

## 基于SDTM/ADaM 的跨学科可视化监查平台

王治东/统计总监 北京信立达医药科技有限公司

2025.08.29



王治东

统计总监

北京信立达医药科技有限公司

硕士研究生毕业于空军军医大学（原第四军医大学），流行病与卫生统计学专业，研究方向为临床试验设计与统计分析方法，先后在医院、药企、CRO从事生物统计师工作，参与80多项临床试验设计与统计分析工作，涵盖治疗领域：抗感染I、II期，消化系统I、II、III期，内分泌系统I、II、III、IV期，肿瘤I、II、III期，免疫系统I、II、III期等。

# 目录

- 解构现状 – 数据孤岛
- 技术基石 – 元数据驱动
- 价值绽放 – 可视化层
- 总结与展望

## 解构现状—数据孤岛

- **当前痛点：**数据分散在EDC不同页面，或者不同系统和文档中，形成信息孤岛。
- **深层瓶颈：**统计流程（DVP、SAP、TFL Shells、Cut Off Plan、Patient Profile Shell、SDTM/ADaM Specs、SDTM、ADaM、TFL.....），在Timeline极限压缩下存在大量人工映射与转换，导致效率低下、错误率高、可追溯性差。
- **核心问题：**数据分散不标准、数据宏观整合效率低、数据时效性差、关联漏查率高、缺乏一个机器可读可执行的统一语义层。

“语义层”技术通过对数据库里的有关数据项定义，把数据库中的数据定义成有明确的业务含义的名称。业务人员所面对的不再是表、字段和它们之间复杂的关联、计算关系，而是他所熟悉的业务术语和指标名称。

假如

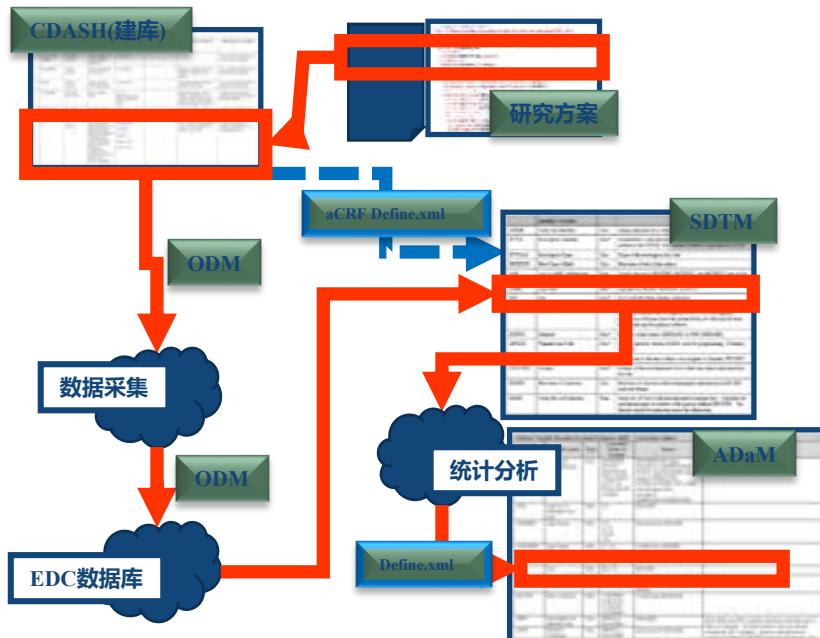
- EDC数据“及时”进行了SDTM/ADaM/TFL的转换
- TFL能以可视化的方式及时展现
- 数据能穿透到个体数据
- SAP&TFL Shell、Patient Profile Shell能和SDTM/ADaM/TFL数据相穿透
- .....

将有效解决上述痛点

# 解决方案？

# 技术基石：元数据驱动—从“文档”到“可执行代码”的跃迁

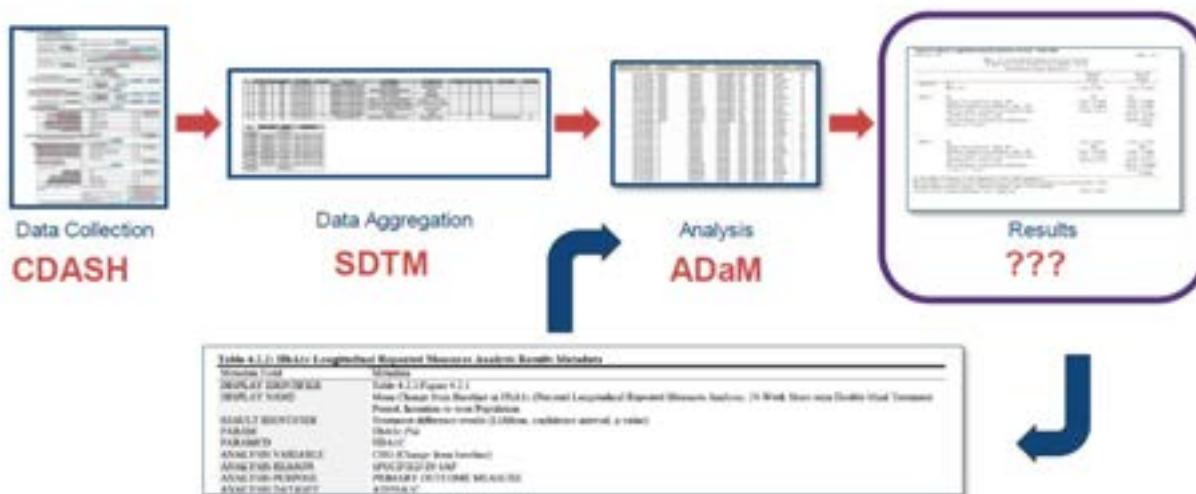
## 架构核心：以元数据为轴的自动化数据流水线



端到端元数据驱动架构图

# 技术基石：元数据驱动—从“文档”到“可执行代码”的跃迁

根据这样的思路： CDISC Foundational Standards



# 技术基石：元数据驱动—从“文档”到“可执行代码”的跃迁

**输入层:** Protocol 、 DVP 、 SAP 、 TFL Shells 、 Cut Off Plan 、 Patient Profile Shell 、 SDTM/ADaM Specs — 所有这些不再是Word/PDF文档，而是结构化的、可被机器解析的元数据。



