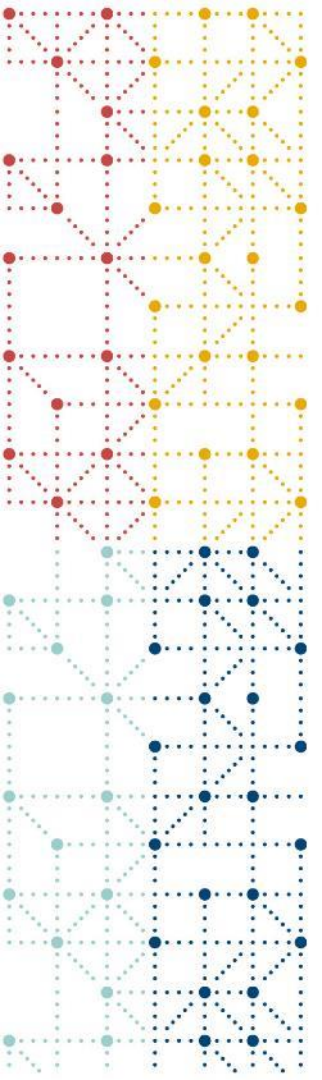


解读CDISC ARS分析结果标准

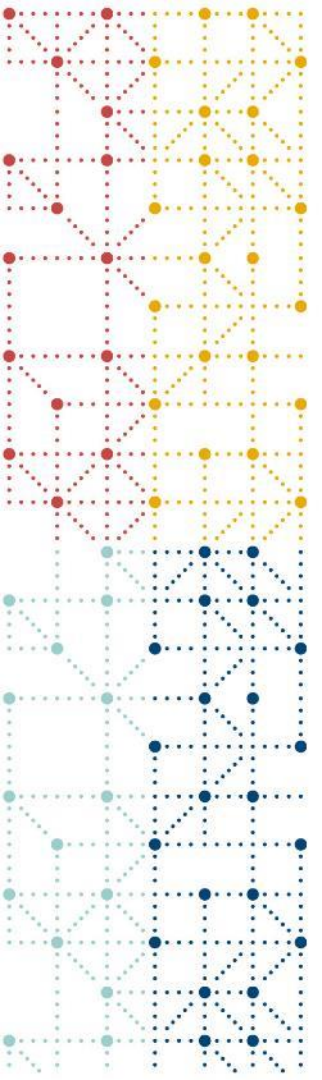
Yi Lu, Tigermed-MacroStat





Agenda

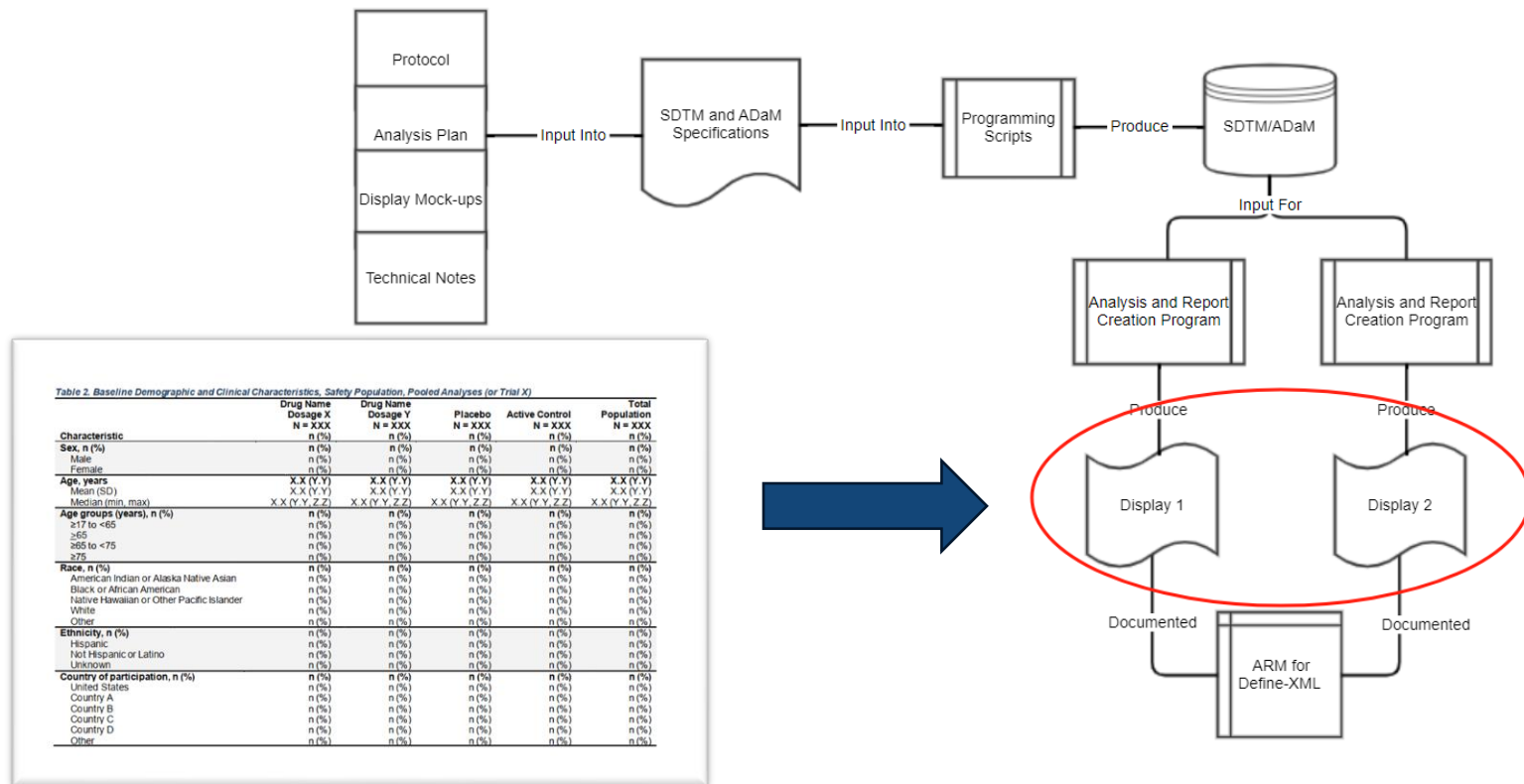
1. ARS模型介绍
2. ARS的核心组件
3. ARS应用场景
4. ARS面临的挑战



ARS模型介绍

- 背景和目的
- ARS的主要内容

当前得到分析结果的工作流程



当前工作流程存在的问题

