



## SEND数据递交案例分享-SEND与SDTM生成过程中的异同点解析

Presented by

吴超珺，统计编程部副总监，凯莱英临床(凯诺)



# Meet the Speaker

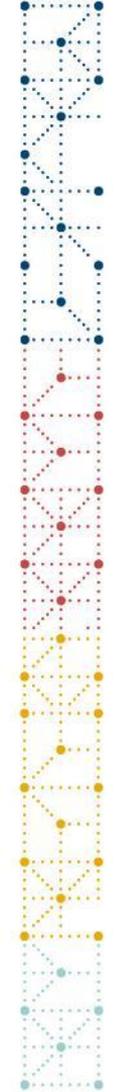
吴超珺

**Title:**统计编程部副总监

**Organization:**凯莱英临床(凯诺)

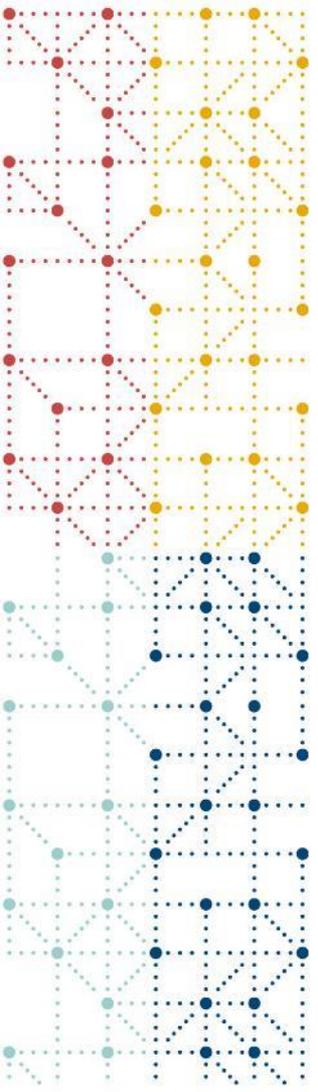
从事统计编程工作十余年，曾任职于泰格医药、太美医疗，在肿瘤、血液病、糖尿病、高血压等治疗领域均有丰富的编程经验。

2019年7月加入凯诺医药，目前主要负责项目管理与执行、团队管理等工作。



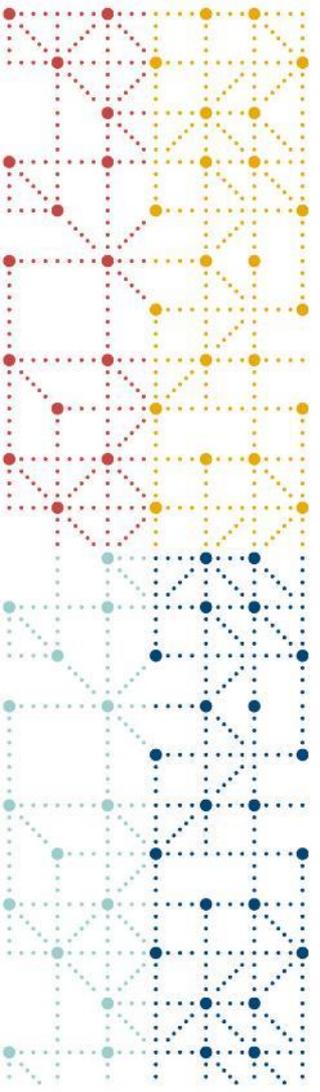
# Disclaimer and Disclosures

- *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.*
- *The author(s) have no real or apparent conflicts of interest to report.*



## Agenda

1. SEND基本介绍
2. 从实施指南看SEND与SDTM的异同点
3. FDA的sdTCG文档中对SEND的相关要求
4. 案例分析



# SEND基本介绍

SEND的定义与作用

# SEND (Standard for Exchange of Nonclinical Data)

## SEND的定义

SEND是SDTM标准在非临床研究中的实施

- SEND指定了一种以统一格式收集和呈现非临床数据的方法
- SEND是向FDA提交数据所需的标准之一

# SEND (Standard for Exchange of Nonclinical Data)

## SEND的作用

为非临床研究数据列表提供标准框架

- 标准数据集
- 标准变量
- 受控术语
- 公共变量

# SEND实施指南 (from CDISC SEND team)

## SENDIG V3.1.1

该版本的SENDIG旨在支持在单剂量一般毒理学、重复剂量一般毒理学和致癌性研究，以及在安全药理学研究期间进行的呼吸和心血管测试中发现的数据。

- 发布日期: 30 March 2021
- 基于SDTM V1.5衍生而成
- 由于FDA一直是鼓励SEND发展的主要监管机构，因此本文件参考了该机构的一些具体实施要求
- 注：SENDIG中的“subject (受试者)” 指代动物

