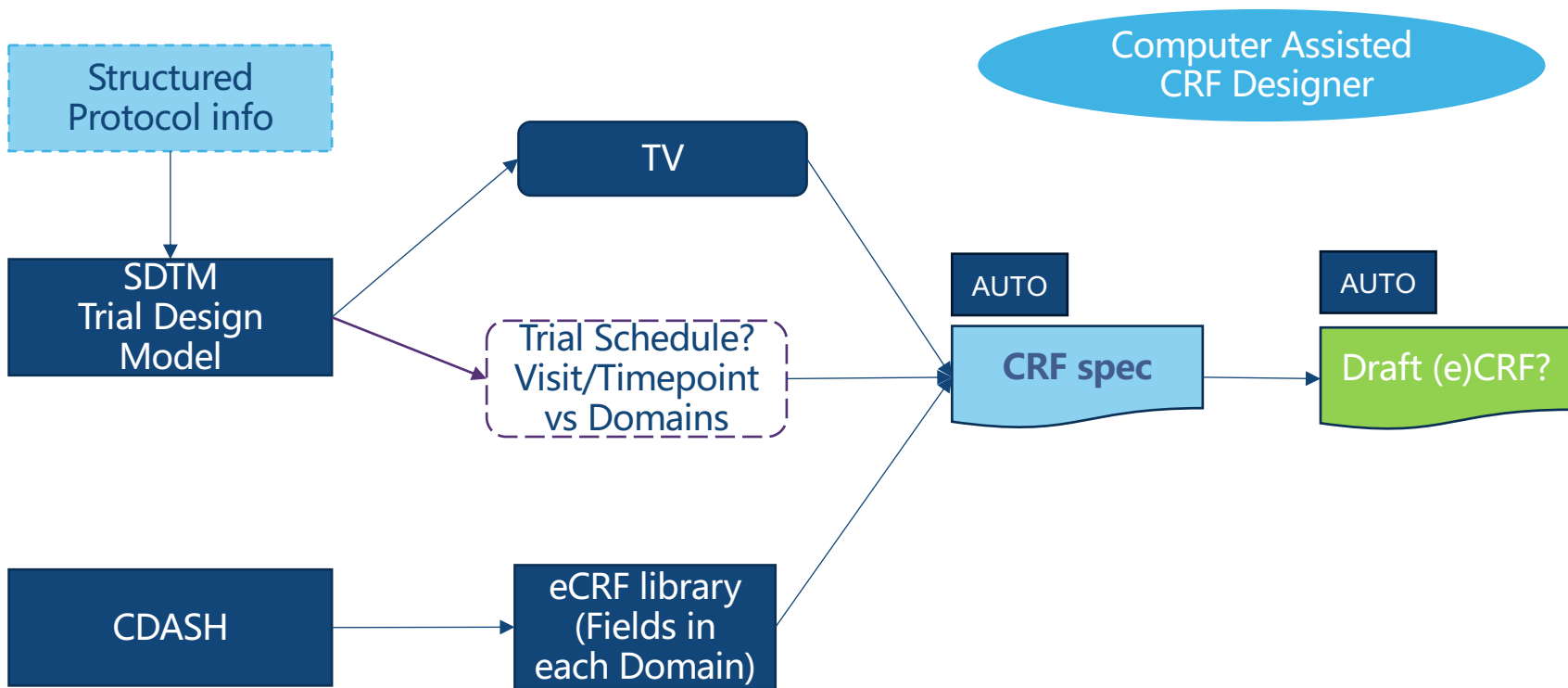
Tools

Metadata

Flexibility

AI Assisted

# Automation from protocol to eCRF?



# How about data clean/review?

Data review, to base on  
standardized or original data?