Mockup Shell

- 表格模板

Mockup Shell

- Mockup的主要元素
  - 流程标题
  - 流程内道
  - 流程外道
  - 表格顺序
  - 表格属性

Specification

- Spec元素
  - 表格属性
  - 表格限制

Tracker

- Programming Tracker
  - 任务类型、任务名称、输出结果、标题备注、关键变量等
  - 责任人、比对过程IssueLog

Code

- Output Metrics
  - 程序-标识符
  - 程序-Programming Code, 比对路径

QC Report

- QC Report
  - 程序日志
  - 比对结果
  - 时间逻辑

#ClearDataClearImpact

cdisc

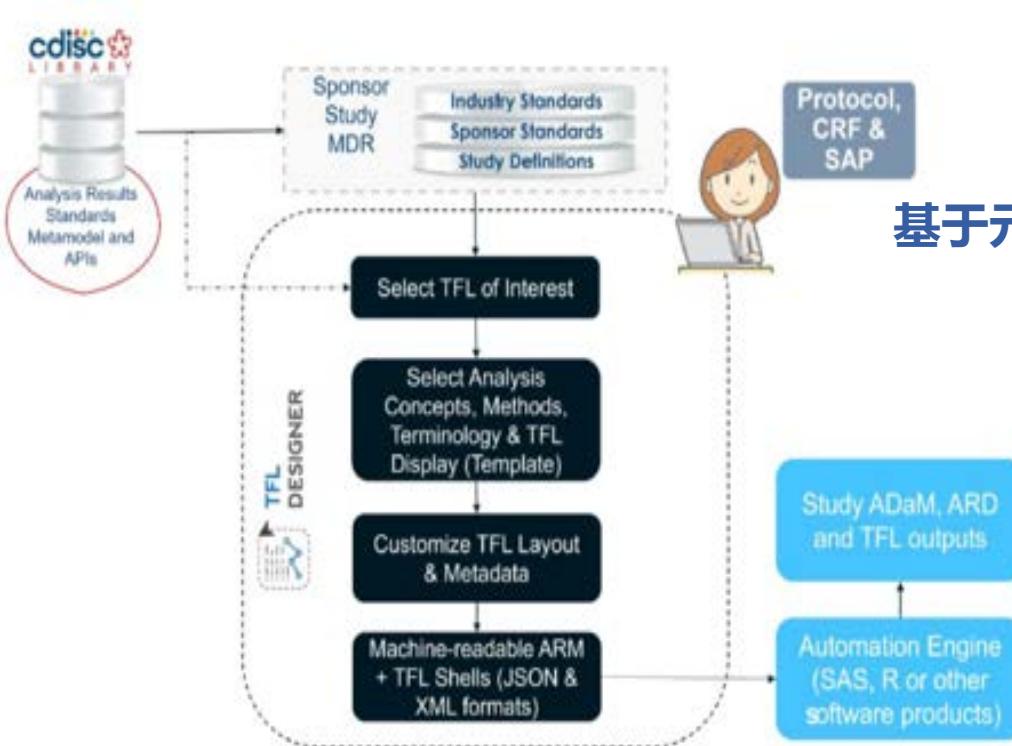
# 技术基石：元数据驱动—从“文档”到“可执行代码”的跃迁

**引擎层:**自动化引擎展示上述元数据

**输出层:**标准化的数据集、表格、清单、以及最重要完整的可追溯的流程元数据，自动生成：

- SDTM/ADaM数据集及其程序(取代手工编程)
- TFL及其程序 (基于TFL Shell)
- Patient Profile
- AE Narrative
- DSUR 图表
- .....