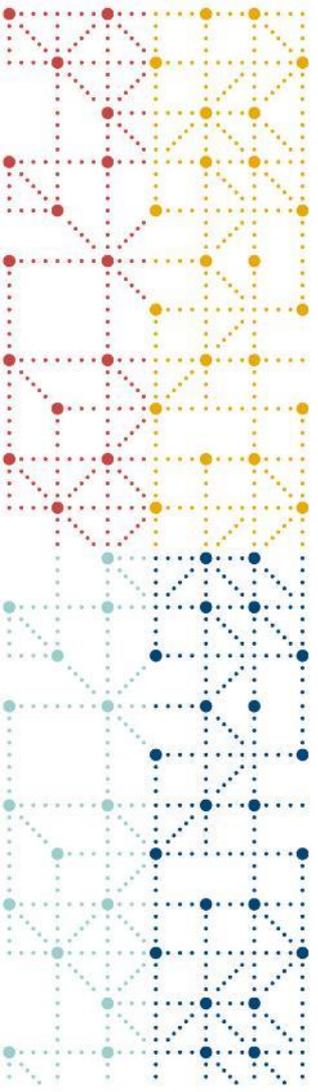
# FDA sdTCG中的eCTD拒收标准

Contains Nonbinding Recommendations



**Table 6: eCTD Technical Rejection Criteria for Study Data Expectations<sup>92</sup>**

Data Type	Modules & Submodules	Center	Application Type	Study Start Date	Requirement
Non-clinical	4.2.3.1, 4.2.3.2, 4.2.3.4	CDER	NDA, BLA, ANDA	On/Prior to December 17, 2016	Submit simplified ts.xpt*
				After December 17, 2016	Comply with CDISC standards
			Commercial IND	On/Prior to December 17, 2017	Submit simplified ts.xpt*
				After December 17, 2017	Comply with CDISC standards
		CBER	NDA, BLA, ANDA, Commercial IND	On/Prior to March 15, 2023	Submit simplified ts.xpt*
				After March 15, 2023	Comply with CDISC standards



# 从实施指南看SEND与SDTM的异同点

SEND与SDTM相同域的异同点解析

(SEND IG V3.1.1 vs SDTM IG V3.2)

# SEND与SDTM共有域总览

SEND Dataset name	SEND Dataset Description	Class
CO	Comments	Special Purpose
DM	Demographics	Special Purpose
SE	Subject Elements	Special Purpose
EX	Exposure	Interventions
DS	Disposition	Events
DD	Death Diagnosis and Details	Findings
EG	ECG Test Results	Findings
LB	Laboratory Test Results	Findings
MI	Microscopic Findings	Findings
PC	Pharmacokinetics Concentrations	Findings
PP	Pharmacokinetic Parameters	Findings
SC	Subject Characteristics	Findings
VS	Vital Signs	Findings
CV	Cardiovascular Test Results	Findings
RE	Respiratory Test Results	Findings
TA	Trial Arms	Trial Design
TE	Trial Elements	Trial Design
TS	Trial Summary	Trial Design
RELREC	Related Records	Relationships
--SUPP	Supplemental Qualifiers	Relationships

# SEND特有域总览

SEND Dataset name	SEND Dataset Description	Class
TX	Trial Sets	Trial Design
BW	Body Weight	Findings
BG	Body Weight Gain	Findings
CL	Clinical Observations	Findings
FW	Food and Water Consumption	Findings
MA	Macroscopic Findings	Findings
OM	Organ Measurements	Findings
PM	Palpable Masses	Findings
TF	Tumor Findings	Findings
POOLDEF	Pool Definition	Relationships