eCRF library  
CDASH

Develop Standard  
Edit-check

Tabulated Data  
SDTM

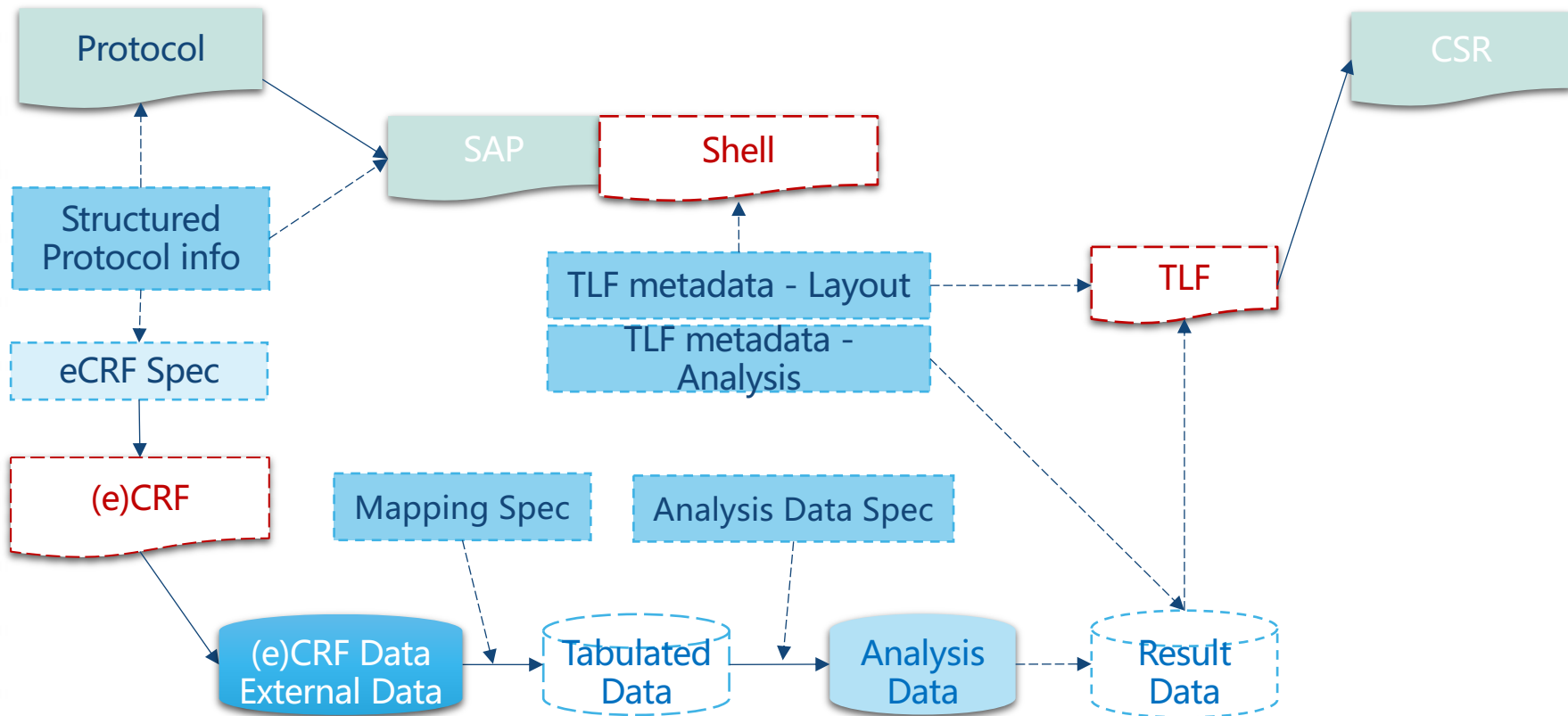
Conformance  
Validation Rules  
(data issue related)