- 昂贵，耗时
- 自动化实施困难
- 缺少可追溯性
- 可能存在分析结果报告之间的差异



需要一个标准的元数据框架指导自动化，存储追溯信息
需要一个标准的，可复用的分析数据集结构促进结果复用，减少差异，降低成本

ARS: Analysis Result Standard

用机器可读语言构建的分析结果标准的逻辑模型。

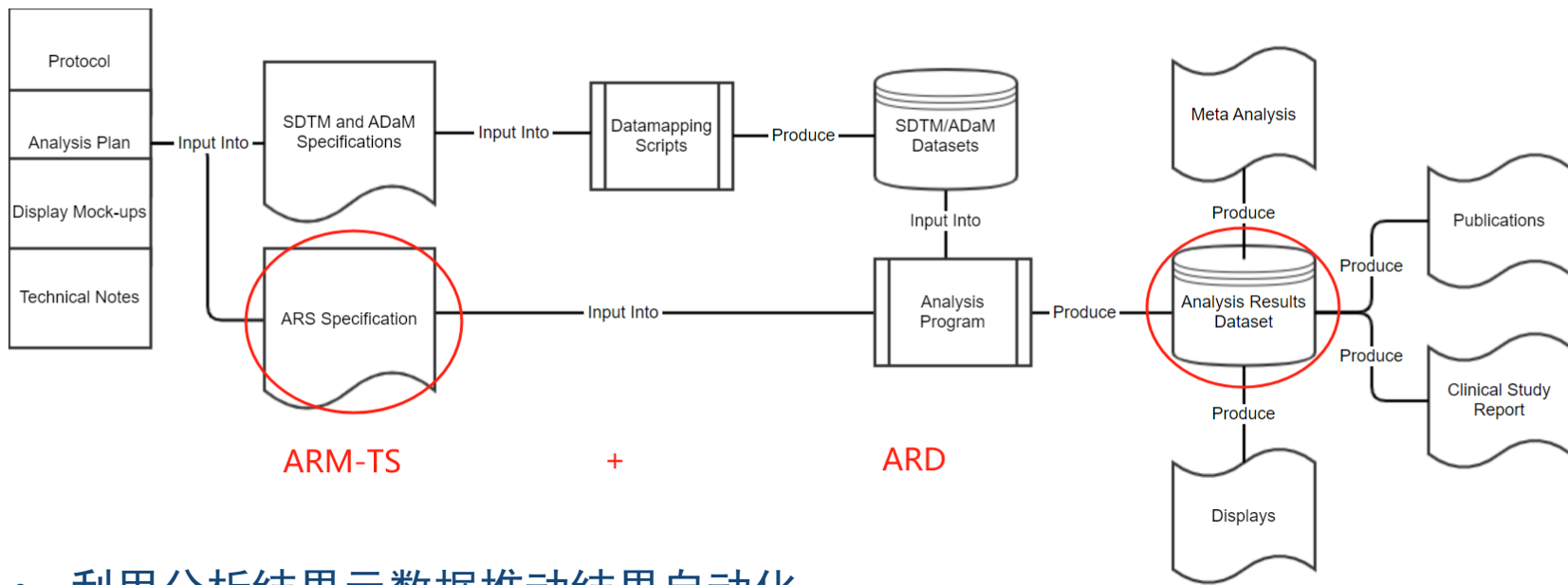
使用LinkML（Linked Data Modeling Language）作为基础，这是一种支持多种机器可读格式（如JSON、XML）的通用建模语言。分析结果元数据