# POOLDEF域 – 汇集具有某一相同特征的受试者

只包含三个Req变量: **STUDYID / POOLID / USUBJID**

POOLID (为SEND开发的变量):

- 如果在某个域中用到了该变量，则需将USUBJID置空；同时应该衍生POOLDEF域

一个池必须包含至少一个受试者

*pooldef.xpt*

Row	STUDYID	POOLID	USUBJID
1	ABC1	CAGE1 P1	ABC1-1001
2	ABC1	CAGE1 P1	ABC1-1002
3	ABC1	CAGE1 P1	ABC1-1003
4	ABC1	POOL1	ABC1-1005
5	ABC1	POOL1	ABC1-1007

# 试验设计域的异同点

## SEND的试验设计域:TA, TE, TX, TS

与SDTM相比:

- 不涉及入排标准(TI)、访视信息(TV)、试验疾病评估(TD)
- TX(Trial Sets)是SEND独有的域
  - 申办者定义的计划的受试者集
  - 通过SET/SETCD的不同值来区分试验集

# TX域示例

*tx.xpt*

Row	STUDYID	DOMAIN	SETCD	SET	TXSEQ	TXPARMCD	TXPARM	TXVAL
1	EXP1	TX	1	Control Group, Vehicle Control once daily	1	ARMCD	Arm Code	1
2	EXP1	TX	1	Control Group, Vehicle Control once daily	2	SPGRPCD	Sponsor-Defined Group Code	1
3	EXP1	TX	1	Control Group, Vehicle Control once daily	3	TCNTRL	Control Type	Vehicle Control
4	EXP1	TX	1	Control Group, Vehicle Control once daily	4	GRPLBL	Group Label	Group 1, Control
5	EXP1	TX	1	Control Group, Vehicle Control once daily	5	TRTDOS	Dose Level	0
6	EXP1	TX	1	Control Group, Vehicle Control once daily	6	TRTDOSU	Dose Units	mg/kg/day
7	EXP1	TX	2	Low-Dose Group, 100 mg/kg Drug a once daily	7	ARMCD	Arm Code	2
8	EXP1	TX	2	Low-Dose Group, 100 mg/kg Drug a once daily	8	SPGRPCD	Sponsor-Defined Group Code	2
9	EXP1	TX	2	Low-Dose Group, 100 mg/kg Drug a once daily	9	GRPLBL	Group Label	Group 2, 100 mg/kg/day
10	EXP1	TX	2	Low-Dose Group, 100 mg/kg Drug a once daily	10	TRTDOS	Dose Level	100
11	EXP1	TX	2	Low-Dose Group, 100 mg/kg Drug a once daily	11	TRTDOSU	Dose Units	mg/kg/day

# 试验设计域的异同点

SEND的试验设计域:TA, TE, TX, TS.

与SDTM相比:

- TS的parameter有所区别-具有不同的CT
- TS少了变量: TSVALCD / TSVCDREF / TSVCDVER
- TE域未发现明显差异
- TA域中部分变量的Core不同:
  - TABRANCH (Exp-Perm)
  - TATRANS (Exp-Perm)
  - EPOCH (Req-Exp)

# DM域的异同点

## SEND的DM域:

比SDTM多了以下变量

- **AGETXT**（比如动物为6-8个月，则填“6-8”）
  - AGE和AGETXT不能同时被填充
- **SETCD**: 申办方定义，用来唯一标识所有不同的“groups”或者“trial sets”
  - 一个受试者只能属于一个试验集
- **SPECIES**, **STRAIN**, **SBSTRAIN**用来描述物种，品种信息

比SDTM少了以下变量

- **RFICDTC / RFPENDTC / DTHDTC / DTHFL / INVID / INVNAM / RACE  
ETHNIC / ACTARM / ACTARMCD / COUNTRY / DMDTC / DMDY**

# DS域的异同点

## SEND的DS域:

提供了受试者最终处置的记录

(SDTM.DS则提供了受试者的全部处置事件，并可能包括方案里程碑和其它事件)

- 每个受试者一条记录
  - 比SDTM少了DSGRPID / DSREFID / DSSPID / DSCAT / DSSCAT / EPOCH / DS DTC
- DSDECOD的CT值和SDTM不同
  - 其含义为动物的最终处置状态。
  - 不一定是死亡，可能为REMOVED FROM STUDY ALIVE / MISSING