# 技术基石：元数据驱动—从“文档”到“可执行代码”的跃迁

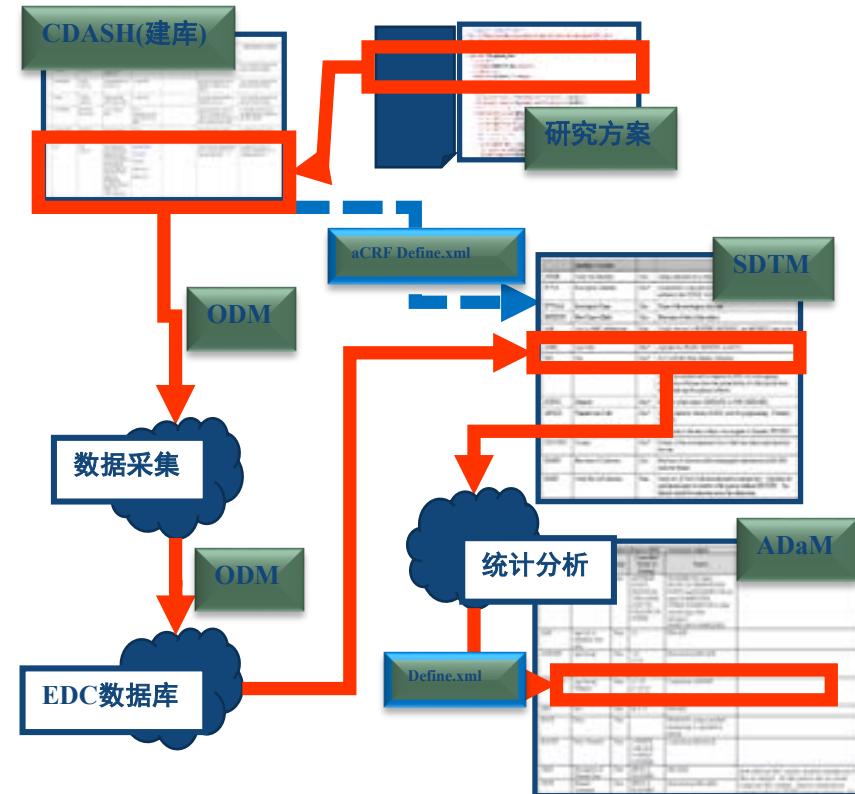


基于元数据的TFL自动化生成

Trust Medlens

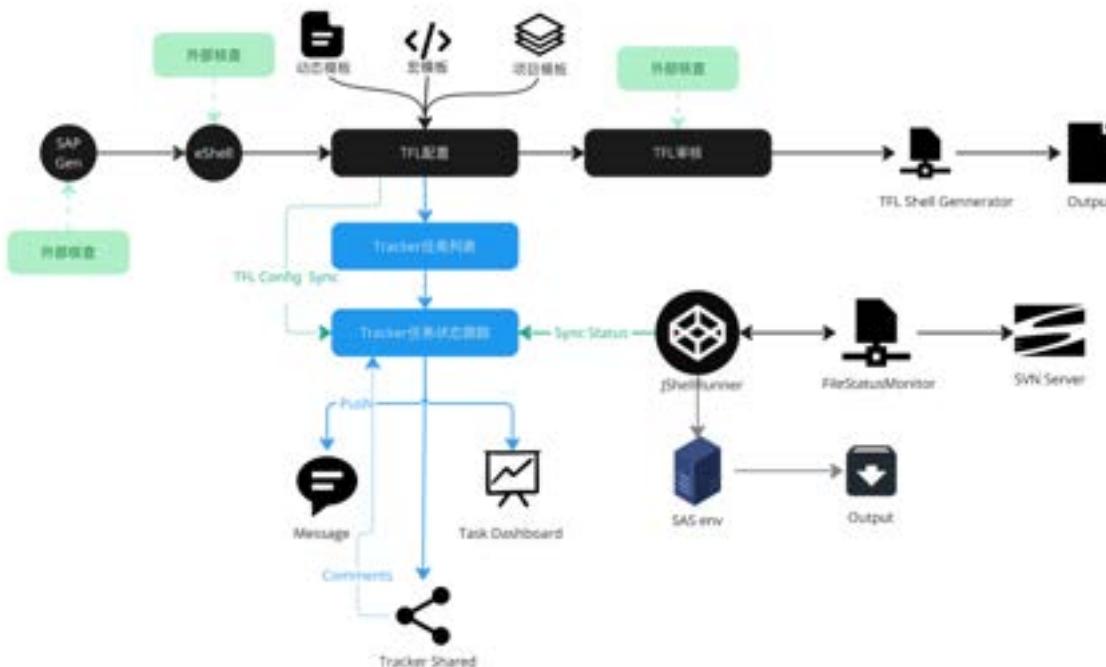
# 技术基石：元数据驱动—从“文档”到“可执行代码”的跃迁

- 结构化文档创建
  - DMP/DVP/SAP/TLF Shell...
- 智能Timeline管理下在线文档审阅
  - CRF/DVP/SAP/SDTM /ADaM ...
- 可视化报告流程
  - 数据清单
  - SDTM/ADaM数据集
  - TFL/动态图
  - 内外部、多部门协同反馈



# 技术基石：元数据驱动—从“文档”到“可执行代码”的跃迁

部分技术路径展示图



# 价值绽放 – 可视化层

## 交互式在线QC

The screenshot displays a medical coding report interface with the following details:

**Medical Coding Report**  
Study Number: 1234567890  
Study Title: Study A  
Sponsor Name: Company A  
Sponsor Company: MedData LLC  
Software Version: WinCoding Global Sep. 2022  
Audit Log:

操作时间	操作员	操作描述
2023-01-01 10:00:00	Admin	登录系统
2023-01-01 10:05:00	John Doe	开始审核
2023-01-01 10:10:00	John Doe	完成审核

**医学编码报告默认设置**  
版本日期: <File Version> | <File Date>

各方代表	姓名/职位/公司	签字	日期
审核人	<Code reviewer> <EMR Title> <EMR Company>		
编码申请人	<Code applicant> <EMR Title> <EMR Company>		

# 价值绽放 – 可视化层

## 交互式在线QC

The screenshot shows a software application window titled "交互式在线QC" (Interactive Online QC). The main area is a grid-based data viewer with numerous rows and columns of data. The columns include headers such as "ID", "名称" (Name), "值" (Value), "状态" (Status), and "操作" (Operation). A vertical toolbar on the left side contains icons for "新建" (New), "编辑" (Edit), "删除" (Delete), "导出" (Export), and "刷新" (Refresh). On the right side, there is a sidebar with sections for "待办事项" (Pending Tasks) and "日志" (Log). The "待办事项" section lists several items, including "H12: 需要将AD00011 + '→AD001'，到AD00011 + '→AD001'，到A..." and "H13: 需要将AD00011 + '→AD001'，到AD00011 + '→AD001'，到A...". The "日志" section shows a log entry: "完成" (Completed) at 16:00 11月 12日.

# 价值绽放 – 可视化层

## 交互式在线QC

The screenshot shows a web-based Quality Control (QC) application interface. On the left, there is a vertical navigation menu with the following items:

- 目录
- 变量页
- 校正历史
- 报告
- 数据源
- 1. 背景
- 2. 研究目的
- 3. 试验设计
- 4. 研究终点
- 5. 统计分析模型
- 6. 统计学考虑

  - 6.1 一般建议
  - 6.2 数据分析
  - 6.3 中心作用
  - 6.4 偏倚校正
  - 6.5 数据管理

    - 6.5.1 数据质量的保证
    - 6.5.2 数据报告
    - 6.5.3 数据审核和验证
    - 6.5.4 分析结果解释

  - 6.6 相关指南
  - 7. 统计分析方法
  - 8. 相关与回顾性计划不同的分析
  - 9. 表、图表和表格
  - 10. 参考文献
  - 11. 别册

#ClearDataClearImpact

# 价值绽放 – 可视化层

## 交互式在线QC

The screenshot displays a user interface for an 'Interactive Online QC' application. At the top, there's a navigation bar with tabs like '首页', '录入', '查询', '导出', '帮助', and '退出'. Below the navigation is a toolbar with various icons for file operations. The main content area is titled '14.1.1.1.1 调整随访日期' and '14.1.1.1.1.1 评估是否需要启动日期'.

The interface includes a table with columns: '项目' (Project), '状态' (Status), '备注说明' (Remarks), and '修改' (Modify). The table lists several items, each with a status indicator (green, yellow, red) and a timestamp. A '筛选' (Filter) button is located at the bottom left of the table.