Standard Data  
Clean Rules

Computer Assisted  
EC Designer

Computer Assisted  
Data Review

# Computer/AI Assisted Data Flow



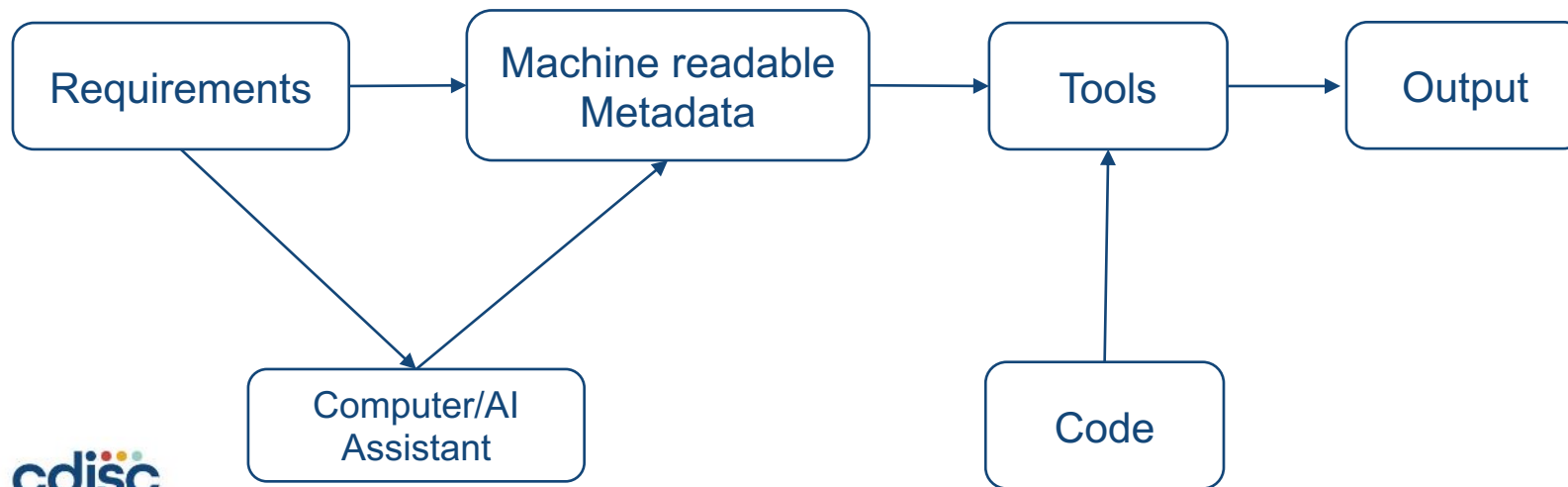
SAP=Statistical Analysis Plan; TLF = Table, Listing and Figure; CSR= Clinical Study report

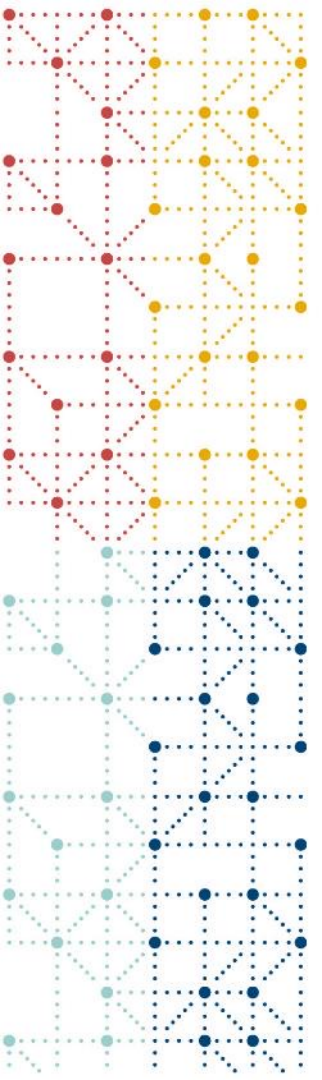
# What to do with programming?