比SDTM多了以下变量

- DSUSCHFL / VISITDY / DSNOMDY / DSNOMLBL

# 其他域（SEND与SDTM共有）的异同点

Send的其他域 (除了部分变量的差异之外):

VS域:

- VSTESTCD对应的CT值有两个：OXYSAT（血氧饱和度），TEMP（体温/肛温）；注意需使用CT (SVSTSTCD)，而非 (VSTESTCD)
- 体重不放VS，放单独的BW域中
- 血压、心率等检查项则放在CV域中；
- 呼吸频率放RE域中

# 其他域（SEND与SDTM共有）的异同点

## Send的其他域 (除了部分变量的差异之外):

### MI域:

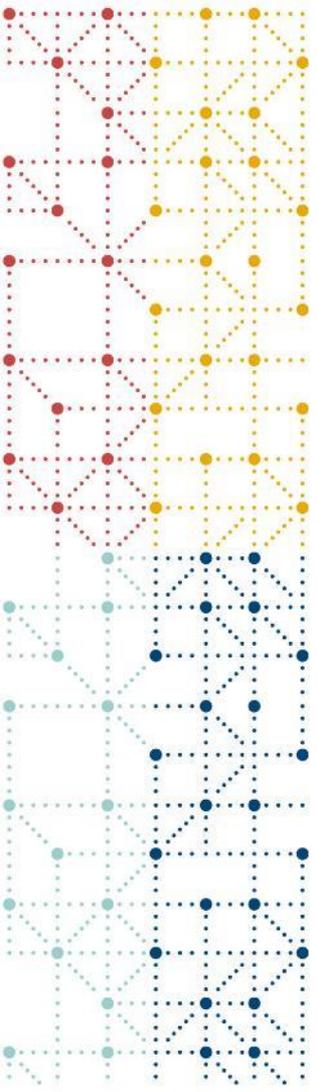
- 为整个研究过程中观察到的每个微观发现提供了记录
  - 大多收集的是动物处死后的数据，如器官的组织病理切片
  - 对于在死后采集或观察的标本，应设为受试者处置的日期/时间，而不是观察标本的时间 (此时MIDTC与DS.DSSTDTC应该相同)

# 其他域（SEND与SDTM共有）的异同点

Domain	SEND	SDTM
DD	DD域应仅用于计划外死亡的受试者； DDTEST只有一个CT：Death Diagnosis。 包含变量：--CSTATE/--EXCLFL/-- REASEX/--USCHFL/--ENDTC/--ENDY/-- NOMDY/--NOMLBL/--EVLINT/--STINT/--	用于保存通常在死亡发生时收集的补充数据，例如官方死亡原因。
EG/CV/RE	ENINT	包含变量：--SCAT/VISIT/VISITNUM
LB	包含变量： LBANTREG/LBSPCUFL/LBLOC/LBLAT/LB DIR/LBPORTOT/LBEXCLFL/LBREASEX/L BUSCHFL/LBENDY/LBNOMDY/LBNOMLBL	包含变量：LBLOINC/VISIT/VISITNUM/ 包含变量：
EX	包含变量 POOLID/FOCID/EXMETHOD/EXTRTV/EX VAMT/EXVAMTU	EXGRPID/EXREFID/EXSPID/EXLNKID/E XLNKGRP/EXCAT/EXSCAT/EXDOSRG M/EXLAT/EXDIR/EXFAST/EPOCH
PC	包含变量： POOLID/PCSPCUFL/PCEXCLFL/PCREAS EX/PCUSCHFL/PCENDY/PCNOMLBL	包含变量：PCULOQ/VISIT/VISITNUM