To the right of the table, there's a '评价 (20)' (Evaluation) section with a checkbox for '通过审核' (Passed Review) and a link '编辑详细' (Edit Details). Below this are three small preview boxes:

- Top-left: '访视开始时间的合理性' (Reasonableness of visit start time) with a timestamp '2023-07-10 10:00:00'.
- Top-right: '访视结束时间的合理性' (Reasonableness of visit end time) with a timestamp '2023-07-10 10:00:00'.
- Bottom: '访视时长合理性' (Reasonableness of visit duration) with a timestamp '2023-07-10 10:00:00'.

At the bottom left, there's a 'cdisc' logo, and at the bottom center, the text '#ClearDataClearImpact'.

# 价值绽放 – 可视化层

## 交互式在线QC

The screenshot displays an interactive online Quality Control (QC) application interface. The top navigation bar includes tabs for '首页' (Home), '我的' (My), '数据' (Data), '报告' (Report), and '帮助' (Help). A search bar is located at the top right.

The main content area contains two tables of data:

**数据 1 (Data 1)**

项目	1月	2月	3月
总值	123.45 (1.23%)	123.45 (1.23%)	123.45 (1.23%)
平均数	123.45	123.45	123.45
标准差	123.45	123.45	123.45
均值 (标准差)	123.45 (1.23%)	123.45 (1.23%)	123.45 (1.23%)
最小值	123.45	123.45	123.45
Q1-Q3	123.45, 123.45	123.45, 123.45	123.45, 123.45
最小值, 最大值	123.45, 123.45	123.45, 123.45	123.45, 123.45

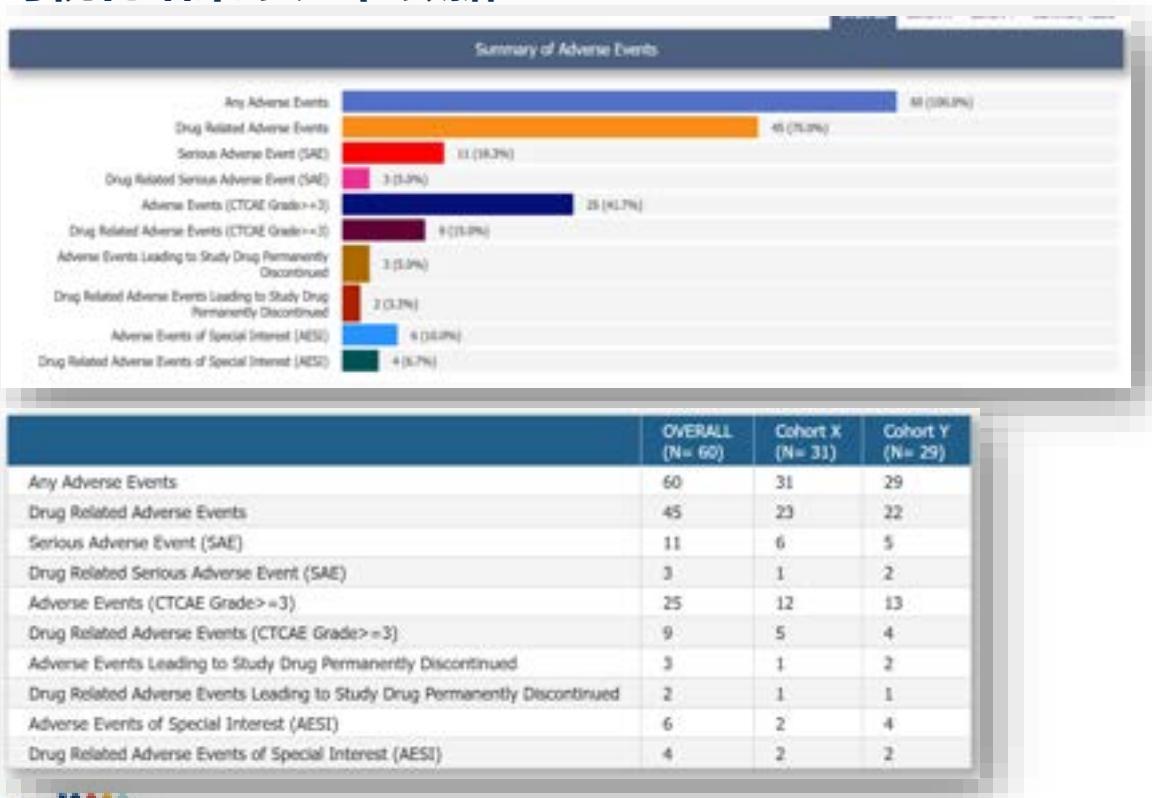
**数据 2 (Data 2)**

项目	1月	2月	3月
总值	123.45 (1.23%)	123.45 (1.23%)	123.45 (1.23%)
平均数	123.45	123.45	123.45
标准差	123.45	123.45	123.45
均值 (标准差)	123.45 (1.23%)	123.45 (1.23%)	123.45 (1.23%)
最小值	123.45	123.45	123.45
Q1-Q3	123.45, 123.45	123.45, 123.45	123.45, 123.45
最小值, 最大值	123.45, 123.45	123.45, 123.45	123.45, 123.45

On the right side of the interface, there is a sidebar with several buttons and a status bar indicating '正在运行...' (Running...) and the date '2024-01-01'.

