

PINNACLE²¹

针对向中国NMPA 递交CDISC数据做准备

Max Kanevsky – 创始人兼首席执行官

2020年9月29日

P21



MAX KANEVSKY

CEO

- ▶ Pinnacle 21创始人
- ▶ FDA和PMDA的验证解决方案架构师
- ▶ FDA JumpStart和coreDF方面的内容专家

议程

- 新NMPA要求
- 当前的数据验证挑战
- Pinnacle 21的实施计划
- 首个NMPA引擎
- 未来规划

P21





新NMPA要求

P21

接受CDISC数据递交

- ▶ 指导于2020年7月20日发布
<http://www.cde.org.cn/news.do?method=viewInfoCommon&id=7a43c3abfde95950>
- ▶ 于2020年10月1日生效，适用药品和生物制剂
- ▶ 中医部分待定

关键要求

- ▶ 数据需以中文呈现
- ▶ SAS XPORT V5格式
- ▶ SDTM和ADaM
- ▶ Define (XML或PDF)
- ▶ 必须递交DM和ADSL,
其他数据如果适用也需递交

外文数据库

- 中文为主，与《药品注册管理办法》一致
- 翻译内容的考虑：提高审评效率并考虑申办者工作量
- 翻译最低要求如下：
 - 数据库：数据集标签和变量标签；在临床总结报告等文件中出现的不良事件名称、合并用药名称、病史名称。
 - 数据说明文件：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。
 - aCRF：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。
- **注意事项：申报资料不同文件之间的中文保持一致**

翻译最低要求

- ▶ 数据集内容
 - ▶ 数据集和变量标签
 - ▶ 不良事件和病史术语
 - ▶ 合并用药
 - ▶ 临床终点
- ▶ Define.xml
- ▶ aCRF
- ▶ 数据审阅说明
- ▶ 数据集和文件（如临床总结报告）
之间的中文翻译需保持一致

数据集及变量相关要求

- 数据集名称的最大长度为8个字节
- 变量名称的最大长度为8个字节
- 数据集标签和变量标签应使用**中文**
- 拆分：存储大小，可仅递交拆分后的数据集

注意事项：在数据审阅说明中，应详细说明数据集的拆分规则及合并的详细步骤

数据集要求

- ▶ 数据集和变量名称不得超过8个字节
- ▶ 变量标签不得超过40个字节
- ▶ 变量长度不得超过200个字节
- ▶ 数据集和变量标签应以中文呈现
- ▶ 依据编码，一个字符可能为1-4个字节
- ▶ 拆分数据集时，务必在数据审阅说明中提供拆分和合并规则细节



