



# 研究数据列表模型

(Study Data Tabulation Model)

CDISC 递交数据标准团队 (SDS Team) 开发

主要贡献者: Wayne Kubick, Fred Wood, Diane Wold, Tom Guintier, Mary Lenzen, Julie Evans, Gary Walker 以及 CDISC SDS 团队

## 中文版读者说明

此版本为研究数据列表模型 (SDTM) V1.2。本版本增加了与人类临床试验和动物实验相关的一些变量、以及对正文的更正和澄清。详细的修订说明参见第 6.1 节。

中文初稿由 Absolute Clinical Data Systems Co., Ltd 翻译, C-STAR SDTM 团队负责校对和审核, 具体详见第 2 页。

## 中文版修订历史

日期	版本	修订小结
2015年2月1日	征求意见稿	在CDISC中国官网发布, 意见征集截止时间为2015年5月31日
2013年12月15日		发布的版本, 反映了在评论时期确定的所有改动和更正。
2013年5月5日		待评论草案。

## 英文版修订历史

日期	版本	修订小结
2008-11-12	V1.2 正式版	正式版本
2007-07-10	V1.2 草案	征求意见稿草案
2005-04-28	V1.1 正式版	正式版本, 集成了在公众审阅期间收到的意见微小修订。
2004-06-25	V1.0 正式版	第一次正式发布版本, 反映了在征求意见稿期间所有确认的修订。

注: 有关陈述和保证、责任限制和免责声明, 参见[附录 A](#)。

## CDISC 研究数据列表模型 (SDTM) V1.2 翻译和审核 流程及致谢

### 翻译

由 Absolute Clinical Data Systems Co., Ltd (AbsCD) 公司提供 V1.2 翻译初稿。

### 审核

中国 CDISC 协调委员会(C3C)于 2011 年 9 月成立 CDISC 数据标准翻译和审阅(C-STAR)项目组, 负责审阅和校对 AbsCD 提供的初稿。中文翻译稿基于英文 V1.2 版, 该英文版本于 2008-11-12 正式发布。

校对审核遵循 C-STAR 项目组批准的流程。任何部分内容都由两个独立的翻译校对。小组负责人、SDTM 审阅团队负责人和 C-STAR 项目组负责人分别依次进行核对和讨论决定。具体流程如下:

1. 第一位翻译校对员独立审阅, 并用 MS WORD Track Changes 记录
2. 第二位翻译校对员独立审阅, 也用 MS WORD Track Changes 记录
3. 其中一位翻译校对员负责合并, 校对并与另外一位讨论, 直至双方意见一致
4. 如果两个校对员意见不一致, 递交给 SDTM 审阅团队负责人讨论判定; 如果还不能定夺, 交由 C-STAR 项目组负责人会同 CDISC 专家评定
5. 所有各部分校对完成, 再进行合并后的全文一致性 QC 及和与 CDISC 其它标准(CDASH, ADaM) 的术语一致性 QC 检查
6. C-STAR 负责人进行最后审阅

### 翻译征求意见稿与终稿

经过 C-STAR 所有流程审阅后, 于 2013 年 5 月 5 日在 CDISC 网站上发布征求意见稿。2013 年 11 月 20 日, C-STAR 项目组根据收到的意见对翻译进行了修改, 并发布翻译终稿。

由于这是第一次尝试对 SDTM 进行中文翻译和核对, 不足之处在所难免。请将您的反馈意见经由 CDISC 发给我们, 以便在下一版中及时更正。

### 致谢!

非常感谢 AbsCD 和全体 SDTM IG 审核团队的努力, 使得 SDTM V1.2 中文稿终于面世。谢谢所有志愿者的奉献和 (审核时) 所属公司的支持!

C-STAR SDTM 翻译审阅组

## 目录

53			
54			
55	1	前言.....	4
56	1.1	目的.....	4
57	1.2	与之前 CDISC 模型的关系.....	4
58	1.3	与之前 SDTM 版本的重要变化.....	4
59	1.4	与 HL7 模型的关系.....	5
60	2	模型基础.....	6
61	2.1	模型概念和术语.....	6
62	2.2	通用观测数据类别 (General Observation Classes) .....	7
63	2.2.1	干预类 (Interventions Observation Class) .....	8
64	2.2.2	事件类 (Events Observation Class) .....	10
65	2.2.3	发现类 (Findings Observation Class) .....	11
66	2.2.3.1	事件或者干预的“相关发现” (Findings About) .....	14
67	2.2.4	用于所有类的标识符 (Identifiers For All Classes) .....	14
68	2.2.5	用于所有类的时间变量 (Timing Variables For All Classes) .....	15
69	2.2.6	临床的人口学数据域 (DM) .....	16
70	2.2.7	注释域 (CO) .....	17
71	2.2.8	受试者元素表 (SE) .....	18
72	2.2.9	受试者访视表 (SV) .....	19
73	3	试验设计模型 (Trial Design Model) .....	20
74	3.1	介绍.....	20
75	3.2	计划元素、分组和访视.....	20
76	3.2.1	试验元素表 (TE) .....	20
77	3.2.2	试验分组表 (TA) .....	20
78	3.2.3	试验访视 (TV) .....	21
79	3.3	试验入选/排除标准 (TI) .....	21
80	3.3.1	试验入选/排除表.....	21
81	3.4	试验概要信息 (TS) .....	22
82	3.4.1	试验概要表.....	22
83	4	数据集之间和记录之间的关系描述.....	23
84	4.1	关联记录数据集 (RELREC) .....	23
85	4.2	补充修饰语数据集 (SUPQUAL) .....	23
86	5	利用此模型进行数据监管递交.....	25
87	6	SDTM 版本历史 (暂未翻译) .....	26
88	6.1	从 SDTM V1.1 到 SDTM V1.2 的改动.....	26
89	7	附录 A 陈述和保证、责任限制和免责声明 (暂未翻译) .....	27
90			

