



2023

CHINA

INTERCHANGE

BEIJING | 25-26 AUGUST



Build SDTM Automation Platform

Presented by Chong Liu, SAS programmer II, Program Department, GCP ClinPlus Co., Ltd



Meet the Speaker

Chong Liu

Title: Build SDTM Automation Platform

Organization: GCP ClinPlus Co., Ltd

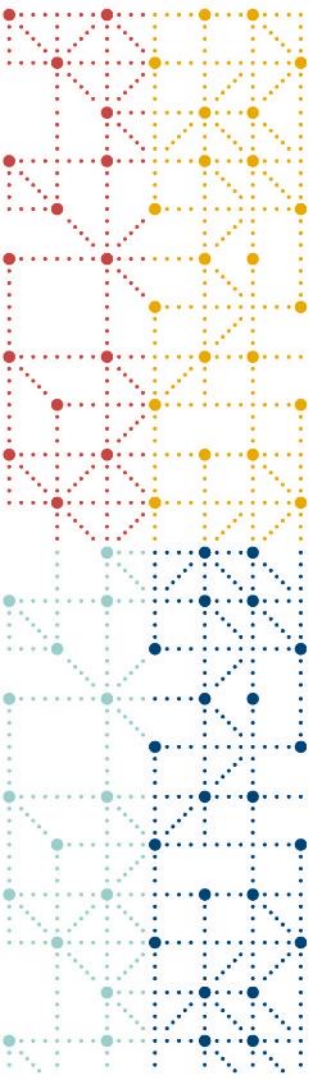
Disclaimer and Disclosures

- *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

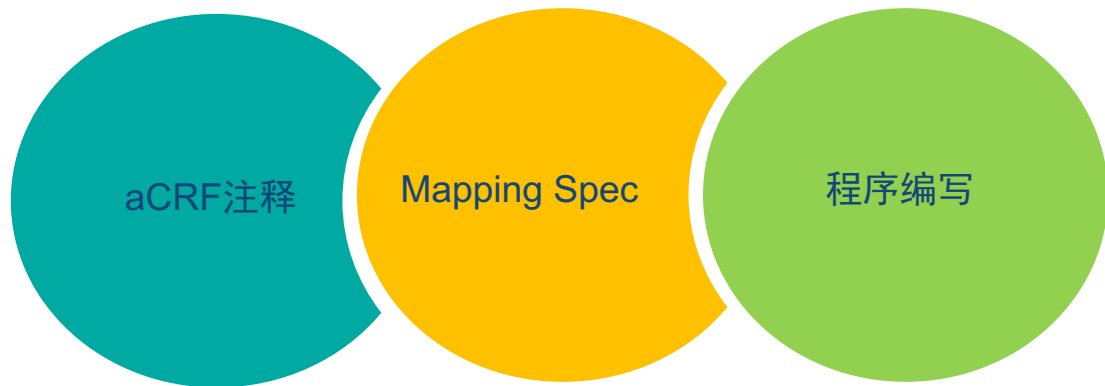
1. 背景
2. 构建基础资源库
3. 模块化工具
4. 多模块整合



背景

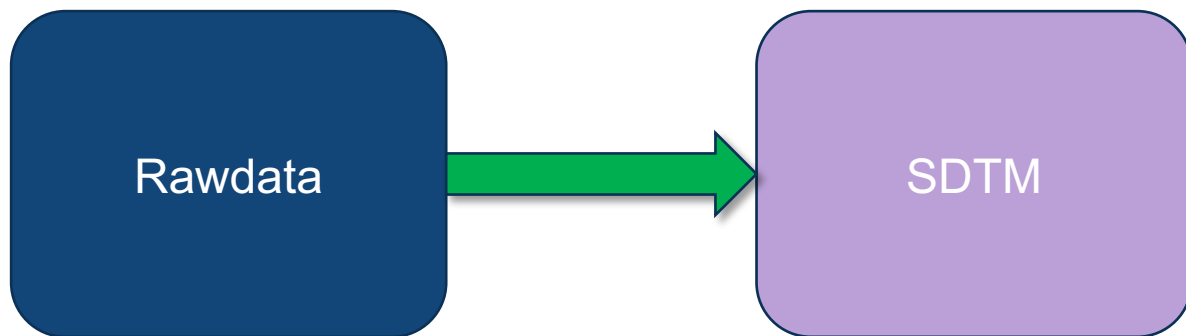
背景

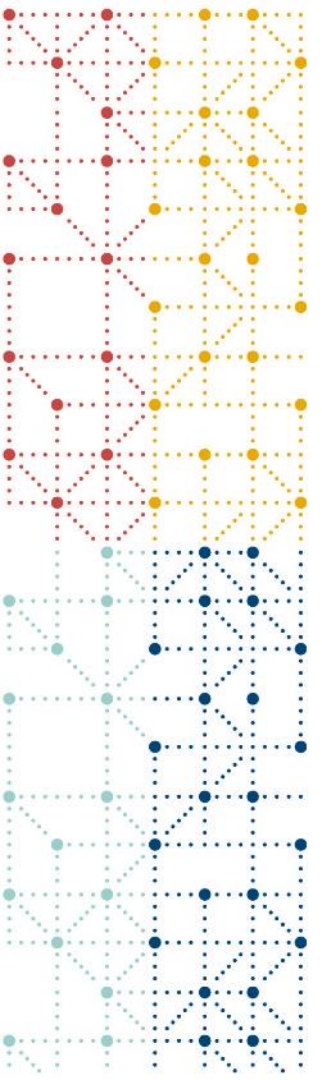
数据转换流程，主要是分为：aCRF注释、Mapping Spec、程序编写三个模块，减少人工编写操作，在模块中插入数据核查，优化这三部分能有效提高项目数据转换的速度与质量。



背景

目前越来越多的临床研究项目采用CDISC标准，将原始数据转换为SDTM数据集的过程，是一个标准且费时的的工作，将SDTM的转化使用工具标准化将会大大提高我们的效率与质量。





构建基础资源库