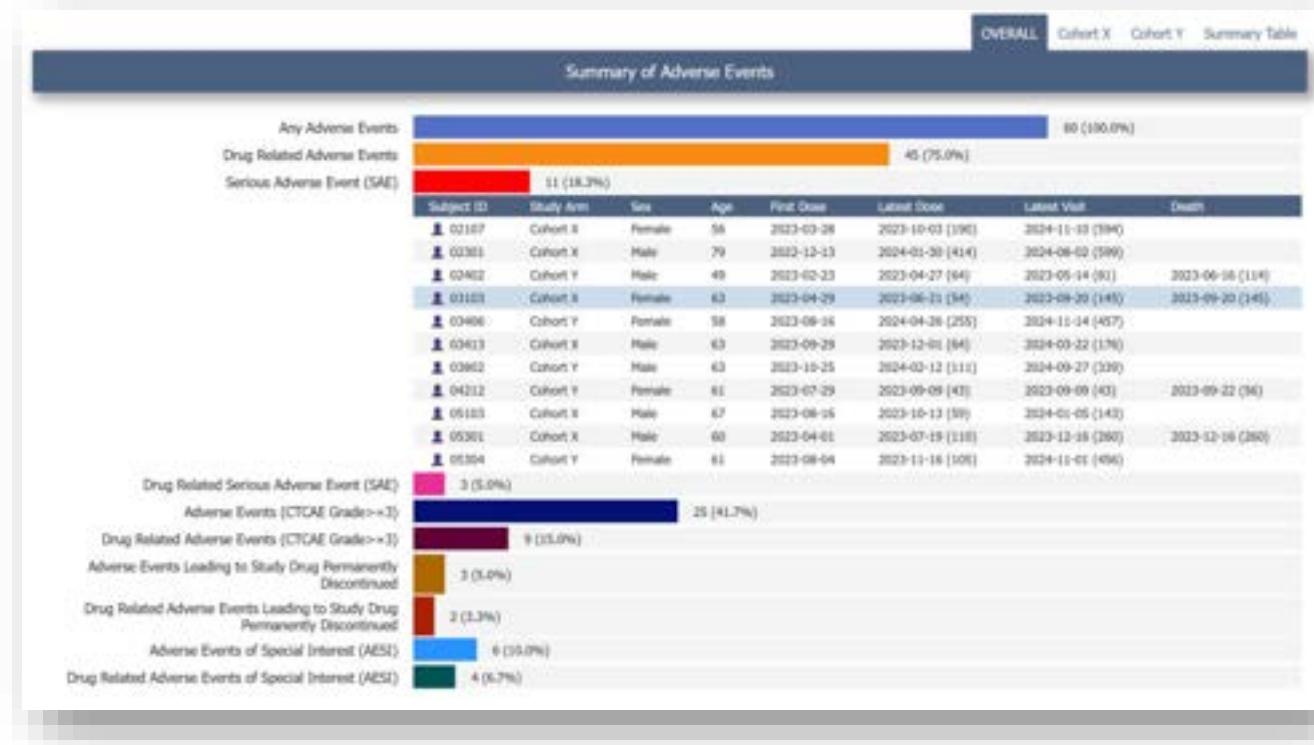
# 价值绽放 – 可视化层

## 可视化结果-安全性数据



# 价值绽放 – 可视化层

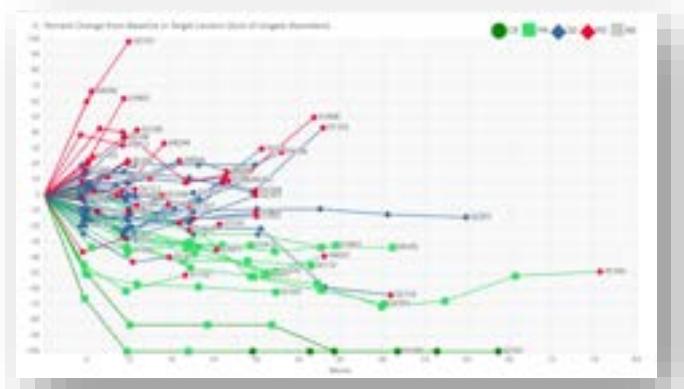
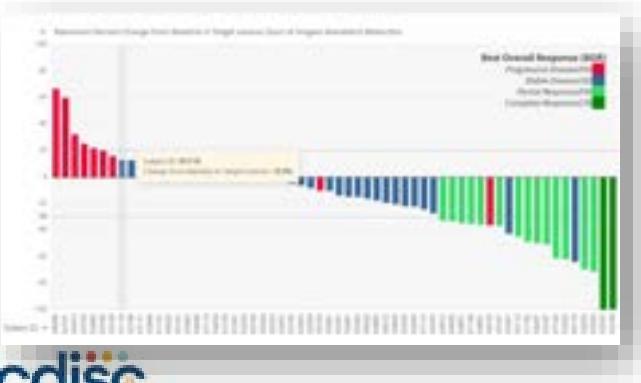
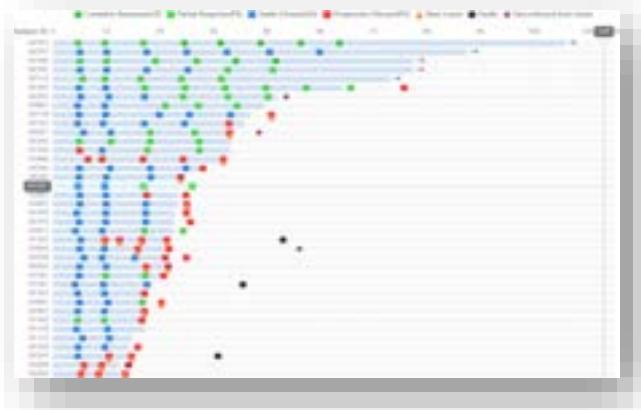
## 可视化结果-安全性数据



穿透：  
- 受试者列表 (Listing)