LinkML中的两个核心概念是Class和Slot。Class代表模型中的组件，Slot则是这些组件的属性，比如ID、名称、描述等。

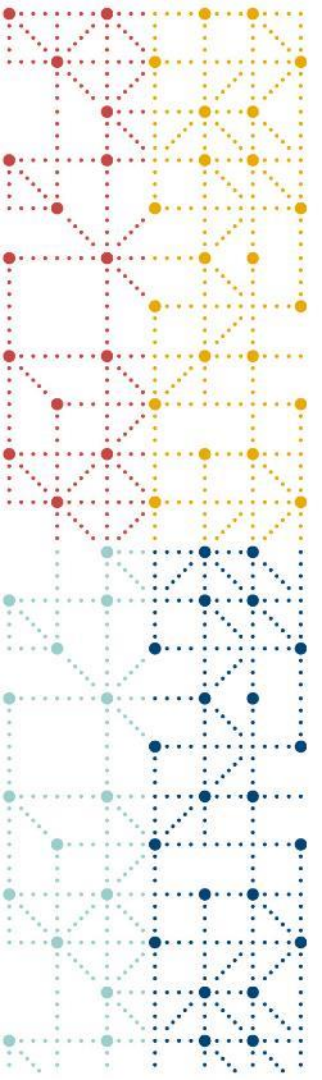
```
mainListOfContents: Class
  name: List of Planned Analyses
  label: LOPA Slot
  contentsList:
  > listItems: ...
  version: 1
  > otherListsOfContents: ...
  > referenceDocuments: ...
  > analysisSets: ...
  > analysisGroupings: ...
  methods:
  > - name: Count by group for a categorical variable...
  > - name: Summary by group of a categorical variable...
  > - name: Summary by group of a continuous variable...
  > analyses: ...
  > outputs: ... 多个属性值，使用'-'
```