# 1 前言

## 1.1 目的

本文档描述研究数据列表模型 (SDTM)，旨在为药物监管机构（如美国 FDA）的研究数据列表 (study data tabulation) 提供一个标准结构，作为产品递交申请的一部分。此文档是基于 CDISC 递交数据标准 (Submission Data Standards, SDS) 小组准备的材料创建，将取代所有之前版本，包括自前一版本 1.1 的诸多修改，详见[第 6.1 节](#)中的描述。

数据列表数据集是向 FDA 进行申报时递交的人类受试者病例报告列表数据 (Case Report Tabulation, CRT) 与相似动物数据的四种主要方式之一。CRT 也可以受试者概要 (patient profiles)、数据清单式列表 (data listing)、或分析数据集 (analysis datasets) 的形式进行申报递交。递交数据列表数据集对于企业的一个好处是：遵守该标准化结构可以最大程度地减少以不同格式进行递交同一数据的需要。

来自药物监管机构的新药评审员将从数据列表数据集中获益匪浅。通过针对评审员的培训，使其了解标准化数据集的原理和相关工具软件的使用，这样评审员用于数据准备的时间得以减少，使工作效率显著提高。另外一个好处是此标准数据集支持 FDA 为所有递交的临床试验数据建立统一的数据仓库所做出的努力，即使用一套统一的标准评审工具来读取、操作和查看相关数据。

此文档是为需要采集、准备、分析研究数据并提交给监管机构的公司和个人准备的。应用此模型的相关指导、说明与规则将由监管机构另行提供，建议读者在以 SDTM 为基础构建申报数据之前参考这些资料。

## 1.2 与之前 CDISC 模型的关系

如前所述，本文档基于我们之前熟知的版本，即 CDISC 递交数据标准和递交域模型 (Submission Domain Models) 而创建。自 SDTMV1.0 被指定为相关临床研究领域首个可实施版本以来，为支持更广泛的监管产品，包括满足非临床动物毒理研究的需求，在后续版本中，SDTM 已进行了大量的改进与增强。CDISC 将继续进一步评估该模型，以涵盖其它监管产品包括食物添加剂、治疗性生物制品、血液衍生制品、疫苗、细胞、(人或动植物) 组织、基因治疗和医疗器械；关于这些产品利用 SDTM 的结构化试行评估已在计划当中，评估结果将用来对标准进一步的优化与加强。有关此模型应用于各种数据的具体实施指南和受控术语集将分别单独加以发布。

## 1.3 与之前 SDTM 版本的重要变化

SDTM 设计时已考虑到向后兼容，即基于 V1.2 版创建的数据集可完全兼容之前的版本。在大多数情况下，这亦意味着在新版本中，通常只是增加新变量或更正文字错误，而不会删除之前版本中已有的变量或改变其现有结构。V1.2 版引入了以下新变量：--PRESPT, --VAMT, --VAMTU, --OBJ, --STRPT, --STPT, --ENRPT, --ENTPT, TIVERS, 和 TSGRPID。V1.2 版也包括大量的文字修正、说明和部分内容顺序的调整，包括受试者元素域与受试者访视域从[第3章](#)调整到[第2章](#)。

有关版本变化的更多细节，参见[第6.1节](#)。

#### 1.4 与 HL7 模型的关系

HL7 的读者应该注意到 SDTM 并不是作为在 HL7 框架内开发的一个 HL7 模型。SDTM 具体用来描述递报数据的内容，而不像 HL7，通常被用来规范实体组织、行为、角色及参与者等。CDISC 开发 SDTM 是为了制定一个向监管机构进行数据递交时的（统一）标准，该标准符合当前业界的通常实践以及监管者的要求，并特别指定利用 SAS V5 作为此类数据的传输格式。然而，CDISC 相信该模型的核心基本概念能与 HL7 V3 参考信息模型（RIM）相对应，并在将来能适应 HL7 的数据类型。

## 2 模型基础

### 2.1 模型概念和术语

SDTM为用户提供了一般性框架,用以描述来自于人类与动物研究所采集的数据及向监管部门递交的数据。该模型基于观测记录这一概念而建立,观测记录(observations)可由某一研究过程中采集到的各类信息组成。观测记录通常对应数据集中的行。某一特定主题下的一组观测记录被称为域(domain)。例如,“受试者101在研究的第六天开始出现轻度恶心症状”是一个临床研究中不良事件域的一个观测记录结果。

每一观测记录都可由一组命名变量来描述,每一变量(相当于数据集中的一列)都可以依据它的角色(Role)进行归类。角色描述了每一单独的观测记录中该变量要传达的信息类型以及该变量如何使用。SDTM变量主要有五种不同角色:

- 标识符(Identifier)变量:例如,用来标识研究本身、参与研究的受试者(人或动物个体)、域名及记录序号等。
- 主题(Topic)变量:指明观测记录的主要目的(例如,某一实验室检查的名称)。
- 时间(Timing)变量:描述观测记录发生的时间(例如,开始日期与结束日期)。
- 修饰语(Qualifier)变量:包括用来进一步描述结果的说明性文字或数值,或观测记录的更多特征(例如,单位或描述性形容词)。
- 规则(Rule)变量:在试验设计模型里,表达一种算法或可执行的方法,来定义其开始、结束或循环条件。

修饰语(Qualifier)变量可被进一步细分为五个类别:

- 分组修饰语(Grouping Qualifiers):对同一域中的数据进行分组。例如:--CAT和--SCAT。
- 结果修饰语(Result Qualifiers):在发现类域中,用来描述与主题变量相关的特定结果。用来回答主题变量所要表达的问题。结果修饰语包括--ORRES, --STRESC, 和--STRESN。
- 同义词修饰语(Synonym Qualifiers):指定了观测记录中某一特定变量的其它可用名称。例如,--MODIFY和--DECOD,与主题变量--TRT or --TERM是等同的术语;--TEST和--LOINC与--TESTCD是等同的。
- 记录修饰语(Record Qualifiers):从记录水平(而不是变量水平)定义某一观测记录的附加属性。例如--REASND, AESLIFE, 及在不良事件域(AE)中所有其他严重不良事件(SAE)标识变量;人口学数据域(DM)中年龄、性别和种族;发现类域中--BLFL, --POS, --LOC, --SPEC 与--NAM。
- 变量修饰语(Variable Qualifiers):用来进一步修饰或描述某一观测记录的特定变量,并且只有在它所修饰的变量中才有意义。例如--ORRESU, --ORNRHI与--ORNRL0是--ORRES的变量修饰语;而--DOSU则是--DOSE的变量修饰语。

例如,“受试者101在研究的第六天开始出现轻度恶心症状”,其主题变量值为该不良事件术语“恶心”。标识符变量值是该受试者编号“101”。时间变量值是该事件出现时研究已开始的天数,即“开始于研究第六天”。该事件严重程度可视为记录修饰语的例子,其值为“轻度”。其它时间与修饰语变量可视情况加入,以提供必要的细节来丰富一个观察记录的描述。

通常情况下，所有受试者的观测结果会采集在一系列不同的域中。域定义为一组具有共同主题并且逻辑上相关的观测结果的集合。其关系的内在逻辑可能涉及到数据的科学主题性质或者它在试验中的角色有关。每一个域采用唯一的，两个字符长的代码加以区别，该代码在整个数据递交过程中要保持始终一致。域代码储存在SDTM标准变量DOMAIN中，并可以有以下四种方式加以应用：作为该数据集的名称，该数据集中变量DOMAIN的值，该数据集中大部分变量的前缀，及在关系型数据集（relationship tables）作为变量RDOMAIN相应的值。

所有数据集均以平面文件（flat files）的形式存在，行代表观测记录，列代表变量名称。每一数据集通过元数据（metadata）定义来提供该数据集使用的所有变量信息。元数据将在一个名为“Define”的数据定义文档中进行描述，并与研究数据一起递交到监管机构（参见CDISC网站www.CDISC.org上有关病例报告表数据定义说明书[define.xml]）。Define.xml指定了七个不同的元数据属性来描述SDTM数据。

- 变量名（最多8个字符长度，为兼容SAS传输文件格式）。
- 描述性变量标签，最多40个字符长度，对于数据集中的每个变量都是唯一的。
- 数据类型（字符型或数值型）。
- 为数据值或数据呈现格式而规定的受控术语集（受控术语或者格式）
- 变量来源。
- 变量角色，决定了在数据集中如何使用该变量。对于V3.x域模型，角色用于表示变量的分类，如标识符变量、主题变量、时间变量、或五种类型的修饰语变量等。
- 注释（Comments）或其它申办者用来与药物监管机构就变量及其内容进行交流的相关必要信息。

这些变量中存储的数据既包括原始值（例如，最初采集的），也包括衍生值（例如，转化为标准单位，或由计算得来，如年龄）。SDTM描述了标准变量的名称、标签、角色和类型；来源和术语集由申办者根据每一个特定的试验自行定义。注意目前数据类型为了与SAS V5的传输文件格式兼容而被限制为字符型或数值型，将来更多的更具描述性的数据类型（例如：整型、浮点型、日期、日期/时间）有望随着SAS v5升级到新版本或者通过采纳不同的数据格式如XML而加以使用。

在准备数据递交时，申办者可以从标准递交数据集中去掉那些没有相关数据采集或未使用的变量（仅限于在SDTMIG实施指南中定义为可允许类别的变量）及其在define.xml中的相关描述，但不允许申办者自行添加变量，并且已存在的变量不可更名或修改为其它用途。申办者应参阅相应的实施指南（Implementation Guide），以了解在通用观测数据类别各域中，哪些变量是必需的（required）、哪些变量是期望的（expected）、以及哪些变量是许可的（permissible）。