构建基础资源库

aCRF资源库主要用于自动注释aCRF，根据不同项目的CRF中的内容描述，不断完善资源库，提高aCRF注释的匹配度。

| 序列 | 字段 | 参数 |
|----|---------------|--------------------------------------------------|
| 1 | Key | 新增受试者、研究中心编号（只读） |
| 2 | EDC.system | TAIMEI、RAVE |
| 3 | Domain | DM（人口学）、肿瘤疾病史 |
| 4 | Language | Chinese、English |
| 5 | 页面纸张 | A4 |
| 6 | ➤ KeyPosition | (x,y) |
| 7 | ➤ ExpPosition | (x,y) |
| 8 | Type | Domain、Variable |
| 9 | ➤ Variable | SITEID、MHTERM |
| 10 | ExpBG | # FACD91# |
| 11 | ExpLineType | |
| 12 | ExpWord | (字型, 字号, 字体颜色, 底纹, 字体粗细(加粗), 斜体, 下划线, 字间距, 对齐方式) |
| 13 | 创建人 | |
| 14 | 创建日期 | |

构建基础资源库

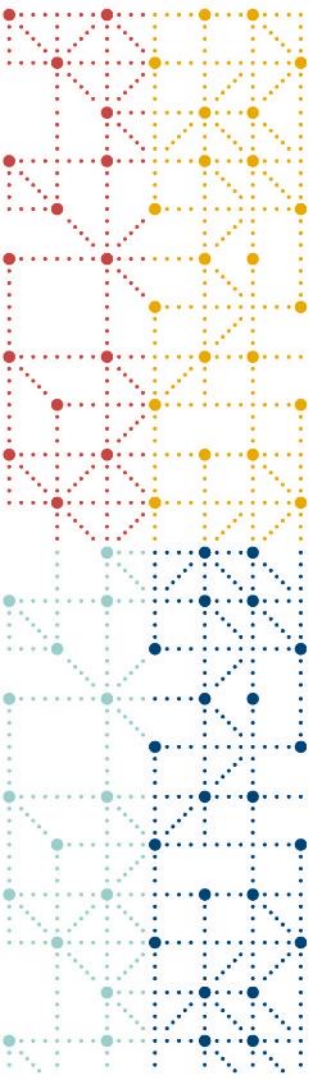
Mapping Spec中的Metadata资源库, 收集各个项目中可能出现的实验室、生命体征的指标值及其相关信息, 根据SDTM Terminology文件, 完善标准值, 能有效提高项目中如实验室模块等的效率, 以及保证因人理解不同造成的标准值不同的问题, 提高项目质量。