引入ARS模型以后的工作流程



- 利用分析结果元数据推动结果自动化
- 支持结果的存储、访问、处理、可追溯性和可重现性



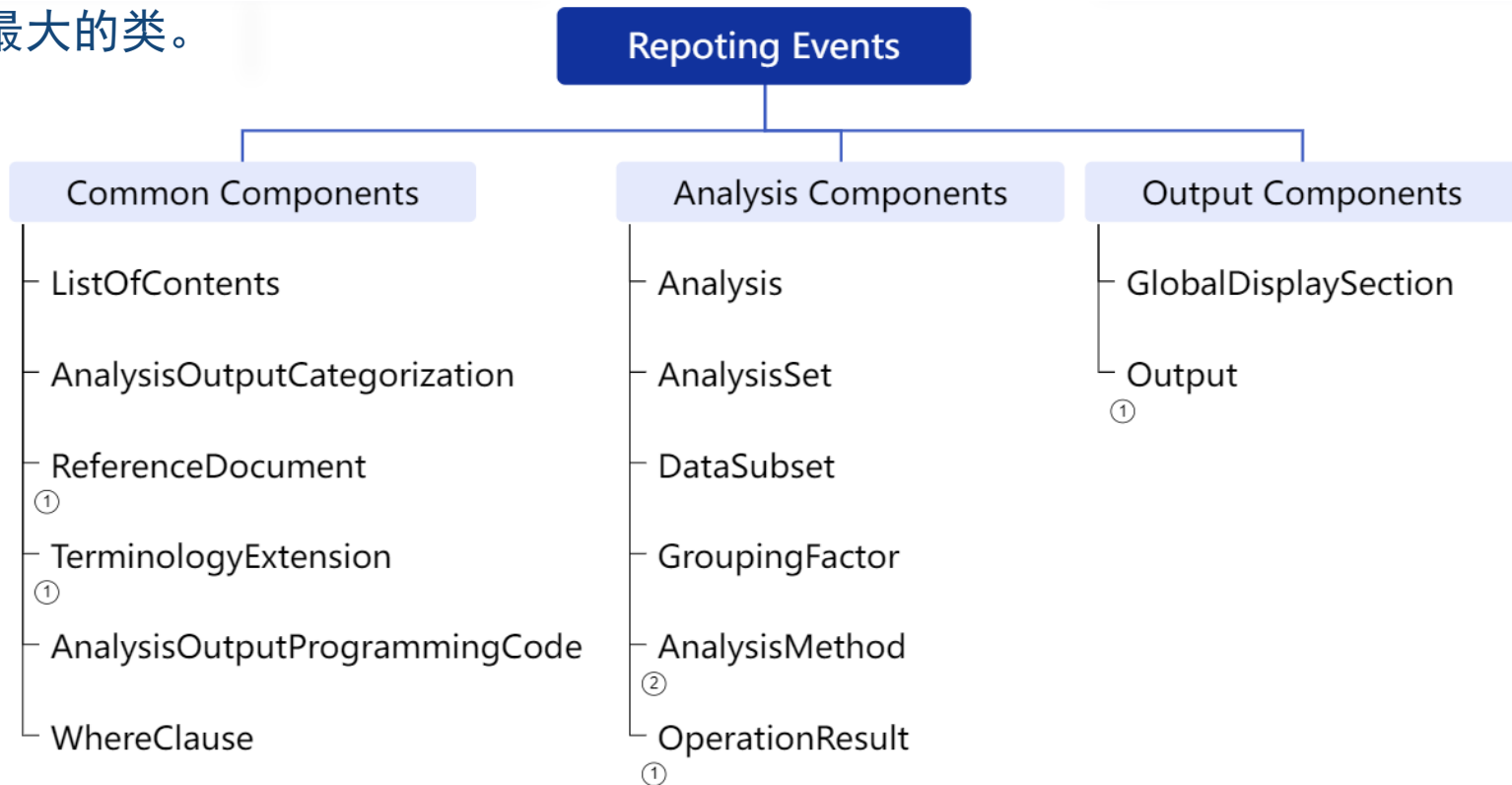
ARS的核心组件

Reporting Events

- Common Components
- Analysis Components
- Output Components

ARS的核心组件

Reporting Events: 定义一类特定分析需求，CSR或者中期分析。
ARS中最大的类。



ARS的核心组件

ListOfContents: 定义Reporting Events下的表名以及每张表输出的分析子列表

```
mainListOfContents:
  name: List of Planned Analyses
  label: LOPA
  contentsList:
    listItems:
      name: Table 2. Baseline Demographic and Clinical Characteristics, Safety Population,
        Trial CDISCPilot01
      level: 1
      order: 1
      outputId: O_T2
      sublist:
        listItems:
          - name: Summary of Subjects by Treatment...
          - name: Sex...
          - name: Age...
          - name: Age groups...
          - name: Race...
          - name: Ethnicity...
```

ARS的核心组件

DocumentReference: 用于创建报告事件中指定的任何参考文档的引用。
可以为以下任何内容创建文档引用:

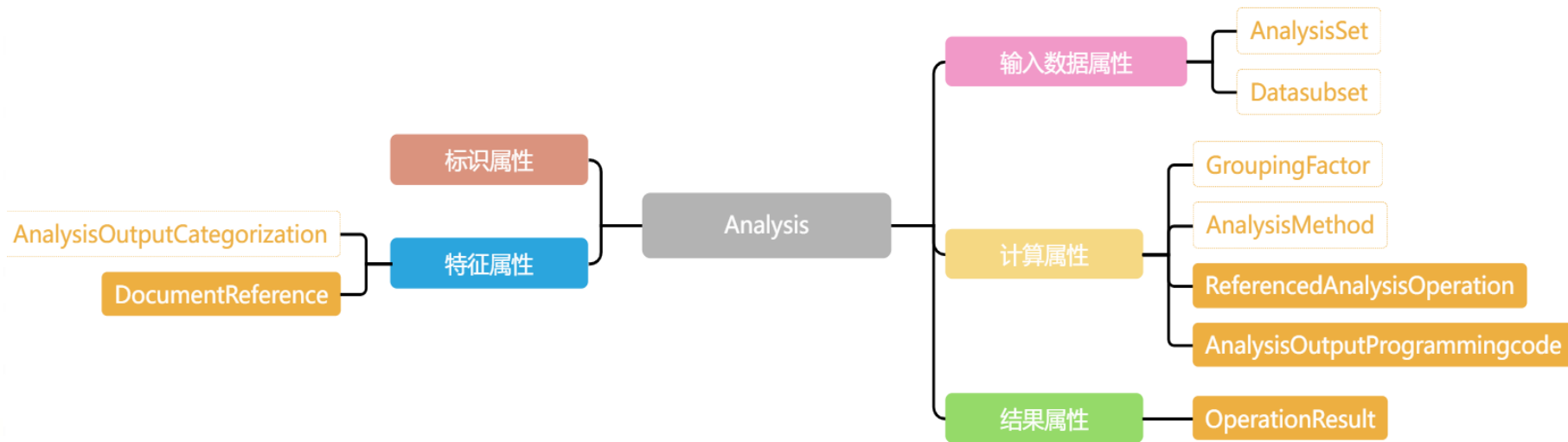
- 根据**SAP**和/或**方案**规定的分析文档
- **临床研究报告或综合总结**中的输出显示
- 包含特定分析方法、分析或输出的编程代码的**程序文件**

★该部分内容可提取 用于生成ARM for Define.XML

ARS DocumentReference Attributes	Corresponding Define-XML def:DocumentRef Attribute/Element
referenceDocumentId	leafID attribute
pageRefs	def:PDFPageRef elements (see below)

ARS的核心组件

Analysis: 用于定义和描述分析。



ARS的核心组件

AnalysisSet: 用于指定在方案/SAP中定义的分析集或分析人群。

Table 2. Baseline Demographic and Clinical Characteristics, Safety Population, Pooled Analyses (or Trial X)

Drug Name

Drug Name

Baseline Analysis

```
analysisSets:
```

```
- name: Safety Population
```

```
id: AS_SAF
```

```
level: 1
```

```
order: 1
```

```
condition:
```

```
dataset: ADSL
```

```
variable: SAFFL
```

```
comparator: EQ
```

```
value:
```

```
- Y
```

```
analyses:
```

```
- name: Summary of Subjects by Treatment
```

```
id: A_SAF_SUM_USUBJID_TRT
```

```
> reason: ...
```

```
> purpose: ...
```

```
methodId: M_GRP_CNT
```

```
version: 1
```

```
> documentRefs: ...
```

```
dataset: ADSL
```

```
variable: USUBJID
```

```
analysisSetId: AS_SAF
```

ARS的核心组件

DataSubset: 用于限制任何给定分析中包含的受试者或记录的附加数据子集，但未被视为定义的分析集。

比如：具有非缺失基线值的受试者，或在 AE 摘要分析中仅包含治疗中出现的不良事件

```
dataSubsets:  
- name: Treatment-Emergent Adverse Events  
  id: Dss01_TEAE  
  level: 1  
  order: 1  
  condition:  
    dataset: ADAE  
    variable: TRTEMFL  
    comparator: EQ  
    value:  
    - Y
```

```
analyses:  
- name: Summary of Subjects with At  
  > documentRefs: ...  
  > categoryIds: ...  
  dataset: ADAE  
  variable: USUBJID  
  analysisSetId: AnalysisSet_02_SAF  
  dataSubsetId: Dss01_TEAE
```