## 2.2 通用观测数据类别（General Observation Classes）

临床研究中采集的大多数观测数据通常可分为干预（Interventions），事件（Events）与发现（Findings）三大类：

- 干预类：如表2.2.1中描述，获取受试者接受的研究性、治疗性或其他治疗（伴有实际或者期望生理效果），可以由研究方案指明的研究治疗（例如，暴露）、研究期间的伴随治疗（例如，伴随用药），或受试者自我服用的其它物质（如酒精、烟草或咖啡因）。
- 事件类：如表2.2.2中描述，获取研究方案中计划的重要事件，如随机化、研究完成、及在试验期间或试验之前发生的独立于临床研究计划评估的状况、事件等（前



者如不良事件，后者如疾病史)。

- 发现类：如表2.2.3中描述，获取临床研究中计划评估的观测结果，如特定的实验室检查、ECG检查或调查意问卷中问题的回答结果等。

基于上述通用观测类别的数据集拥有相同的标识符变量与时间变量。表2.2.4与表2.2.5分别对其进行了描述。作为一条基本法则，任何有效的标识符变量或时间变量均可在符合通用观测数据类别的递交数据集中使用。

表格中变量名前的两个连字符（例如--TRT），表明其前面须有两个字符长度的域代码前缀。使用域代码作前缀有助于最小化报表输出时对域进行合并/链接时变量无法区分的风险。

除了上述三种通用观测数据类别外，数据递交时还常包括一系列具有特殊标准结构的特殊目的的数据集，以呈现临床研究中其它重要的信息，有关例子包括：

- 临床研究与动物研究中的人口学特殊目的域，在第2.2.6节中描述。
- 其他特殊目的域如注释（第2.2.7节），受试者元素（第2.2.8节），受试者访视（第2.2.9节）。
- 描述试验设计的数据集，在第3章描述。
- 代表数据集间与记录间关系的数据集，在第4章描述。

SDTM支持临床与临床前研究数据的展示，两种数据拥有许多相同的变量，但也有部分变量是SDTMIG（人类临床试验研究数据表格模型实施指南）或者SEND（非临床研究数据交换标准）中独有的。在下述表格中，最后两列（如果适用）以符号“√”来分别表示是否适用于SDTMIG或者SEND。

## 2.2.1 干预类 (Interventions Observation Class)

表 2.2.1: 干预——主题和修饰语变量，每一个固定剂量期间或每一次干预事件，一条记录。

变量名称	变量标签	类型	角色	描述	SDTM IG	SEND
主题变量						
--TRT	治疗名称	字符型	主题	干预观测的主题，通常是该观测给药期间中所给予的治疗、药物或疗法的字面名称。	√	√
修饰语变量						
--MODIFY	治疗修正名称	字符型	--TRT 的同义修饰语	如果为了编码的目的修改了--TRT 的值，则修改后的名称放在这里。	√	√
--DECOD	治疗标准名称	字符型	--TRT 的同义修饰语	主题变量--TRT 或主题变量修正词(适用时) (--MODIFY) 的标准化或字典衍生名称，等于 WHO Drug 词典中的药物通用名，或 SNOMED、ICD9、其它发表或申办者定义词典中的术语。	√	√
--CAT	分类	字符型	分组修饰语	用于定义主题变量值的类别。	√	√
--SCAT	子类	字符型	分组修饰语	用于对--CAT 进一步分类。	√	√
--PRES	预设	字符型	记录修饰语	当特定干预是在 CRF 中事先规定时使用。值应为“是” (Y) 或空 (null)。	√	√



变量名称	变量标签	类型	角色	描述	SDTM IG	SEND
--OCCUR	是否发生	字符型	记录修饰语	用于记录预设干预是否发生,当特定干预是否发生的信息是引出时使用。	√	√
--STAT	完成状态	字符型	记录修饰语	用于表明未回答“预定干预是否发生”的问题。值应为空或“未做”(NOT DONE)。	√	√
--REASND	未做原因	字符型	记录修饰语	未做原因,当--STAT 的值是“未做”(NOT DONE) 时与--STAT 一起使用。	√	√
--INDC	适应症	字符型	记录修饰语	说明干预的适应症(例如某疗法使用或给予的原因)。	√	√
--CLAS	类别	字符型	--TRT 的变量修饰语	药物或治疗的类别,通常来源于编码词典。	√	√
--CLASCD	类别编码	字符型	--TRT 的变量修饰语	用来代表--CLAS 的编码。	√	√
--DOSE	剂量	数值型	记录修饰语	--TRT 给予的量。--DOSTXT 有值时则不赋值。	√	√
--DOSTXT	剂量描述	字符型	记录修饰语	以文本格式采集的给药信息。例如: <1 每天, 200-400。--DOSE 有值时则不赋值。	√	√
--DOSU	剂量单位	字符型	--DOSE 的变量修饰语	--DOSE, --DOSTOT 或--DOSTXT 的单位。例如: ng, mg, mg/kg。	√	√
--DOSFRM	剂型	字符型	--DOSE 的变量修饰语	治疗的剂型,例如: 片剂、胶囊。	√	√
--DOSFRQ	期间给药频率	字符型	--DOSE 的变量修饰语	通常表示为特定期间的给药次数,例如 QD, BID, TID, QID。	√	√
--DOSTOT	每日总剂量	数值型	记录修饰语	--TRT 的每日总剂量,单位是--DOSU 的值。当采集的是每日总剂量时使用。	√	√
--DOSRGM	计划给药方案	字符型	--DOSE 的变量修饰语	干预的(计划)治疗方案的文字描述。例如给药两周、间歇两周。	√	√
--ROUTE	给药途径	字符型	--TRT 的变量修饰语	干预的给予途径。例如: 口服、静脉注射。	√	√
--LOT	批号	字符型	记录修饰语	--TRT 所述干预的批号。	√	√
--LOC	给药部位	字符型	记录修饰语	干预的解剖学位置,如注射部位。例如: “右上臂”注射。	√	√
--TRTV	溶媒	字符型	记录修饰语	实施治疗的赋形剂,如用于溶解药物的液体。例如: 盐水。	√	√
--VAMT	溶媒剂量	数值型	记录修饰语	治疗赋形剂使用或者给予的数量	√	√
--VAMTU	溶媒剂量单位	字符型	--VAMT 的变量修饰语	治疗赋形剂的单位。如: mL, mg。	√	√
--ADJ	剂量调整原因	字符型	记录修饰语	描述或者解释剂量调整的原因。例如: 不良事件、反应不足、非医学原因。	√	√