| | | | | | | | | | |
|---------------------------|----------|-------------------------|----------|--------|-------|--------|--------|----------|---------------|
| Domain | | Original Test Name | --TESTCD | --TEST | --CAT | --SCAT | --SPEC | --METHOD | Original Unit |
| Factor to Convention Unit | --ORRESU | Factor to Standard Unit | --STRESU | | | | | | |

构建基础资源库

QRS资源库也是作为Mapping Spec的主要部分，主要收集量表相关的信息，将量表的Code值做一个标准化。

| Source | VARNAME | Lvarlabel | ORG_CAT | ORG_SCAT | ORG_SCAT_RES | QSCAT | QSSCAT | QSEVLINT | QSEVINTX | QSTEST | QSTESTCD |
|--------|----------|-----------------|-----------------------------------|----------|--------------|------------------------|-------------------|----------|----------|------------------------------------------|----------|
| CSS3 | CSSACTYN | Whether Actual | C-SSRS-SUICIDAL BEHAVIOR-Lifetime | | | C-SSRS BASELINE/SCREEN | SUICIDAL BEHAVIOR | | LIFETIME | CSS04-Actual Attempt-Life | CSS0412A |
| CSS3 | CSSACTNT | Total number of | C-SSRS-SUICIDAL BEHAVIOR-Lifetime | | | C-SSRS BASELINE/SCREEN | SUICIDAL BEHAVIOR | | LIFETIME | CSS04-Number of Actual Attempts-Life | CSS0413A |
| CSS3 | CSSNSYN | Has subject enj | C-SSRS-SUICIDAL BEHAVIOR-Lifetime | | | C-SSRS BASELINE/SCREEN | SUICIDAL BEHAVIOR | | LIFETIME | CSS04-Non-suicidal Self-injur Behav-Life | CSS0414A |