# 其他域（SEND与SDTM共有）的异同点

Domain	SEND	SDTM
PP	包含变量： POOLID/VISITDY/PPNOMDY/PPNOMLBL/PP TPTREF	包含变量：PPDTC
SC	包含毛发颜色、物理标记、测试动物供应商名称等信息	包含变量： SCSPID/SCCAT/SCSCAT/SCSTAT/S CREASND；包含受教育程度、婚姻状况、国籍等信息
CO	包含变量：POOLID/CODY；当RDOMAIN和USUBJID没有被填充时，假设comments属于整个研究。	USUBJID不能为空
SE	SESEQ表示实际的元素顺序（与SESTDTC/SEENDTC表示的原始顺序一致）	包含变量：TAETORD/EPOCH； TAETORD表示计划的元素顺序； SESEQ表示实际的元素顺序
RELREC	包含变量：POOLID； <b>POOLID被填充的记录不能被包含在域对域的关系中</b> (RELTYPE应为空)； RELTYPE为可有变量	RELTYPE为期望变量



# 从实施指南看SEND与SDTM的异同点

## SEND特有域及变量介绍

# SEND特有的域

以下为未在SDTM中出现的域：

SEND Domain	SEND Domain Description	Class	Notes
TX	Trial Sets	Trial Design	如前所示
BW	Body Weight	Findings	体重放在BW中，而非VS中
BG	Body Weight Gain	Findings	体重增加 (包括增加值，增加百分比，增加的均值等)；BGORRESU可能发生变化，BGSTRESU在试验中需保持一致
CL	Clinical Observations	Findings	临床体征信息；也包含眼科检查，体格检查，皮肤检查等
FW	Food and Water Consumption	Findings	食物和水的消耗量
MA	Macroscopic Findings	Findings	大多数情况下，MATEST为大体病理检查（动物处死后的数据）；对于死后采集或观察的标本，MADTC应设为受试者处置的日期/时间，同DS.DSSTDTC

# SEND特有的域

以下为未在SDTM中出现的域：

SEND Domain	SEND Domain Description	Class	Notes
OM	Organ Measurements	Findings	包含器官测量和相对器官重量的详细信息（动物处死后的数据）
PM	Palpable Masses	Findings	包含在研究试验阶段检查的任何可触肿块的信息
TF	Tumor Findings	Findings	肿瘤发现（动物处死后的数据）；TFTEST仅有CT: Tumor Examination；为MI域的子集
POOLDEF	Pool Definition	Relationships	池定义，如前所示

# SEND特有域中-SEND特有的变量

SDTM IG 2.7章节中明确规定不能在SDTM中使用的变量：

Variable name	Dataset Label	Values / Notes
--EXCLFL	Exclusion Flag	Y/null
--REASEX	Reason for Exclusion	--EXCLFL为Y时使用
--DETECT	Time in Days to Detection of Tumor	从开始给药到试验阶段最早发现肿瘤的天数；涉及TF域 “Y”=导致死亡，“N”=未导致死亡，“U”=未知。涉及MA/MI/TF域
--DTHREL	Relationship to Death	涉及MA/MI/TF域
--ANTREG	Anatomical Region of Specimen	标本的解剖区域
--USCHFL	Unscheduled Flag	Y/null
--NOMDY	Nominal Study Day for Tabulations	整数
--NOMLBL	Label for Nominal Study Day	--NOMDY的label，如：Day 28
--DISTR	Distribution Pattern of Finding	分布模式，如：漫射，聚焦，多焦点；有CT；涉及MI域
--CHRON	Chronicity of Finding	急慢性；有CT

# SEND特有域中-非SEND特有的变量

SDTM中可以按要求使用的变量：

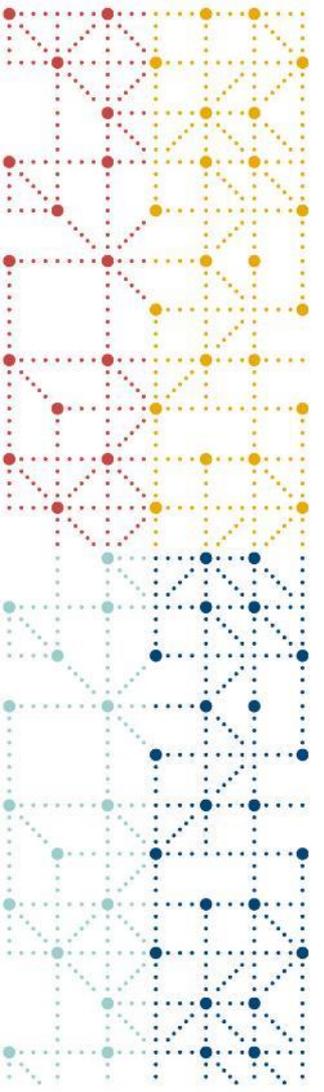
SEND Variable name	SEND Dataset Label	Values / Notes
--SPCCND	Specimen Condition	标本状况的描述
--SPCUFL	Specimen Usability for the Test	标本的可用性； N/null
--PORTOT	Portion or Totality	解剖位置或标本的限定词 (进一步说明部分或整体)； 有CT
--CSTATE	Consciousness State	测量时被测者的意识状态； 有CT
--STINT	Planned Start of Assessment Interval	相对于--TPTREF的评估间隔的计划开始时间； in ISO 8601 duration format
--ENINT	Planned End of Assessment Interval	相对于--TPTREF的评估间隔的计划结束时间； in ISO 8601 duration format