ARS的核心组件

AnalysisGroupings: 用于定义分组因素。

分组因素的示例包括受试者特征（例如性别或年龄组）、计划或实际治疗组、不良事件分析数据集中的“系统器官类别”或“首选术语”，以及生命体征分析数据集中的“参数”或“访视”。

```
analysisGroupings:
- name: Treatment
  id: AnlsGrouping_01_Trt
  dataDriven: false
  groupingDataset: ADSL
  groupingVariable: TRT01A
> groups: ...
- name: Gender
  id: AnlsGrouping_02_Sex
  dataDriven: false
  groupingDataset: ADSL
  groupingVariable: SEX
  groups:
> - name: Male ...
> - name: Female ...
```

```
analyses:
- name: Summary of Subjects by Treatment
  dataset: ADSL
  variable: USUBJID
  analysisSetId: AnalysisSet_02_SAF
  orderedGroupings:
  - order: 1
    groupingId: AnlsGrouping_01_Trt
    resultsByGroup: true
```

ARS的核心组件

AnalysisMethod: 用于报告事件中任何分析使用的分析方法。

```
methods:
- name: Summary by group of a continuous variable
  description: Descriptive summary statistics across groups for a continuous variable
  label: Grouped summary of continuous variable
  id: M_GRP_SUM_CONTIN
  operations:
    - name: Mean
      label: Mean
      id: M_GRP_SUM_CONTIN_1_MEAN
      order: 1
      resultPattern: X.X
    - name: Standard deviation
      label: SD
      id: M_GRP_SUM_CONTIN_2_SD
      order: 2
      resultPattern: (Y.Y)
```

```
analyses:
- name: Summary of Age by Treatment
  id: A_SAF_SUM_AGE_TRT
  reason:
    controlledTerm: SPECIFIED IN SAP
  purpose:
    controlledTerm: PRIMARY OUTCOME MEASURE
  methodId: M_GRP_SUM_CONTIN
  version: 1
```

ARS的核心组件

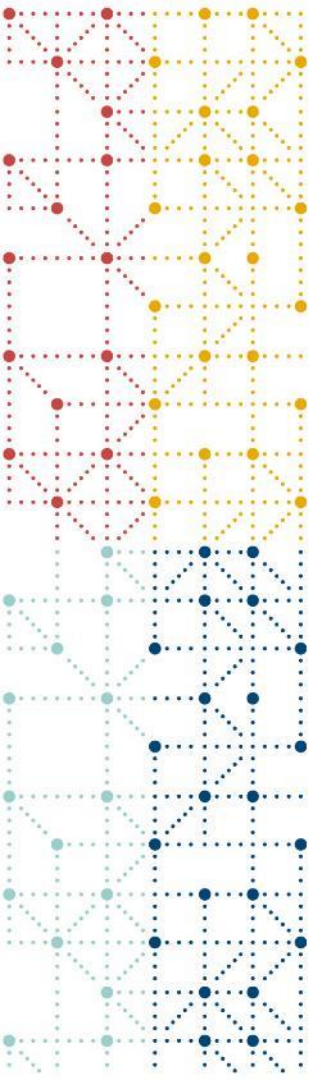
OperationResult: 分析的结果。也就是分析结果数据集ARD。

```
    "results": [  
      {  
        "resultGroups": [  
          {  
            "groupingId": "AnlsGrouping_01_Trt01An",  
            "groupId": "AnlsGrouping_01_Trt01An_01"  
          }  
        ],  
        "rawValue": "",  
        "formattedValue": "(N=)"  
      },  
      {  
        "resultGroups": [  
          {  
            "groupingId": "AnlsGrouping_01_Trt01An",  
            "groupId": "AnlsGrouping_01_Trt01An_02"  
          }  
        ],  
        "rawValue": "",  
        "formattedValue": "(N=)"  
      }  
    ]  
  }
```

ARS的核心组件

Output: 用于指定表格显示和输出的相关内容。
比如标题，脚注，输出格式，输出文件名等等。

```
fileSpecifications:
  - name: t2-demog (RTF)
    label: t2-demog
    fileType:
      controlledTerm: rtf
    location: ./t2-demog.rtf
  - name: t2-demog (PDF)
    label: t2-demog
    fileType:
      controlledTerm: pdf
    location: ./t2-demog.pdf
  - orderedSubSections: ...
    - sectionType: Footer
    - orderedSubSections:
```