当前的数据验证挑战

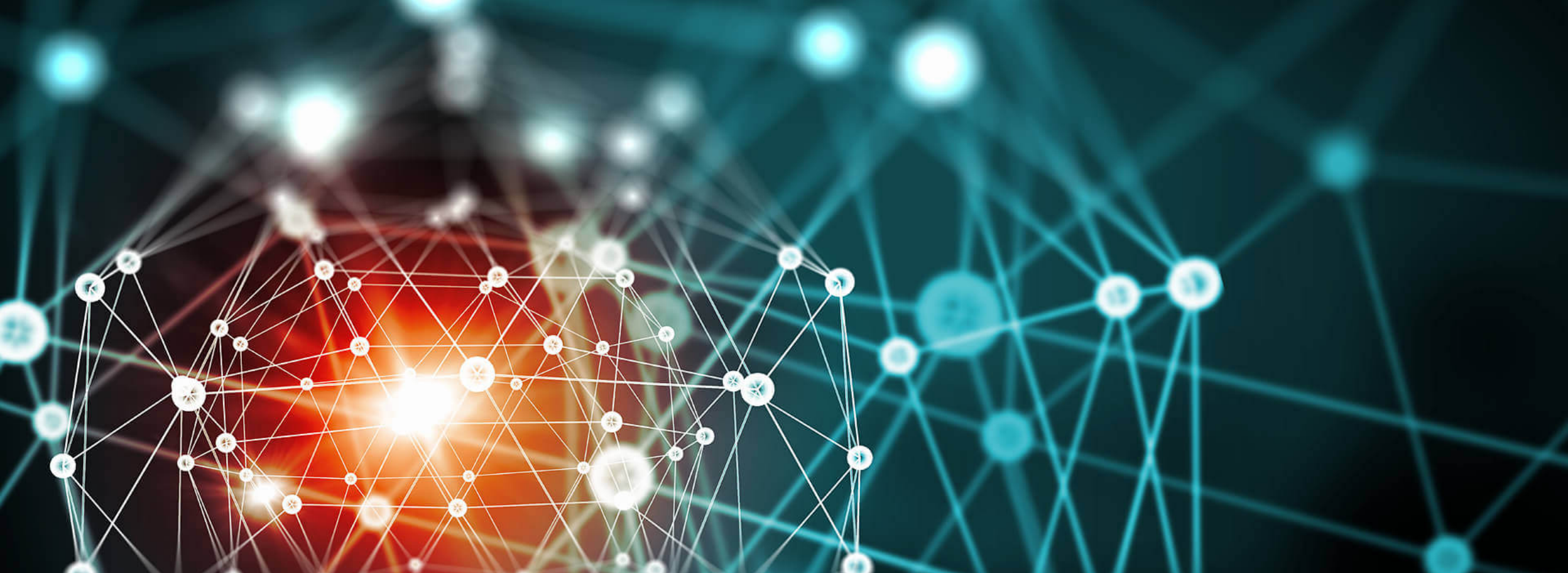
P21

新的验证要求

- ▶ 对于中文字符编码的支持
 - 原支持：LATIN1
 - 需支持：UTF-8, EUC-CN, GBK, ...
- ▶ 中文提示及描述
- ▶ NMPA数据递交规则子集

挑战

- ▶ CDISC仍在进行标准和术语的中文翻译
 - ▶ SDTM-IG 3.2已经可用，SDTM-IG 3.3 进行中
 - ▶ ADAM-IG 1.1 进行中
 - ▶ 常见编码表的CDISC受控术语翻译也在进行中
- ▶ 还需要其他文件的翻译
 - ▶ Define.xml样式表
 - ▶ PHUSE数据审阅说明模板
- ▶ NMPA业务规则尚未定稿