# --SPID变量的特殊用法

## MASPID (Mass Identifier)

肿块标识符：

- 标记为MASS 1 或 MASS A的格式
- CL/PM/MI/TF域中的--SPID与MASPID保持一致
- 其它域中的--SPID仍为Sponsor-Defined Identifier

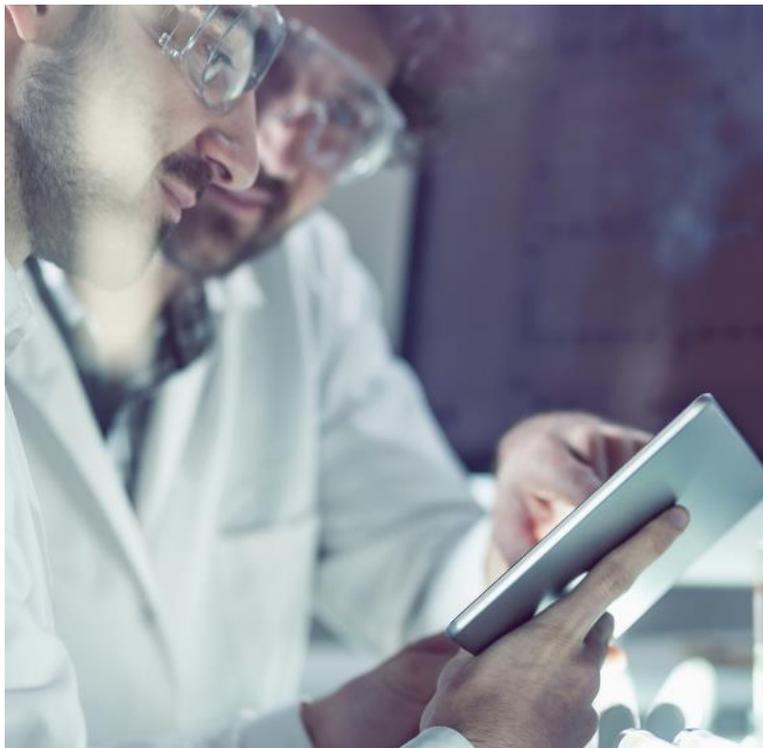


## FDA的sdTCG文档中对SEND的相关要求

# sdTCG中关于SEND的一般考虑

## 关于SEND数据集的更新:

- 如果SEND数据集有更改，需要与最终研究报告一起重新提交，需使用“replace”操作符重新提交更新后的数据集
- 如果SEND数据集没有更新，则不需要重新提交，但建议在最终研究报告中提交更新的nSDRG。



# sdTCG中关于VISITDY/--NOMDY变量

对于有宽限期的事件或检查，对应的VISITDY/--NOMDY，应该包含事件的单个计划的研究日

例如：

- 计划于day 15进行尿常规检查，但当天没有收集到任何动物的尿液，day 16成功收集了所有尿液。在研究报告中，收集第16天尿液样本的数据将与第15天的样本结果进行分析，所以VISITDY/LBNOMDY应该是15