238

## 2.2.2 事件类 (Events Observation Class)

239 表 2.2.2: 事件——主题和修饰语变量, 每一个事件, 一条记录。

变量名称	变量标签	类型	角色	描述	SDTM IG	SEND
主题变量						
--TERM	报告词	字符型	主题	事件观测的主题变量, 是事件的字面或预定名字。	√	√
修饰语变量						
--MODIFY	修正报告词	字符型	--TERM 的同义修饰语	如果为了编码的目的修改了--TERM 的值, 则修改后的文字放在这里。	√	√
--DECOD	词典衍生词	字符型	--TERM 同义修饰语	主题变量--TERM 或主题变量修正词(--MODIFY) 的字典或申办者定义的文字描述。等于首选词 (MedDRA 中的 PT)	√	√
--CAT	分类	字符型	分组修饰语	用于定义主题变量值的类别。	√	√
--SCAT	子类	字符型	分组修饰语	用于对--CAT 进一步分类。	√	√
--PRESF	预设	字符型	记录修饰语	当--TERM 所述事件是在 CRF 中事先规定时使用。如果是预设事件, 则其值应为“是” (Y), 如果是自发报告事件则为空。	√	√
--OCCUR	是否发生	字符型	记录修饰语	用于记录预设事件是否发生, 当预设事件是否发生信息是引出时使用。	√	√
--STAT	完成状态	字符型	记录修饰语	用于表明未回答“预设事件是否发生”的问题。值应为空或“未做” (NOT DONE)。	√	√
--REASND	未做原因	字符型	记录修饰语	未做原因。当--STAT 的值是“未做” (NOT DONE) 时和--STAT 一起使用。	√	√
--BODSYS	身体系统或器官分类	字符型	记录修饰语	事件的来自标准层级体系 (如 MedDRA) 的相关系统或系统器官分类。例如: 消化系统失调。	√	√
--LOC	事件的部位	字符型	记录修饰语	描述事件相关的解剖学位置。例如: “左上臂” 皮疹。	√	√
--SEV	严重程度/强度	字符型	记录修饰语	事件的严重程度或强度。例如: 轻度, 中度, 重度。	√	√
--SER	严重事件	字符型	记录修饰语	是否为严重事件? 值为“是” (Y) 或“否” (N)。	√	√
--ACN	对研究治疗采取的措施	字符型	记录修饰语	描述因事件而对研究治疗的改变。例如: 增加剂量、剂量未改变。	√	√
--ACNOTH	采取的其它措施	字符型	记录修饰语	描述历事件而采取的研究药物剂量调整无关的措施。	√	√
--REL	因果关系	字符型	记录修饰语	记录研究者对于事件和治疗之间关系的观点。ICH E2A 和 E2B 的例子包括: 无关, 可能无关, 可能有关, 有关。	√	√