ARS应用场景

- eTFL Portal Package
- 基于ARS的自动化工具
 - TFL Output
 - ADaM
 - ARM for Define.XML

- eTFL Portal Package

Table Of ContentsSummary of Analysis Popul... X

Summary of Analysis Populations

Safety Population

Study Test

Analysis Population, n (%)	Placebo (N=XX)	Xanomeline High Dose (N=XX)
Enrolled	XX (XX.XX)	XX (XX.XX)
Safety	XX (XX.XX)	XX (XX.XX)
Intent-To-Treat (ITT)	XX (XX.XX)	XX (XX.XX)
Full Analysis Set	XX (XX.XX)	XX (XX.XX)
Per-Protocol	XX (XX.XX)	XX (XX.XX)
Enter Label		

PropertiesOutline

Analysis DatasetADSL

Analysis CategorySafety

Analysis Of InterestDisposition

> Data Subset for TableConfigure

Split Statistics Column

Population SetSafety Population

Population DatasetADSL

Where Clause for Safety Population
(SAFFL EQ Y)

- 基于ARS的自动化工具 <生成ADaM>

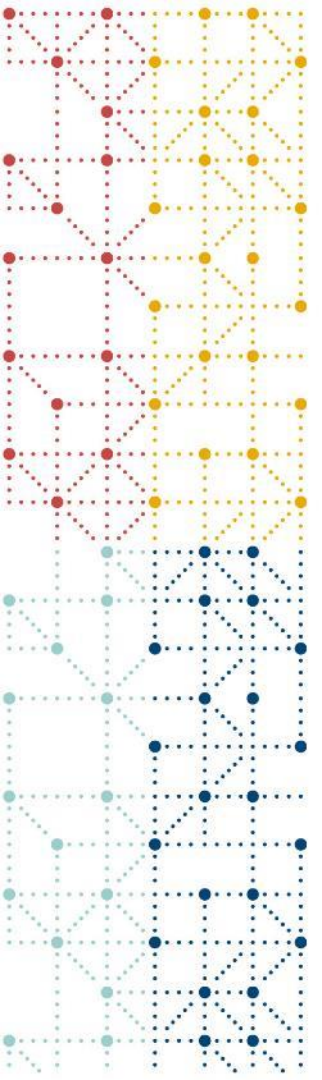


```
{'dataset': 'ADSL', 'variable': 'SAFFL', 'values': ['Y']}  
{'dataset': 'ADSL', 'variable': 'TRT01A', 'values': ['Placebo']}  
{'dataset': 'ADSL', 'variable': 'TRT01A', 'values': ['Xanomeline Low Dose']}  
{'dataset': 'ADSL', 'variable': 'TRT01A', 'values': ['Xanomeline High Dose']}  
{'dataset': 'ADSL', 'variable': 'SEX', 'values': ['M']}  
{'dataset': 'ADSL', 'variable': 'SEX', 'values': ['F']}  
{'dataset': 'ADSL', 'variable': 'AGEGR2', 'values': ['<65']}  
{'dataset': 'ADSL', 'variable': 'AGEGR2', 'values': ['65+']}  
{'dataset': 'ADSL', 'variable': 'AGEGR3', 'values': ['17-<65']}  
{'dataset': 'ADSL', 'variable': 'AGEGR3', 'values': ['65-<75']}  
{'dataset': 'ADSL', 'variable': 'AGEGR3', 'values': ['75+']}
```

• 基于ARS的自动化工具 <生成ARM>



	ARS	ARM
目的和用途	前瞻性地利用分析结果，以实现自动化、可重用性和可追溯性。	为了满足监管需求而开发的，用于向监管机构递交，以提供特定分析结果到用于生成该结果的特定ADaM数据的可追溯性。ARM通常仅针对关键分析进行回顾性创建。
创建过程	YAML数据建模，同时添加了更多组件，可以用于支持创建ARM，	定义XML标准而创建的，包含对分析目的、每个分析的原因和目的的描述
数据准备要求	假设输入数据集已经是“分析就绪的”（analysis-ready），不需要复杂的数据处理，重点描述如何从这些数据集生成结果。	假设输入数据是经过处理、符合标准的分析数据，通常是从准备好的分析数据集生成。