# 价值绽放 – 可视化层

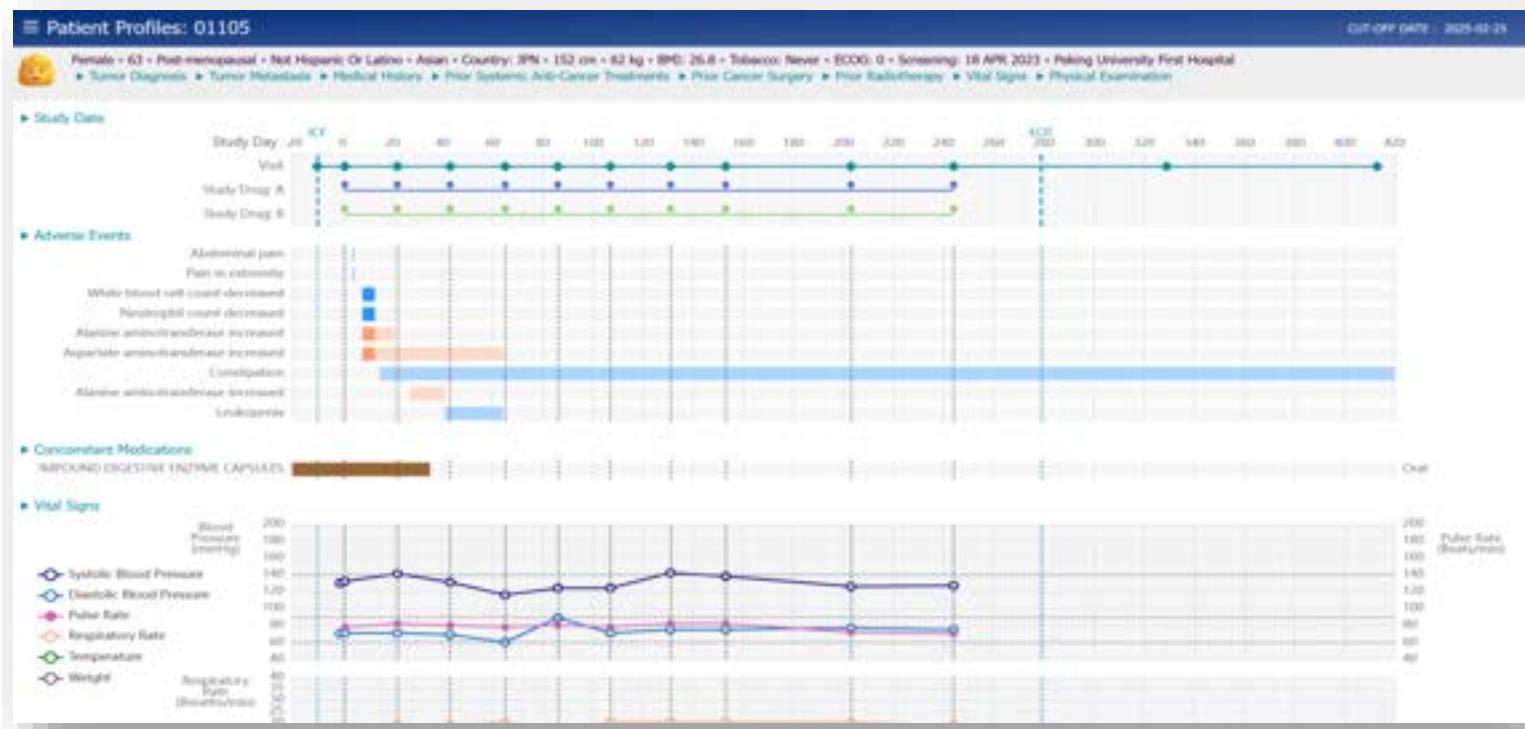
## 可视化结果-疗效性: Oncology



	Overall (N= 60)	Cohort X (N= 31)	Cohort Y (N= 29)
Censored Summary, n (%)	Censored Event 14 (23.3) 46 (76.7)	6 (19.4) 25 (80.6)	8 (27.6) 21 (72.4)
Progression-Free Survival (Months), Kaplan-Meier Estimate	Median (95% CI) Q1 (95% CI) Q3 (95% CI) 4.73 (2.89, 6.60) 1.76 (1.41, 2.89) 11.30 (6.90, -)	4.73 (2.89, 6.60) 1.71 (1.28, 2.89) 9.13 (3.62, -)	4.83 (1.04, 11.30) 1.81 (1.33, 4.40) 10.14 (5.82, -)
PFS Rate (%), Kaplan-Meier Estimate	3 Months (95% CI) 6 Months (95% CI) 9 Months (95% CI) 12 Months (95% CI) 15 Months (95% CI) 18 Months (95% CI) 21 Months (95% CI) 24 Months (95% CI) 27 Months (95% CI) 61.7 (48.2, 72.6) 39.5 (27.1, 51.6) 30.9 (19.7, 42.0) 24.0 (14.1, 35.5) 24.0 (14.1, 35.5) 24.0 (14.1, 35.5) 21.4 (11.7, 33.0) 21.4 (11.7, 33.0) 19.4 (7.9, 34.6) 19.4 (7.9, 34.6) 19.4 (7.9, 34.6) 19.4 (7.9, 34.6) -	56.8 (36.0, 70.3) 38.7 (22.0, 55.1) 29.0 (14.5, 45.3) 19.4 (7.8, 34.6) 19.4 (7.8, 34.6) 19.4 (7.8, 34.6) 19.4 (7.8, 34.6) 19.4 (7.8, 34.6) -	68.0 (48.8, 82.5) 40.0 (22.3, 57.2) 32.0 (16.6, 50.0) 29.1 (13.9, 46.3) 29.1 (13.9, 46.3) 29.1 (13.9, 46.3) 21.8 (7.5, 40.9) -