变量名称	变量标签	类型	角色	描述	SDTM IG	SEND
--RELNST	与非研究治疗的关系	字符型	记录修饰语	记录研究者对事件是否由研究药物之外的其它治疗引起的意见。例如：“更可能与阿司匹林的使用有关”。	√	√
--PATT	事件模式	字符型	记录修饰语	用来表明事件发生的时间模式。例如：间歇的，连续的，单次事件。	√	√
--OUT	事件结果	字符型	记录修饰语	描述事件的结果。例如：已恢复/已痊愈，致命。	√	√
--SCAN	涉及癌症	字符型	记录修饰语	事件是否和癌症发展有关？有效值包括“是”（Y）和“否”（N）。	√	√
--SCONG	先天性畸形或出生缺陷	字符型	记录修饰语	事件是否与先天性异常或出生缺陷有关？有效值包括“是”（Y）和“否”（N）。	√	√
--SDISAB	持续或明显的残疾或失能	字符型	记录修饰语	事件是否导致永久或显著的残疾/失能，有效值可为“是”（Y）和“否”（N）。	√	√
--SDTH	导致死亡	字符型	记录修饰语	事件是否导致死亡？有效值包括“是”（Y）和“否”（N）。	√	√
--SHOSP	需要或延长住院	字符型	记录修饰语	事件是否导致需要或延长了住院时间？有效值包括“是”（Y）和“否”（N）。	√	√
--SLIFE	危及生命	字符型	记录修饰语	事件是否危及生命？有效值包括“是”（Y）和“否”（N）。	√	√
--SOD	药物过量引发	字符型	记录修饰语	事件是否由过量引发？有效值包括“是”（Y）和“否”（N）。	√	√
--SMIE	其它医学上重要的严重事件	字符型	记录修饰语	是否属其它种类的严重事件？有效值包括“是”（Y）和“否”（N）。	√	√
--CONTRT	是否给予伴随或额外的治疗	字符型	记录修饰语	是否由于事件的发生给予了另外一种治疗？有效值包括“是”（Y）和“否”（N）。	√	√
--TOX	毒性	字符型	--TOXGR 的变量修饰语	--TOXGR 所定量修饰的毒性的描述，如 NCI CTCAE 短名称。例如：高钙血症，低钙血症。申办者应在 Define.xml 文档的申办者备注列注册使用的评分标准和版本。	√	√
--TOXGR	毒性等级	字符型	记录修饰语	使用标准毒性尺度（例如 NCI CTCAE）记录的毒性等级。申办者应在 define.xml 文档的申办者备注列中注明使用的评分标准和版本。	√	√

240

## 2.2.3 发现类 (Findings Observation Class)

241

表 2.2.3：发现——主题和修饰语变量， 每一个发现， 一条记录

变量名称	变量标签	类型	角色	描述	SDTM IG	SEND
<b>主题变量</b>						
--TESTCD	测量、检验或检查简称	字符型	主题	--TEST 的缩写值，当将一个数据集从垂直格式转为水平格式的时候作为列名，此缩写最多 8 个字符型。例如：PLATELET，SYSBP，RRMIN，EYEEXAM。	√	√
<b>修饰语变量</b>						
--TEST	测量、检验或检查名称	字符型	--TESTCD 的同义修饰语	--TESTCD 的长名称。例如，血小板计数，收缩压，RR 间期综合（最小值），眼部检查。	√	√
--MODIFY	修正词	字符型	--ORRES 的同义修饰语	如果因为编码的目的修改了--ORRES 的值，则修改后的文本放在这里。	√	√
--CAT	分类	字符型	分组修饰语	用于定义主题变量值的类别，例如：血液学、尿液分析、化学、汉密尔顿抑郁量表、SF36。	√	√
--SCAT	子类	字符型	分组修饰语	用于定义对--CAT 的进一步分类，例如：白细胞分化（DIFFERENTIAL）。	√	√
--POS	观测时受试者体位	字符型	记录修饰语	研究受试者在测量或检查时的体位，例如：仰卧位，立位，坐位。	√	√
--BODSYS	身体系统或器官分类	字符型	记录修饰语	发现涉及的系统或器官类别，来源于一个标准等级分类，如 MedDRA 的 SOC。	√	√
--ORRES	结果/发现-原始单位	字符型	结果修饰语	收到或采集的测量的原始结果。如：120，<1，POS	√	√
--ORRESU	原始单位	字符型	--ORRES 的变量修饰语	--ORRES 的单位，例如，IN、LB、kg/L。	√	√
--ORNRLO	正常范围下限-原始单位	字符型	--ORRES 的变量修饰语	--ORRES 结果对应的正常范围或参考范围的下限。	√	√
--ORNRHI	正常范围上限-原始单位	字符型	--ORRES 的变量修饰语	--ORRES 结果对应的正常范围或参考范围的上限。	√	√
--STRESC	结果/发现-标准格式	字符型	结果修饰语	包含所有发现的结果，从--ORRES 拷贝或衍生为标准格式或标准单位。--STRESC 字符型应存储所有字符型的结果或发现。若结果是数值型，应同时以数值型存于--STRESN。例如：如果在多个测试中--ORRES 结果是 'NONE'、'NEG' 和 'NEGATIVE'，而且这些值实际有相同的含义，那么它们可在--STRESC 以标准格式表示为 "NEGATIVE"。	√	√
--STRESN	数值型结果/发现-标准单位	数值型	结果修饰语	用于标准格式的连续性或数值型结果或发现，以数值格式由--STRESC 拷贝而来。--STRESN 应存储所有数值型的结果或发现。	√	√
--STRESU	标准单位	字符型	--STRESC 和--STRESN 的变量修饰	用于--STRESC 或--STRESN 的标准单位，例如：mol/L。	√	√