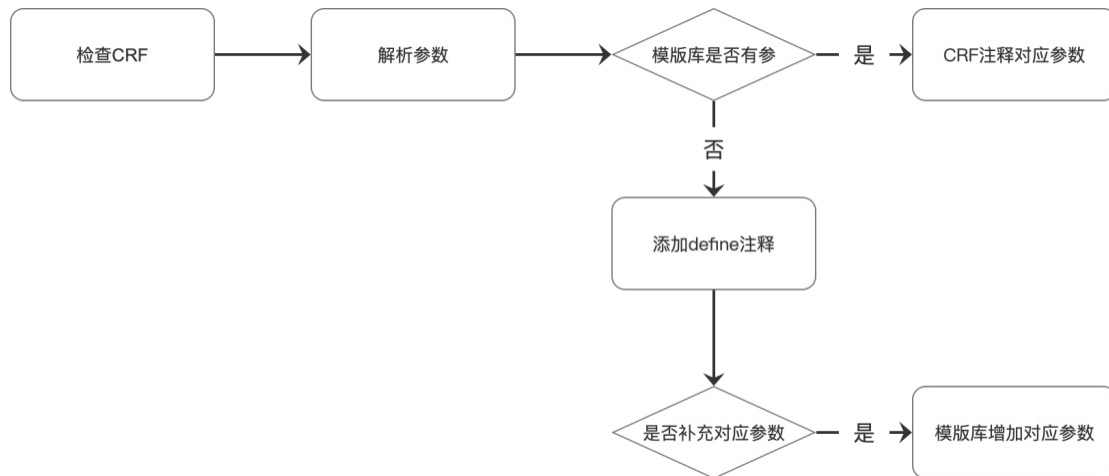


模块化工具

模块化工具

aCRF自动注释工具

CRF注释是转换SDTM数据集时比较费时的一步工作，自动注释可以极大提高效率以及准确性。



模块化工具

aCRF工具界面

The screenshot displays the aCRF-Tool interface. At the top, there are navigation tabs: "标准工具", "CRF-Tool", and "模板库", with "3232-ocrf" selected under "模板库". The user "admin" is logged in. Below the tabs are search filters for "Key", "创建时间" (Creation Time), and "Domain". There are buttons for "重置" (Reset) and "查询" (Search). A "展开" (Expand) dropdown is also present. Below the filters are three buttons: "批量删除" (Batch Delete), "批量导入" (Batch Import), and "创建模板" (Create Template). The main area contains a table with the following columns: Key, EDC-system, Domain, Language, Page Paper, KeyPos, VarPos, Type, Variable, ExpBG, ExpLine, ExpWord, Creator, Time, and 操作 (Action). The table lists several "新增受试者" (New Subject) entries with various parameters. At the bottom, there are pagination controls: "首页" (Home), "上一页" (Previous), "2", "3", "4", "5", "10", "下一页" (Next), "尾页" (End), "共 10 页 共 300 条记录" (Total 10 pages, 300 records).

| <input type="checkbox"/> | Key | EDC-system | Domain | Language | Page Paper | KeyPos | VarPos | Type | Variable | ExpBG | ExpLine | ExpWord | Creator | Time | 操作 |
|--------------------------|-------|------------|----------|----------|------------|-----------|-----------|--------|------------------|-------|----------------|---------|------------|------------------------------------------|----|
| <input type="checkbox"/> | 新增受试者 | TAIMEI | DM (人口学) | Chinese | A4 | (102,998) | (102,998) | Domain | # FACD91# | 虚线 | (xx,yy,zz,n,d) | 卢俊义 | 2023/03/03 | 编辑 删除 | |
| <input type="checkbox"/> | 新增受试者 | TAIMEI | DM (人口学) | Chinese | A4 | (102,998) | (102,998) | Domain | SITEID # FACD91# | 虚线 | (xx,yy,zz,n,d) | 卢俊义 | 2023/03/03 | 编辑 删除 | |
| <input type="checkbox"/> | 新增受试者 | TAIMEI | DM (人口学) | Chinese | A4 | (102,998) | (102,998) | Domain | # FACD91# | 虚线 | (xx,yy,zz,n,d) | 卢俊义 | 2023/03/03 | 编辑 删除 | |
| <input type="checkbox"/> | 新增受试者 | TAIMEI | DM (人口学) | Chinese | A4 | (102,998) | (102,998) | Domain | # FACD91# | 虚线 | (xx,yy,zz,n,d) | 卢俊义 | 2023/03/03 | 编辑 删除 | |
| <input type="checkbox"/> | 新增受试者 | TAIMEI | DM (人口学) | Chinese | A4 | (102,998) | (102,998) | Domain | # FACD91# | 虚线 | (xx,yy,zz,n,d) | 卢俊义 | 2023/03/03 | 编辑 删除 | |
| <input type="checkbox"/> | 新增受试者 | TAIMEI | DM (人口学) | Chinese | A4 | (102,998) | (102,998) | Domain | # FACD91# | 虚线 | (xx,yy,zz,n,d) | 卢俊义 | 2023/03/03 | 编辑 删除 | |
| <input type="checkbox"/> | 新增受试者 | TAIMEI | DM (人口学) | Chinese | A4 | (102,998) | (102,998) | Domain | # FACD91# | 虚线 | (xx,yy,zz,n,d) | 卢俊义 | 2023/03/03 | 编辑 删除 | |
| <input type="checkbox"/> | 新增受试者 | TAIMEI | DM (人口学) | Chinese | A4 | (102,998) | (102,998) | Domain | # FACD91# | 虚线 | (xx,yy,zz,n,d) | 卢俊义 | 2023/03/03 | 编辑 删除 | |
| <input type="checkbox"/> | 新增受试者 | TAIMEI | DM (人口学) | Chinese | A4 | (102,998) | (102,998) | Domain | # FACD91# | 虚线 | (xx,yy,zz,n,d) | 卢俊义 | 2023/03/03 | 保存 取消 | |