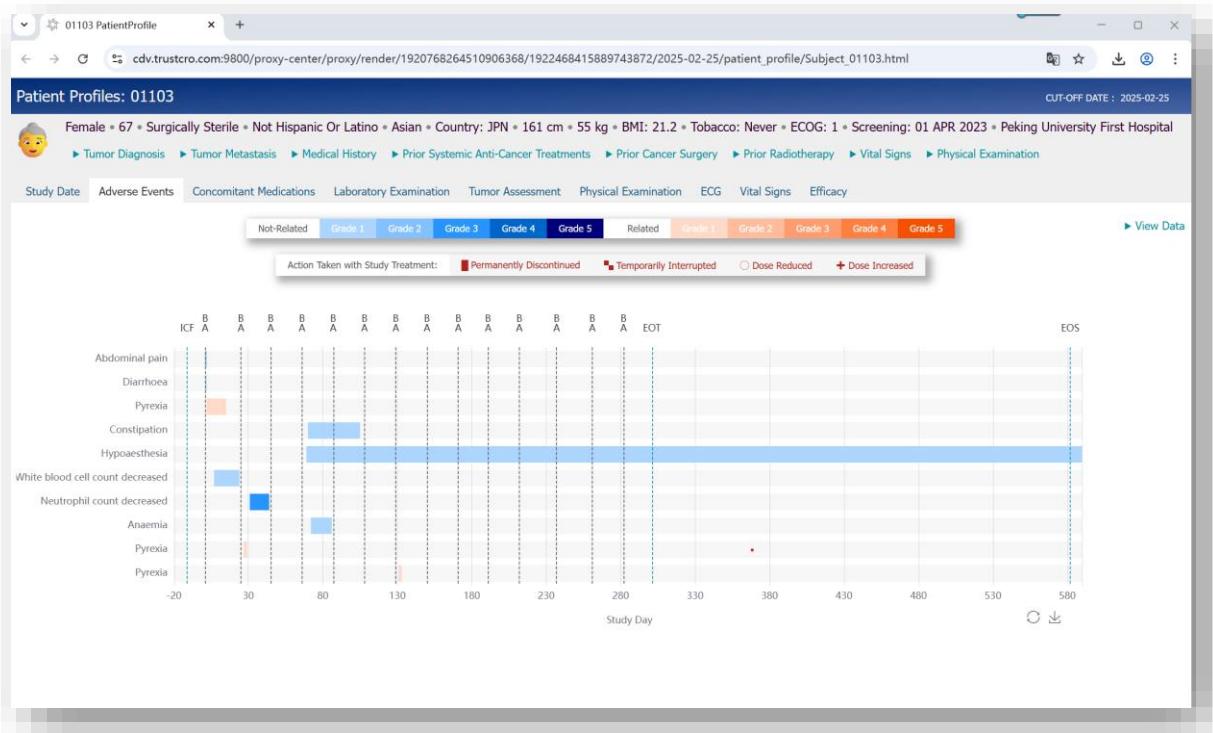
# 价值绽放 – 可视化层

## 可视化结果-Patient Profile



# 价值绽放 – 可视化层

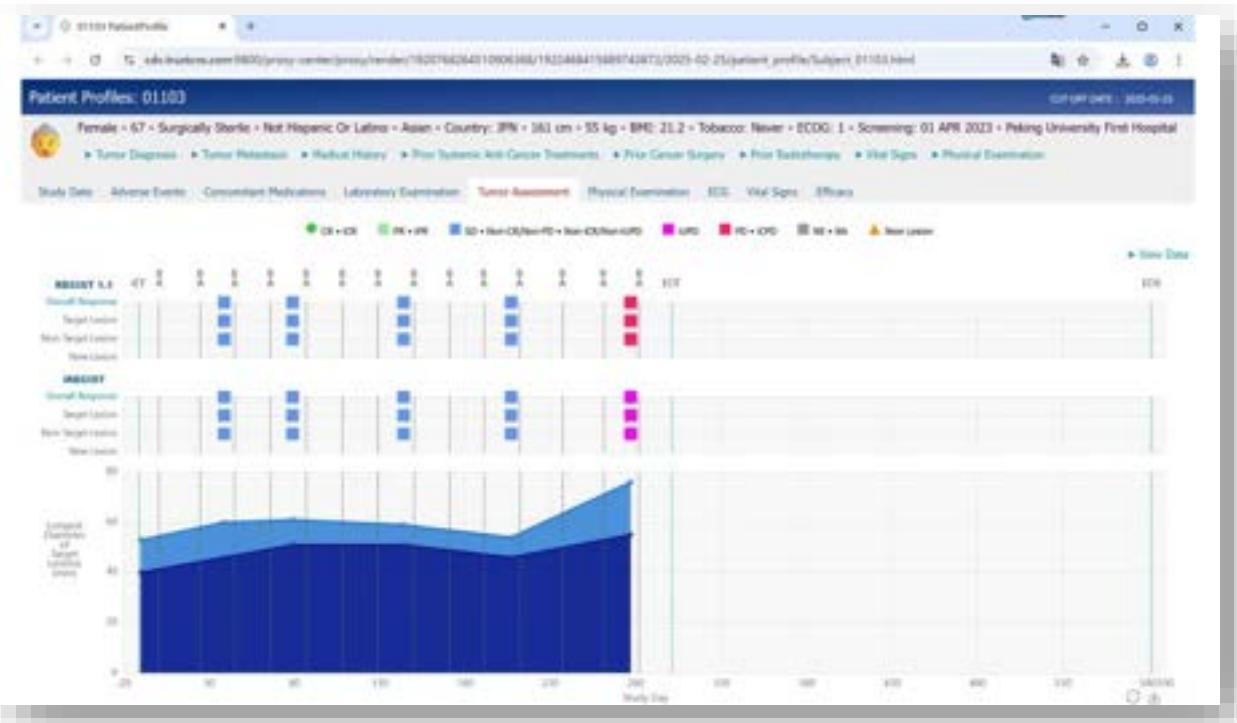
## 可视化结果-Patient Profile



- 根据医学审核的习惯，生成标准化的个体受试者全生命周期数据 (Patient Profile)