old



new

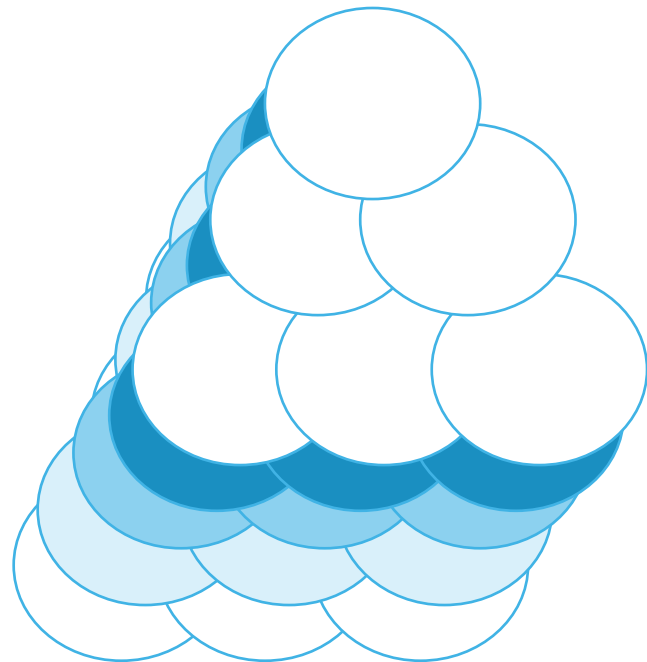




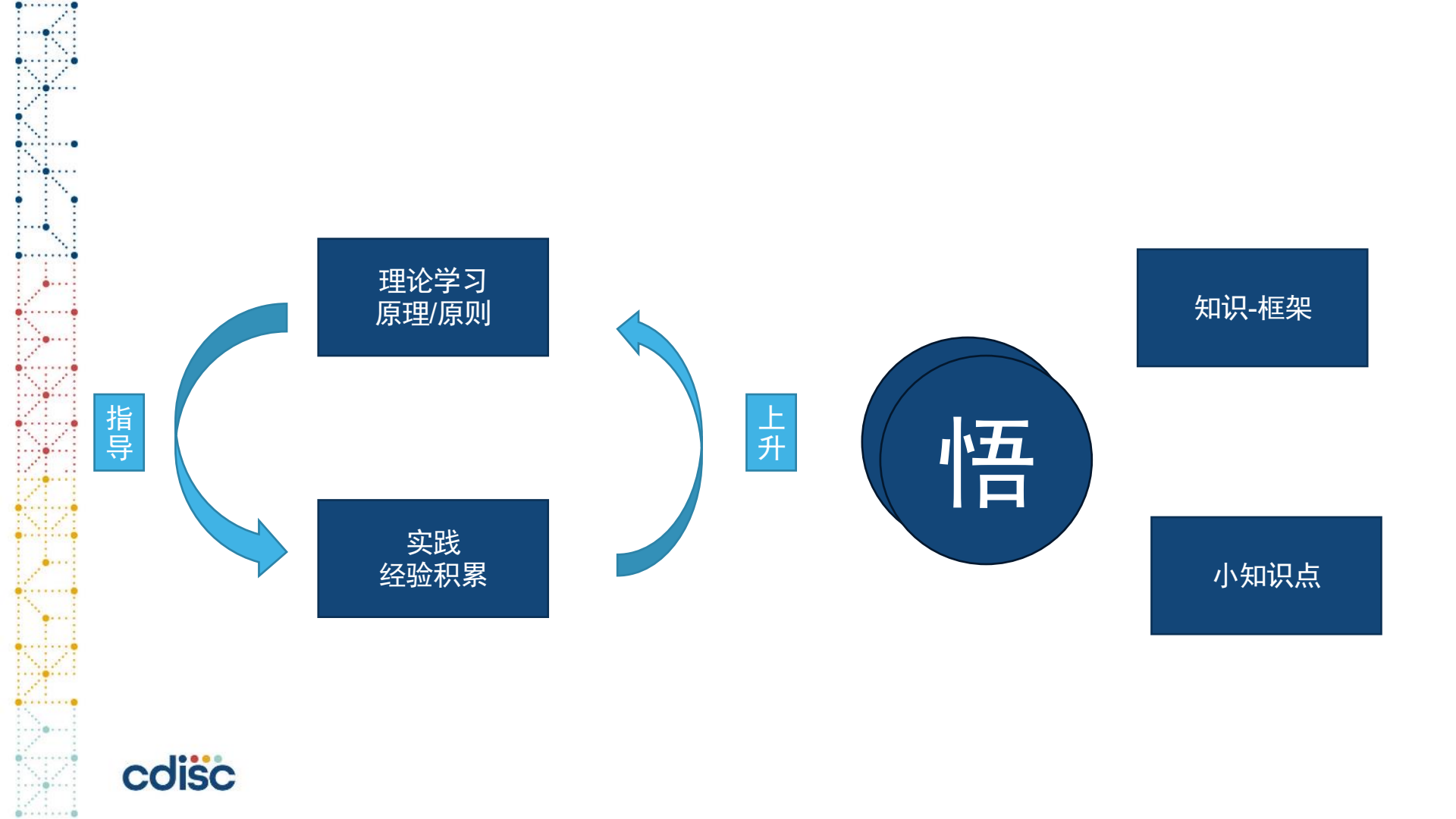
### 3. How to Study CDISC

# Path of CDISC Study

- 1.1 高度概括的概述 – 知识框架（标准化）
- 1.2 相对细节的概述 - 知识框架（具体标准）
  - ✓ 如：迪时的Overview Webinar/研讨班
  - ✓ IG目录
- 2.1 细节的讲解与精读
  - ✓ CDISC官方培训班
  - ✓ IG的精读
- 2.2 揣着IG实战，随时翻阅
- 2.3 交流、讨论、Paper、分享 – 经验积累
  - ✓ CDISC interchange：官方组织，每年
  - ✓ CDISC用户活动：北、上、广...线下活动即将恢复
  - ✓ 迪时组织的advanced webinar、研讨班：尽量每年
  - ✓ 微信群：CDISC用户群、数据科学研讨群...
- 3.0 升华，回归到理论、框架 – 思与悟