# sdTCG中关于SEND域

SEND Domain	Notes
SUPPQUAL	若申办方打算将重要变量 (如支持关键分析的信息)放在SUPP域, 则需与 审查部门讨论, 并在nSDRG中呈现
MI	申办者应确保从MIORRES到MISTRESC的研究结果的转化严格遵守 SENDIG的说明; 有些内容需放入主域的结果限定符变量中 (如MISEV, MIDTHRL, MICHRON), 而非在MISTRESC或SUPPMI中
CL	仅包含发现记录, 不包含事件和干预的记录; CLSTRESC的标准化应严格 遵守SENDIG的说明
LB	分类的、不连续的结果应放入LBSTRESC中, 即使类别是数字的, LBSTRESN也应该为空
PC	对于单次剂量、重复剂量或致癌性研究, 如果实际/收集的信息 (PCDTC/PCDY)不容易纳入数据集, 则这些变量可能为空或填充计算或名 义日期/时间。应在nSDRG中提及计算或名义日期和时间的使用

# sdTCG中关于SEND域

SEND Domain	Notes
custom	Questions about custom domains should be addressed in pre-submission meetings and documented in the SDSP.
TDM	所有TDM数据集都应递交；恢复/毒性动力学动物试验的TA和TX域，应严格遵循SENDIG上的示例
TX	SPGRPCD/GRPLBL/PLANMSUB/PLANFSUB等参数在每个试验集中应只有一条记录
TS	附录C中包含了SEND TS域的参数；可按需添加
Tumor.xpt	致癌性研究应包括一个肿瘤发现的电子数据集 (tumor.xpt)；若tumor.xpt和SEND同时提交，则需保持数据的一致性和可追溯性，相关信息在nSDRG中呈现；若SEND数据集是创建tumor.xpt的基础，则需衍生TF域，或至少应在nSDRG中包含TFDETECT (从开始给药到试验阶段最早发现肿瘤的天数)的算法
BG	It is <b>not necessary</b> to include a BG domain in submissions.
CO	在CO域中提交的评论应与研究解释相关。为了减少歧义，应避免在任何自由文本字段中使用缩写或应在nSDRG中概述缩写

# sdTCG中关于SEND的适用范围及简化的TS

- 4.1.3.4.1章节中对SEND的适用范围做了详细的阐述
- 如果因为以上章节中概述的原因而未提交SEND时，可能需要使用简化的ts.xpt文件，其中应在TSVALNF字段中填充值“NA”

## Examples of simplified ts.xpt datasets for a nonclinical study:

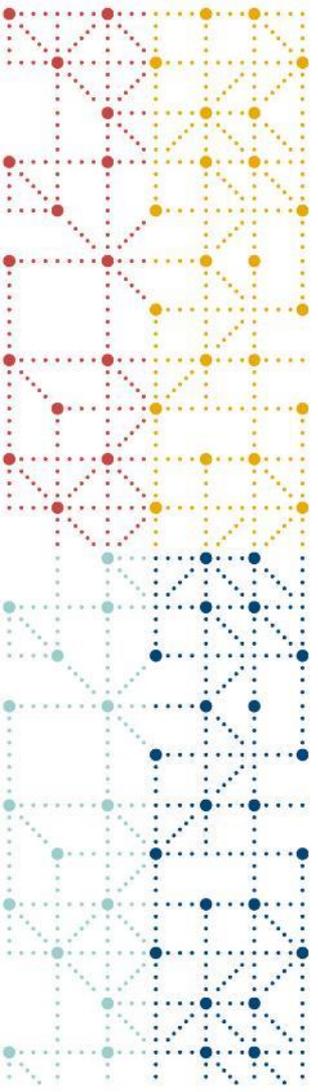
Sponsors should submit a dataset named ‘ts.xpt’ with four variables (STUDYID, TSPARMCD, TSVAL, and TSVALNF) and one row of information. Example datasets are shown below.<sup>93</sup>

### For a study with a valid SSD (example C):

STUDYID	TSPARMCD	TSVAL	TSVALNF
study ID in STF	STSTDTC	yyyy-mm-dd	

### For a study without a valid SSD (example D):

STUDYID	TSPARMCD	TSVAL	TSVALNF
study ID in STF	STSTDTC		Use the value ‘NA’



## 案例分析

SEND项目实施过程中的要点及解决方案

# 原始数据的沟通及预处理

## 原始数据特点

(通常需花费较多时间于对外沟通与数据整理)