ARS面临的挑战

- 模型本身较为复杂
- 环境：
 - 现有技术支持不足
 - 对监管机构优势不明显

ARS面临的挑战

? 模型本身:

- 模型细节多，组件较为复杂。存在一定的学习曲线。推广和普及难度高。现有的用户手册更偏向于提供给开源工具开发者使用。

★ 增强文档支持:

区分不同使用角色。也可以考虑提供不同层次的培训资源，从初学者到高级用户不等。

ARS面临的挑战

? 模型本身:

- 模型细节多，组件较为复杂。存在一定的学习曲线。推广和普及难度高。现有的用户手册更偏向于提供给开源工具开发者使用。

★ 分阶段，分步骤实施&推广:

根据市场需求，各公司现有资源利用ARS开发一些自动化工具。

- 1.Word/Excel Shell -> 线上TFL库 (eTFL)
- 2.TFL自动化工具 (自动化生成ARD;根据ARD自动化生成不同显示)
- 3.其他配套的自动化小工具

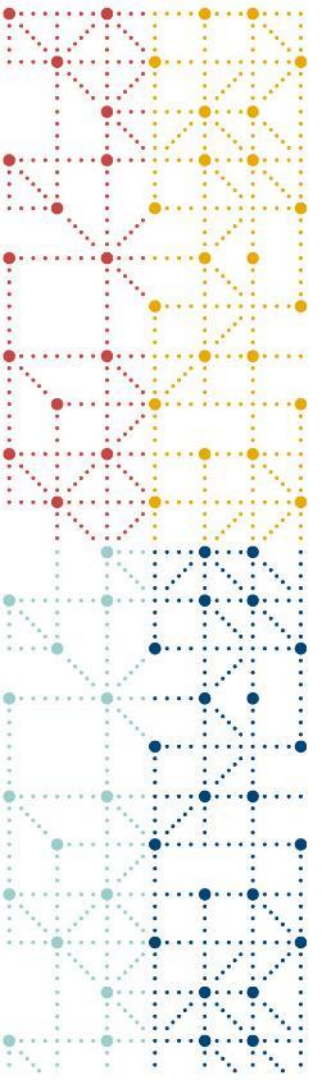
ARS面临的挑战

? 环境:

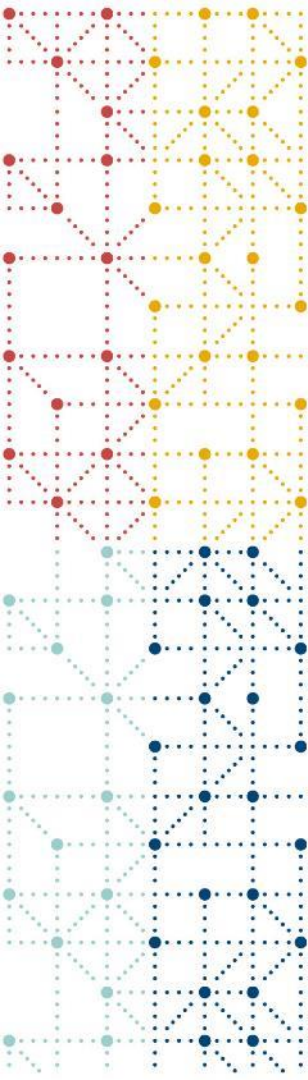
- 缺乏成熟的技术工具和支持平台;
- 模型的优势主要体现在使用者层面, 对监管机构价值和推动作用不突出

★ 加大对ARS的宣传和展示:

通过更多的沟通和合作, 展示ARS标准在提升数据质量、减少重复工作、加速药物研发等方面的价值, 获得多方认可。从而促进开源工具开发, 进一步争取监管机构的支持。

A decorative vertical bar on the left side of the slide, featuring a grid of dots in red, yellow, and blue, connected by thin lines to form a complex geometric pattern.

Q&A



Thank You!

参考资料:

1.<[Analysis Results Standard User Guide v1.0](#)>

2.CDISC EU 2024 <[Getting Started with the New CDISC Analysis Results Standard](#)> Bess
LeRoy/Richard Marshall