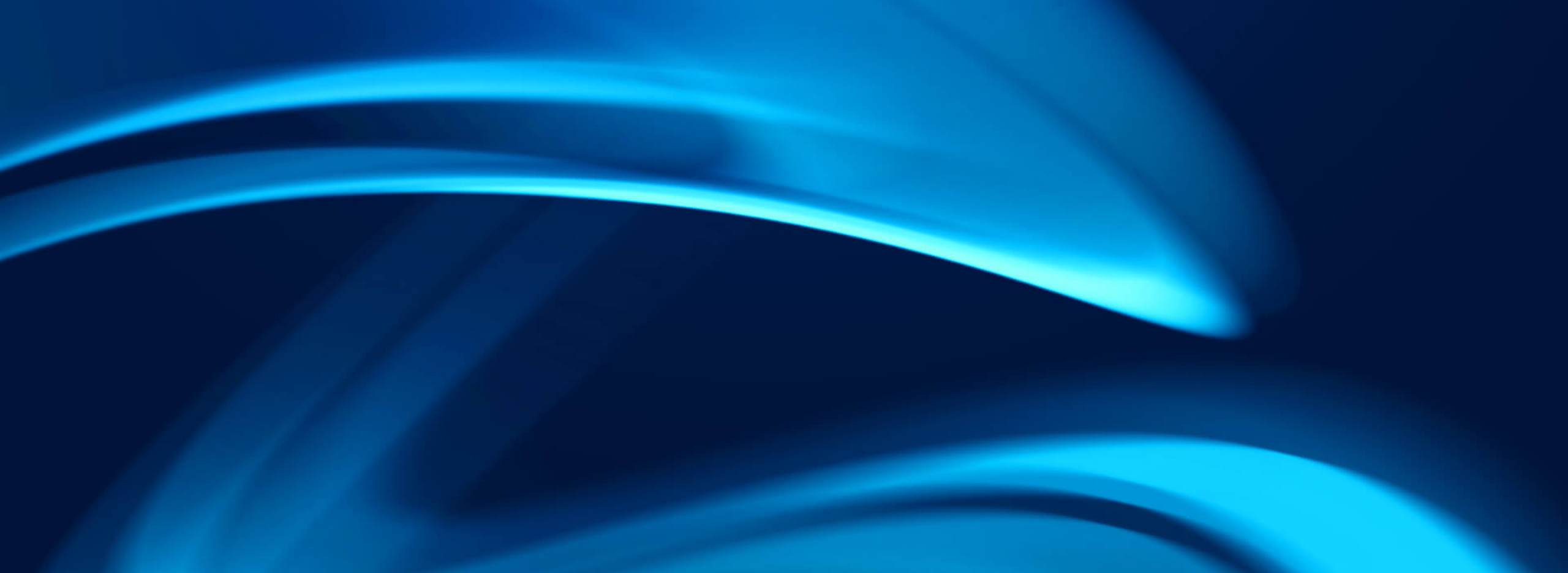


P21

PINNACLE 21的实施计划

加快实施速度

- ▶ 快速发布初始版本
 - ▶ 基本功能
- ▶ 每月更新
 - ▶ 新功能
 - ▶ 添加对新标准和版本的支持
 - ▶ 调整（改善）规则
- ▶ 与所有利益相关方密切合作
 - ▶ NMPA
 - ▶ CDISC
 - ▶ 用户



首个NMPA引擎

P21



八月

- ▶ NMPA 1907.2 引擎已于2020年8月24日发布
 - ▶ Community: NMPA 1907.3
- ▶ SDTM-IG 3.2和3.1.3
 - ▶ 288条规则
 - ▶ 中文MedDRA
- ▶ UTF-8编码
- ▶ 简体中文提示及描述
- ▶ Pinnacle 21中文论坛

选择新的NMPA引擎

Community

Enterprise

Validator | *check compliance with SDTM, SEND, ADaM, and Define.xml*

✓ Validate Data

Engine	NMPA (1907.3) ▲	Standard
Source Format	NMPA (1907.3)	Configuration
Source Data	FDA (1907.2)	
	FDA Legacy (1903.1)	
	PMDA (1810.3)	
	PMDA Legacy (1511.6)	

Drop your files here to select them or

Home > Configure > Studies > My First NMPA submission > New Data Package

New Data Package

Details

Compatibility

Validation Engine

- P21 1907.2
- FDA 1907.2
- FDA 1903.1 (Legacy)
- PMDA 1810.3
- PMDA 1511.6 (Legacy)
- NMPA 1907.2

Version: 1907.2
Chinese-language validation engine. Use to prepare for NMPA submission.

P21

中文数据的问题详细说明

Domain	Record	Count	Variables	Values	Pinnacle 21 ID	Publisher ID	Message
AE		3	AEOUT	痊愈/缓解	SD0037		未能在(Outcomes of Events for AE)用户定义的代码列表中找到AEOUT变量的值
AE		2	AEPRESP	是	SD0037		未能在(No Yes Response (Yes only))用户定义的代码列表中找到AEPRESP变量的值
AE		1	AEREL	可能有关	SD0037		未能在(Causality)用户定义的代码列表中找到AEREL变量的值
AE		2	AEREL	有关	SD0037		未能在(Causality)用户定义的代码列表中找到AEREL变量的值
AE		1	AESEV	严重	SD0037		未能在(Severity/Intensity Scale for AE)用户定义的代码列表中找到AESEV变量的值
AE		1	AESEV	中度	SD0037		未能在(Severity/Intensity Scale for AE)用户定义的代码列表中找到AESEV变量的值
AE		1	AESEV	轻度	SD0037		未能在(Severity/Intensity Scale for AE)用户定义的代码列表中找到AESEV变量的值
DM		1	ETHNIC	西班牙或拉丁裔	SD0037		未能在(Ethnic Group)用户定义的代码列表中找到ETHNIC变量的值
DM		5	ETHNIC	非西班牙或拉丁裔	SD0037		未能在(Ethnic Group)用户定义的代码列表中找到ETHNIC变量的值
DM		1	RACE	夏威夷原住民或其他太平洋岛国人	SD0037		未能在(Race)用户定义的代码列表中找到RACE变量的值
DM		2	RACE	白人	SD0037		未能在(Race)用户定义的代码列表中找到RACE变量的值
DM		1	RACE	美国印第安人或阿拉斯加原住民	SD0037		未能在(Race)用户定义的代码列表中找到RACE变量的值
DM		1	RACE	阿拉斯加加人	SD0037		未能在(Race)用户定义的代码列表中找到RACE变量的值
DM		1	RACE	黑人或非洲裔美国人	SD0037		未能在(Race)用户定义的代码列表中找到RACE变量的值
DM			VARIABLE	DTHDTC	SD0057	CG0016	未能找到SDTM规定的期望变量DTHDTC
DM			VARIABLE	DTHFL	SD0057	CG0016	未能找到SDTM规定的期望变量DTHFL
DM			VARIABLE	RFICDTC	SD0057	CG0016	未能找到SDTM规定的期望变量RFICDTC
DM			VARIABLE	RFICTTC	SD0058	CG0013, CG0542, CG0564, CG0565, CG0566, CG0567	变量在数据集中出现, 但是不在SDTM模型中