变量名称	变量标签	类型	角色	描述	SDTM IG	SEND
--STNRLO	正常范围下限- 标准单位	数值型	--STRESC 和 --STRESN 的 变量修饰语	标准化结果的正常范围或参考范围的下限 (如: --STRESC, --STRESN), 以标准单位 (--STRESU) 呈现。	√	√
--STNRHI	正常范围上限- 标准单位	数值型	--STRESC 和 --STRESN 的 变量修饰语	标准化结果的正常范围或参考范围的上限 (如: --STRESC, --STRESN), 以标准单 位 (--STRESU) 呈现。	√	√
--STNRC	文本结果的正 常值范围	字符型	--STRESC 的 变量修饰语	存于--STRESC 中的顺序或分类的字符型尺度 的正常范围和参考范围字符型。例如: 由阴 性到微量。	√	√
--NRIND	正常值/参考值 范围标志	字符型	--ORRES 的变 量修饰语	用于标明值是否超出正常范围或参考范围。 可由--ORNRL0 和--ORNRI 或其他客观评判 标准定义。例如: 是 (Y), 否 (N); 高, 低; 正常; 异常。	√	√
--RESCAT	结果分类	字符型	--ORRES 的变 量修饰语	用于分类发现的结果。例如: 用恶性或良性 来描述肿瘤发现。	√	√
--STAT	完成状态	字符型	记录修饰语	用于表明一个问题未问及或一个检查未做, 或者检查没有得到结果。应为空 (null) 或 “未做” (NOT DONE)。	√	√
--REASND	未做原因	字符型	记录修饰语	未做原因。当--STAT 的值是“未做” (NOT DONE) 时和--STAT 一起使用。	√	√
--XFN	外部文件名	字符型	记录修饰语	外部文件名, 比如一个 ECG 波形或医学影像 文件。	√	√
--NAM	实验室/供应商 名称	字符型	记录修饰语	提供检测结果的供应方 (例如实验室) 名称 或编号。	√	√
--LOINC	LOINC 编码	字符型	--TESTCD 同 义修饰语	主题变量如实验室检测名称的 LOINC 代码。	√	√
--SPEC	标本的材料类 型	字符型	记录修饰语	定义检测样本的类型, 例如: 血清、血浆、 尿液。	√	√
--SPCCND	标本状况	字符型	记录修饰语	定义样本的性状, 例如: 浑浊。	√	√
--LOC	测量部位	字符型	记录修饰语	检测采集相关的位置, 例如: “直肠”温度, “左臂”血压和“V1”导连的心电图。	√	√
--METHOD	检验或检查方 法	字符型	记录修饰语	检测方法, 例如: 酶免疫测定 (EIA), 电泳 法, 浸渍法。	√	√
--BLFL	基线标识	字符型	记录修饰语	基线值标识符, 应为是 (Y) 或空 (null)。	√	√
--FAST	是否空腹	字符型	记录修饰语	用于标识空腹状态的标识符, 允许值包括是 (Y), 否 (N), 未知 (U) 或如无关则为空 (null)。	√	√
--DRVFL	衍生标识	字符型	记录修饰语	用于标识衍生记录 (例如: 为其它记录平均 值的记录, 如计算出来的基线值), 变量值 应为是 (Y) 或空 (null)。	√	√

变量名称	变量标签	类型	角色	描述	SDTM IG	SEND
--EVAL	评估者	字符型	记录修饰语	评估者的角色。仅用于结果为主观时（例如，由某人或某团体确定）。评估者举例：裁决委员会（ADJUDICATION COMMITTEE），供应者（VENDOR）。	√	√
--TOX	毒性	字符型	--TOXGR 的变量修饰语	--TOXGR 所定量修饰的毒性的描述，如 NCI CTCAE 短名称。例如：高钙血症，低钙血症。申办者应在 Define.xml 文档的申办者备注列注明使用的尺度和版本。	√	√
--TOXGR	毒性等级	字符型	--STRERC 的变量修饰语	使用标准毒性尺度记录的毒性等级（例如：NCI CTCAE 等级）。申办者应该在的 Define.xml 的申办者备注列注明使用的尺度和其版本。例如：2。	√	√
--SEV	严重程度	字符型	记录修饰语	发现的严重程度或强度，例如：轻度、中度、重度。	√	√
--DTHREL	与死亡关系	字符型	记录修饰语	描述一个特定发现与受试者的死亡之间的关系。		√
--LLOQ	定量下限	数值型	--STRERN 的变量修饰语	说明定量检测方法的检测下限 <sup>1</sup> 。与--STRESU 的单位相同。	√	√

### 2.2.3.1 事件或者干预的“相关发现” (Findings About)

事件或者干预相关的发现，利用发现类别变量，增加--OBJ变量，如下表中所述。注意，--OBJ变量仅用于事件或者干预相关发现。

表 2.2.3.1: 相关发现，增加修饰语变量

变量名称	变量标签	类型	角色	描述	SDTMIG	SEND
--OBJ	观测对象	字符型	记录修饰语	在事件相关发现或者干预相关发现模型中使用。描述在--TESTCD/--TEST 测量的事件或者干涉的性质。例如：呕吐事件有其发现，则--OBJ = “呕吐”，呕吐量在--TESTCD = “量”中记录。	√	