  - ✓ 如：迪时的Overview Webinar/培训班

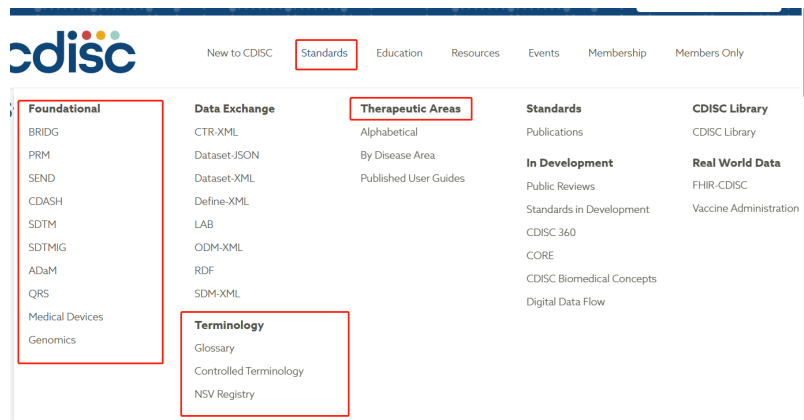






# CDISC resources

- <https://www.cdisc.org>



- FDA

## Study Data Standards Resources

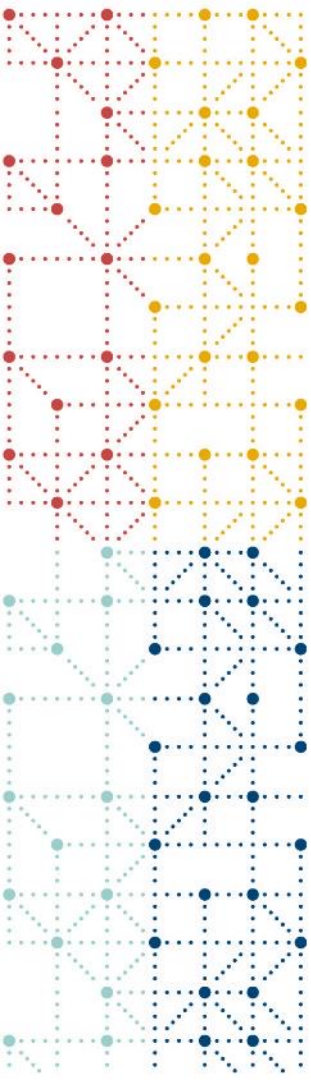
1. FDA Data Standards Catalog

2. FDA Guidances

3. Technical Guides

4. FDA Business and Validator Rules

<https://www.fda.gov/industry/fda-data-standards-advisory-board/study-data-standards-resources>



# Thank You!

[victor.wu@datascie.com](mailto:victor.wu@datascie.com)

**cdisc**