规则提示和描述翻译

Pinnacle 21 ID	Publisher ID	Message	Description	Category
DD0101		define.xml文件缺失	监管机构提供的指南表明, 递交eCTD时, 4号模组(非临床)与5号模组(临床)中每个研究都必须有一个define.xml文件。备注: 验收标准及其版本请参考相关监管机构的数据标准目录。	Presence
SD0001	CG0408	数据源中没有记录	域名表中应至少有一个记录。	Presence
SD0002	CG0014, CG0208	必需Variable变量的值为空值(NULL)	任何记录中必需变量("核心"属性为'Req'的变量)都不能为空(NULL)。	Presence
SD0003	CG0238	ISO 8601 Variable变量值无效	日期/时间变量值(*DTC)必须符合ISO 8601国际标准。	Format
SD0004	CG0413	DOMAIN值不一致	域名缩写(DOMAIN)变量应与数据集的名称一致。	Consistency
SD0005	CG0028	-SEQ变量值重复	在一个域中, 每条记录的顺序变量(-SEQ)的数值都必须是唯一的。该域中存在的每个唯一受试者标识符(USUBJID), 池标识符(POOLID), 或申办者设备标识符(SPDEVID)变量值记录的顺序变量(-SEQ)也必须是唯一的。	Consistency
SD0007		标准单位值不一致	测量、测试、或检查简称(-TESTCD)、类别(-CAT)、样本类型(-SPEC)、以及测试或检查方法(-METHOD)相同的记录, 其标准化单位(-STRESU)必须保持一致。	Consistency
SD0008	CG0379	未能在MedDRA词典中找到-DECOD的值	词典衍生术语(-DECOD)变量的值必须用define.xml指定的MedDRA词典中的词典首选术语(Preferred Term)填写(不分大小写)。	Terminology
SD0008C		-DECOD值大小写错误	词典衍生术语(-DECOD)变量必须首字母大写, 并使用define.xml指定版本的MedDRA词典中的词典首选术语(Preferred Term)(区分大小写)。	Terminology
SD0012		-STDY在-ENDY之后	事件、接触、或观察的开始日期(-STDY)必须小于或等于事件、接触、或观察的结束日期(-ENDY)。	Limit
SD0013		-STDTC在-ENDTC之后	事件、接触、或观察的开始日期/时间(-STDTC)必须小于或等于事件、接触、或观察的结束日期/时间(-ENDTC)	Limit
SD0014		-DOSE为负值	剂量(-DOSE)值没有缺失时必须大于等于0。	Limit
SD0015		-DUR为负值	事件、接触、或观察的持续时间(-DUR)值没有缺失时必须大于等于0。	Limit
SD0021		结束时间点缺失	一次事件或干预发生时, 应填入一个结束时间点变量(比如: 除非是否发生的变量(-OCCUR)='N', 或完成状态(-STAT)!='', 或持续时间(-DUR)!='', 否则事件或干预的结束日期/时间(-ENDTC)、或相对于参照期间的结束(-ENRF)、或相对于参照期间的结束(-ENRTP)不应缺失)。	Consistency