模块化工具

Mapping Spec工具

- 获取标准变量名，标签等
- 分模块上传Mapping Spec信息
- 检查变量信息是否合规
- 获取EDC变量名
- *生成模板SAS Code*



模块化工具













SDTM模板程序

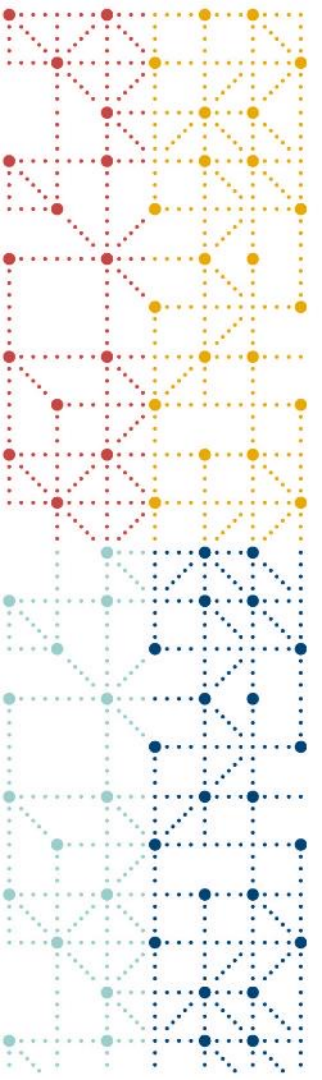
- 固定各个Domain的加工程序
- 标准化数据处理
- 数据点核查
- 逻辑优化，提高速率
- *匹配预处理变量加工程序*

模块化工具

SDTM模板程序

部分程序示例

-  m-check-cdisciso8601-v1-0
-  m-combine-supp-v1-0
-  m-create-cdisciso8601-v1-0
-  m-create-orresn-v1-0
-  m-create-relrec-v1-0
-  m-findds2var-v1-0
-  m-get-dslist-v1-0
-  m-get-epochcode-v1-0
-  m-get-labelname-v1-0
-  m-remove-specialchar-v1-0
-  m-sdtm-blfl-v1-0
-  m-sdtm-supp-v1-0



多模块整合

多模块整合

模块间整合





多模块整合

aCRF到Mapping Spec

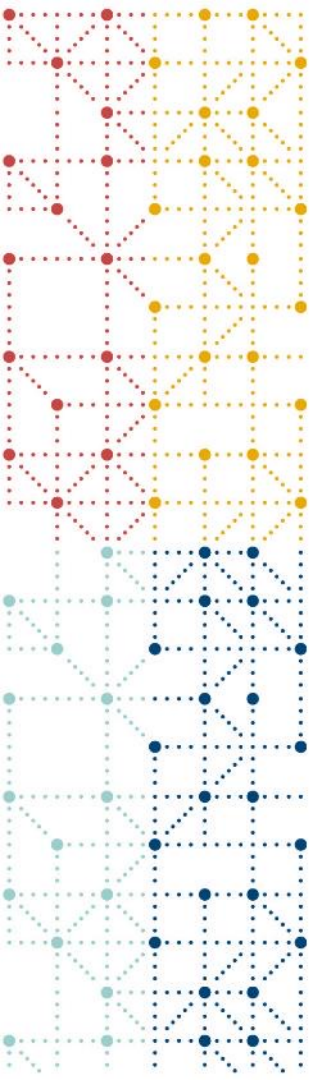
1. 获取注释信息；
2. 解析注释信息；
3. 提取信息填充mapping文件，包括Domain、Variables、SUPP相关信息等；
4. 使用ALS文件，结合aCRF资源库，获取EDC变量信息



多模块整合

Mapping Spec到SAS Code

- 使用宏直接创建模板程序；
- 根据Spec中的变量信息，可以固定变量名，标签，类型等；
- 将获取的EDC变量信息，填充至模板程序；
- 核验程序结构。



Thank You!

cdisc