# 价值绽放 – 可视化层

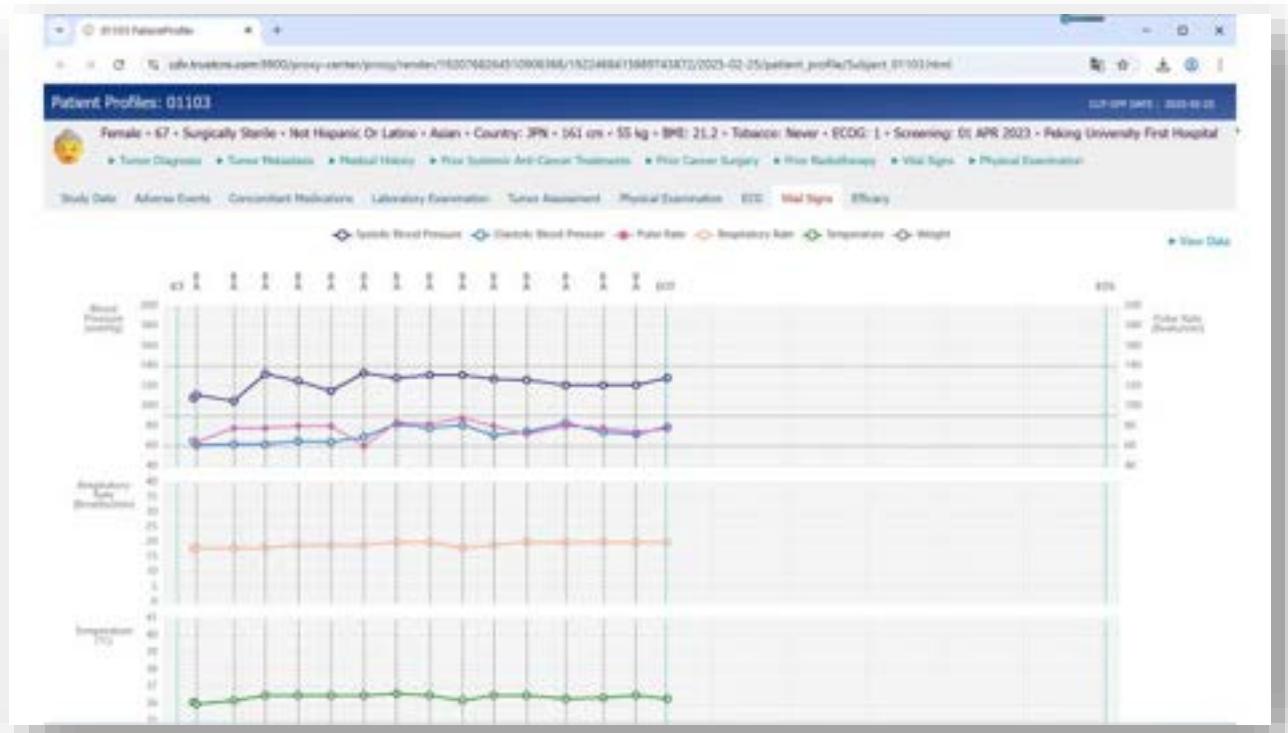
## 可视化结果-Patient Profile



- 根据医学审核的习惯，生成标准化的个体受试者全生命周期数据（Patient Profile）

# 价值绽放 – 可视化层

## 可视化结果-Patient Profile



- 根据医学审核的习惯，生成标准化的个体受试者全生命周期数据（Patient Profile）

# 价值绽放 – 可视化层

## 可视化结果-SDTM数据

AE: Adverse Event													SUBJECT ID: 03103	SITE: Zhongshan Hospital	SITE ID: 03	CUT-OFF DATE: 2025-03-25
RECORDPOSITION AETERN		AESTDAT	ARONOD	AEENDAT	AERAE	AERELIST	AEISER	AEIEL	AERELB	AERELDOS	AESEV	AEANOA	AEACNB	AEOUT		
Record number	Adverse Event	Start Date (DD-MMM-YYYY)	Ongoing	End Date (DD-MMM-YYYY)	Was the event an immunotherapy-related AE (iAE)?	Was the event an AE of Special Interest?	Was the event serious?	Relationship to Drug A	Relationship to Drug B	Was this event related to overdose?	CTCAE Grade	Action Taken with Drug A	Action Taken with Drug B	Outcome		
1	GASTROINTESTINAL HEMORRHAGE	14 MAY 2023	No	21 JUN 2023	No	No	Yes	Not Related	Not Related	No	Grade 3	Temporarily Interrupted	Temporarily Interrupted	Recovered/Resolved		
2	ANOREXIA	20 JUN 2023	Yes		No	No	No	Not Related	Not Related	No	Grade 2	None	None	Not Recovered/Not Resolved		
3	GGT INCREASED	20 JUN 2023	Yes		No	No	No	Related	Related	No	Grade 2	None	None	Not Recovered/Not Resolved		
4	URINARY PROTEIN POSITIVE	27 APR 2023	Yes		No	No	No	Not Related	Not Related	No	Grade 1	None	None	Not Recovered/Not Resolved		
5	ACUTE CEREBRAL INFARCTION	29 SEP 2023	No	29 SEP 2023	No	No	Yes	Not Related	Not Related	No	Grade 5	Not Applicable	Not Applicable	Fatal		
6	WEIGHT LOSS	21 JUN 2023	Yes		No	No	No	Not Related	Not Related	No	Grade 1	None	None	Not Recovered/Not Resolved		

穿透至个体：

- SDTM数据
- 数据库**基本**完成SDTM转换

# 价值绽放 – 可视化层

## 可视化结果-不同时期的SDTM数据的比较

Change Data Capture (2025-02-25 vs. 2025-02-11)											Search Subject ID ...	
Note: The subjects with a data row that is added, modified or deleted are listed below. Click the Subject ID to view details. ■ Data Changed ■ Rows Deleted ■ Rows Added ■ Subjects Newly Enrolled												
Subject ID	Study Arm	Site ID	Sex	Age	Screening	First Dose of Study Drug (Day 1)	Latest Dose of Study Drug	Latest Visit	End of Treatment (EOT)	End of Study (EOS)	Death	SAE
01101	Cohort X	01	Female	65	2023-03-14 (-11)	2023-03-25	2023-09-13 (173)	2024-02-28 (341)	2023-09-28 (188)	2024-03-20 (362)	2024-03-20 (362)	
02105	Cohort X	02	Female	66	2023-02-03 (-6)	2023-02-09	2023-08-17 (190)	2024-11-13 (644)	2023-08-17 (190)	2024-11-13 (644)		
02108	Cohort Y	02	Male	73	2023-04-14 (-25)	2023-05-09	2023-07-18 (71)	2024-06-16 (405)	2023-08-15 (99)			
02403	Cohort X	02	Male	71	2023-04-11 (-22)	2023-05-03	2023-09-06 (127)	2024-07-25 (450)	2023-09-26 (147)			

## 价值绽放：可视化层——基于标准化基座的认知增强

Trust MedLens可视化平台不是一种“画图工具”，是一种“认知交互界面”。

- **宏观 - 全局态势感知：** 基于各类元数据，通过算法动态构建试验级风险预警。
- **中观 - 关联网络分析：** 对标准化后的数据（如AE、CM）进行关联性挖掘发现隐藏的药物-事件关系，将可视化从“展示”提升到“发现”。
- **微观 - 高保真证据链溯源：** 从TFL→ADaM→SDTM→eCRF的无损、高速穿透。

基于标准的SDTM/ADaM和结构化元数据，没有CDISC标准，这一切也无法实现。

# 总结与展望：迈向认知驱动的临床开发

## 总结：技术架构的升维

- 我们从“文档驱动、人力密集型”的作坊模式，进化到了“元数据驱动、自动化”的流水线模式。
- 标准化（SDTM/ADaM）是基座，元数据驱动是引擎，基于标准化的可视化结果是我们的终极目标。

# 总结与展望：迈向认知驱动的临床开发

## 未来展望：

- **主动式风险干预：**平台基于实时数据流预测风险并采取相应措施。
- **生成式探索：**积极拥抱AI，用户用自然语言询问数据问题，平台自动生成相应的分析图表。
- **最终愿景：**构建一个“自我描述、自我验证、可实时探索”的临床数据平台。



Thank You!