下一步

P21

九月

- ▶ SDTM
 - ▶ 添加新规则，以及对中文受控术语的支持
- ▶ Define-XML
 - ▶ 翻译提示及描述
 - ▶ 中文样式表
- ▶ 其他编码
 - ▶ UTF-8
 - ▶ EUC-CN, GB2312
 - ▶ MS936, CP936, GBK
 - ▶ GB18030
 - ▶ Big5

选择数据集编码

Community

Enterprise

Validator check compliance with SDTM, SEND, ADaM, and Define.xml

Validate Data

Engine: NMPA (1907.3) Standard: SDTM

Source Format: SAS® Transport (XPORT) Configuration: SDTM-IG 3.2 (NMPA)

Source Encoding: EUC-CN, GB2312

Source Data: UTF-8, EUC-CN, GB2312, MS936, CP936, GBK, GB18030, Big5

Drop your files here to select them or [Browse...](#)

Home > Validate Data

Validate Data

Validation Setup

Project: NMPA submission

Study: My First NMPA submission

Data Package: SDTM

Data Encoding: UTF-8

Comment: EUC-CN, GB2312, MS936, CP936, GBK, GB18030, Big5 (Optional)

中文的DEFINE.XML

心病心绞痛治

Define-XML文档的创建日期/时间: 2020-09-11T21:11:07-04:00
 Define-XML版本: 2.0.0
 样式表版本: 2018-11-21

Annotated Case Report For

▶ 补充文件

▼ 数据集

AE (不良事件)

DM (人口学)

TA (试验组)

▼ 受控术语

▼ 代码清单

实际分组描述

实际分组编码

对研究治疗采取的措施

转归

相关性

严重程度

年龄单位

计划分组描述

计划分组编码

国家

元素描述

时期

元素编码

族群

No Yes Response (No o

种族

性别

No Yes Response (Yes c

▶ 外部词典

标准	SDTM-IG 3.2
研究名称	心病心绞痛治
研究概要	医冠心病心绞痛治疗领域
协议名称	002-心心绞
元数据名称	Study 心病心绞痛治 Data Definitions

数据集

数据集	描述	类	结构	目的	按键	文献资料	位置
AE	不良事件	事件类	每个受试者每个不良事件一条记录	Tabulation	STUDYID, USUBJID, AEDECOD, AESTDTC		ae.xpt
DM	人口学	特殊目的	每个受试者一条记录	Tabulation	STUDYID, USUBJID		dm.xpt
TA	试验组	试验设计	每个试验组的每个元素对应一条记录	Tabulation	STUDYID, ARMCD, TAETORD		ta.xpt

回到Define-XML文档的[顶部](#)

TA (试验组) - 试验设计

位置: [ta.xpt](#)

变量	标签/描述	类型	角色	长度或显示格式	受控术语或ISO格式	来源/源/方法/注释
STUDYID	研究标识符	text	标识符	6		Protocol
DOMAIN	域名缩写	text	标识符	2		Assigned
ARMCD	计划分组 编码	text	主题	1	计划分组编码 <ul style="list-style-type: none"> "A" = "药 A" "B" = "药 B" "P" = "安慰剂 ≤" "SCRNFAIL" = "筛查失败" 	Assigned
ARM	计划分组 描述	text	同义词修 饰语	13	计划分组描述 [6 术语]	Protocol

十月

- ▶ ADaM

- ▶ 对ADaM-IG 1.1的初步支持
- ▶ 翻译提示及描述

- ▶ SDTM

- ▶ 将持续添加CDISC受控术语和标准元数据翻译的相关规则
- ▶ 添加任何NMPA具体规则

未来

- 软件本地化
- 将用户界面和报告翻译为中文和日文

合作

- ▶ NMPA
 - ▶ 明确业务规则
- ▶ CDISC
 - ▶ 翻译各项标准
- ▶ PhUSE
 - ▶ 翻译数据审阅说明模板
- ▶ 用户
 - ▶ 测试
 - ▶ 申请新功能
 - ▶ 用户教育

有问题吗？



谢谢

请保持联系!



MAX KANEVSKY

MAX@PINNACLE21.COM

幻灯片张贴至我们的博客后，
将以电子邮件通知您

www.pinnacle21.com/blog

P21