- 多由第三方GLP (Good Laboratory Practice)单位以EXCEL格式提供，文件可能会较零散。来源可能有：
  - 使用GLP公司EXCEL模板手动记录数据
  - 从商业化软件 (如：Provantis系统) 导出的数据 (可以选择excel格式)
- 申办方可能不清楚GLP单位可提供的数据内容，不了解数据格式。不明确做SEND数据集需要哪些数据，以及什么数据结构是更方便处理的
- **解决方案：**对照报告、方案和提供的EXCEL原始数据，查漏补缺，与申办方沟通数缺少的数据内容。对于不满足格式要求的数据 (如仅有中文结果)，要求提供英文数据

# 原始数据的沟通及预处理

## 原始数据可能存在的问题

- EXCEL原始数据存在换行符等特殊字符，和科学计数法等特殊数值型格式
- **解决方案：**在读入EXCEL原始数据时，使用SAS批量处理特殊字符和科学计数法等。检查并确认处理后的原始数据与EXCEL原始数据内容等价
- 数据结构与CDISC标准的SEND数据集结构差异较大
- **解决方案：**为方便编程，在不改变原始数据内容的前提下，可手动增加辅助变量（非必要不添加）

# 增加辅助变量示例

A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	V	W
0		Number of days, relative to the day on which the administration started														SUBJID	ARM
mg/kg	Type of observation results: Cageside observations	0	0	0	0	0	1	2	3	4	5	6	7	8	9		
Sex: Male		am	post dose 1h	post dose 2h	post dose 4h	post dose 30 min	am										
1	No abnormalities observed															001	Group C
2	No abnormalities observed															002	Group C
3	No abnormalities observed															003	Group C
4	No abnormalities observed															004	Group C
5	No abnormalities observed															005	Group C
6	No abnormalities observed															006	Group C
7	No abnormalities observed															007	Group C
8	No abnormalities observed															008	Group C
9	No abnormalities observed															009	Group C
10	No abnormalities observed															010	Group C

SUBJID与ARM为自行添加的辅助变量

# SEND程序的特点

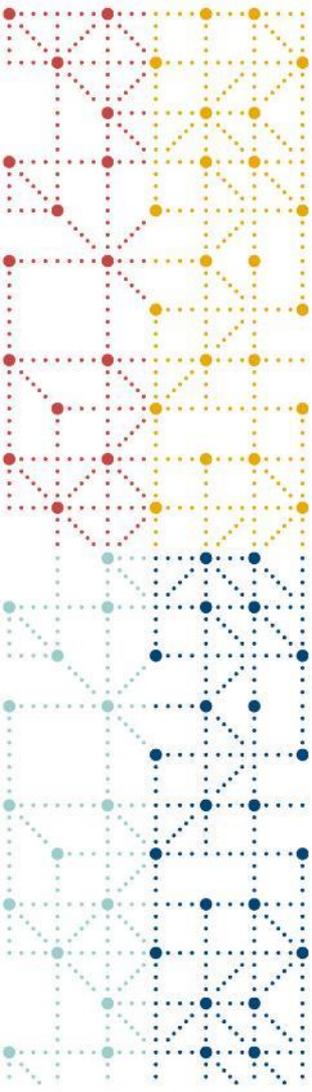
## 根据SEND项目经验总结

- 在编程阶段拿到的原始数据通常是稳定的数据，不随时间动态更新
  - 在编程时可使用if...then...语句来选取特定的数据
- 除了试验设计域之外，其他域仍有部分数据是来自报告或者方案
  - 在编程时，SEND要比SDTM的程序多很多直接赋值的语句
  - 要从报告或方案中获取数据的域：TA, TE, TX, TS, DM, SE, EX, DS, DD（可能）

# SEND程序的特点

## 根据SEND项目经验总结

- SEND程序的难点在于数据结构的整理，因为拿到的原始数据结构可能五花八门
  - 为了整理成SEND标准结构，与SDTM程序相比，需要用到更多转置，拼接和retain语句
- 原始数据收集较为零散时，会导致一个域的数据可能分布在几十个EXCEL sheet中
  - 曾遇到重复毒性的LB域，原始数据分布在84个sheet中，其数据结构有些相似有些则完全不同
  - 针对数据结构相似的sheet，程序中应使用宏，减少重复代码



**Thank You!**

**cdisc**