### 2.2.4 用于所有类的标识符 (Identifiers For All Classes)

以下所有标识符变量可用于任一通用观测类的任何域。在基于三个通用观测类之一的所有域内STUDYID、DOMAIN、USUBJID和--SEQ均为必需变量。

所有标识符变量均可用于SDTMIG和SEND。

表 2.2.4: 所有观测类——标识符

变量名称	变量标签	类型	描述
------	------	----	----

<sup>1</sup>【译者注】原文为the lower limit of quantitation for an assay，也可译为检测限或检测灵敏度。

变量名称	变量标签	类型	描述
STUDYID	研究标识符	字符型	研究的唯一标识符。
DOMAIN	域名缩写	字符型	与观测最相关的域名，两个字符缩写。域名缩写也用作变量的前缀以确保在数据集合并时的唯一性。
USUBJID	受试者唯一标识符	字符型	申报产品相关的所有申报或提交的全部研究中唯一标识一个受试者的标识符。
--SEQ	序号	数值型	序号，确保数据集内一个受试者（或参数，例如试验概要域）的记录的唯一性。可以是任何有效数字（包括小数），不必从 1 开始。应该按照连续的时间次序排列。
--GRPID	分组 ID	字符型	可选的组别标识符，用来关联域内同一个受试者的一组相关记录。--GRPID 也用于关联试验概要域中的一组相关记录（ <a href="#">第 3.4 节</a> ）。
--REFID	参考 ID	字符型	可选的内部或外部标识符，例如实验室样本 ID 或心电图波形或医学影像的 UUID。
--SPID	申办者 ID	字符型	申办者定义的标识符。如预印在 CRF 的伴随用药页上的行标识符。

## 2.2.5 用于所有类的时间变量 (Timing Variables For All Classes)

以下所有时间变量可用于除实施指南标准一域一模型假设中限制外的任一通用观测类的任何域。

所有时间变量均可用于SDTMIG或SEND。

表 2.2.5: 所有观测类——时间变量

变量名称	变量标签	类型	描述
VISITNUM	访视编号	数值型	临床会见码。变量 VISIT 的数值型版，用来排序。
VISIT	访视名称	字符型	方案定义的临床会见的描述。
VISITDY	访视的计划研究日	数值型	计划的访视研究日。应为整数。
TAETORD	治疗组内元素的计划顺序	数值型	表示组内元素的计划顺序的数字（参见 <a href="#">第 3.2.2 节</a> ）。
EPOCH	时期	字符型	与观察的开始日期/时间相关的时期，若开始的日期/时间未采集，则为采集的日期/时间（参见 <a href="#">第 3.2.2 节</a> ）。
--DTC	采集日期/时间	字符型	观测的采集日期和时间，以 ISO 8601 字符格式表示。
--STDTC	观测开始日期/时间	字符型	观测的开始日期/时间，以 ISO 8601 字符格式表示。
--ENDTC	观测结束日期/时间	字符型	观测的结束日期/时间，以 ISO 8601 字符格式表示。
--DY	访视/采集/检查的研究日	数值型	访视/采集/检查的实际研究日，表示为相对于人口领域中申办者定义的 RFSTDTC 的整天数。
--STDY	观测开始研究日	数值型	观测开始的实际研究日，表示为相对于人口统计学领域中申办者定义的 RFSTDTC 的整天数。
--ENDY	观测结束研究日	数值型	观测结束的实际研究日，表示为相对于人口统计学领域中申办者定义的 RFSTDTC 的整天数。
--DUR	持续时间	字符型	采集事件、干预或发现的持续时间，以 ISO 8601 字符格式表示。仅用于 CRF 采集而非衍生的信息。
--TPT	计划时点名称	字符型	方案中定义的观测时间的文字描述。可表示为距离某一固定参照点的时间，如上次用药时间。参见--数值型 TPTNUM 和--TPTREF。



变量名称	变量标签	类型	描述
--TPTNUM	计划时点编号	数值型	计划时间点的数值型版，用以排序。
--ELTM	距离参照时点的计划时间	字符型	距离计划的固定参照(--TPTRF)的 ISO 8601 字符型格式的计划时间，如“前一剂”或“前一次进餐”。这个变量在重复测量的时候很有用。不是钟表时间或日期/时间变量，为一个区间，以 ISO 持续时间表示。
--TPTRF	参照时点	字符型	--ELTM， --数值型 TPTNUM 和 --TPT 变量涉及的固定参照时间点的描述。例如：“前一剂”或“前一次进餐”。
--RFTDTC	参照时点的日期/时间	字符型	--TPTRF 定义的以 ISO 8601 字符格式表示固定数值型的参照时间点的日期/时间。
--STRF	开始时间与参照区间对应关系	字符型	标识观测的开始是在申办者定义的参考周期的“之前”、“之中”或“之后”。申办者定义的参照周期是一个连续的时间段，由人口统计学域中不相连的一个起始点 RFSTDTC 和一个结束点 RFENDTC 确定。
--ENRF	结束时间与参照区间对应关系	字符型	标识观测的结束是在申办者定义的参考周期的“之前”、“之中”或“之后”。申办者定义的参照周期是一个连续的时间段，由人口统计学域中不相连的一个起始点 RFSTDTC 和一个结束点 RFENDTC 确定。
--EVLINT	评估时间区间	字符型	与发现变量如--TESTCD 相关的评估期间，以 ISO 8601 字符格式表示。例如：-P2M 表示问卷如 SF-36 评估的期间是过去 2 个月的周期。
--STRTPT	开始时间与参照时点对应关系	字符型	标识观测开始是在申办者定义参照时间点之前或之后，以变量--STTPT 定义。
--STTPT	开始时间的参照时点	字符型	--STRTPT 代表申办者定义参照点的描述或 ISO 8601 字符格式的日期/时间。例如：“2003-12-15”或者“VISIT 1”。
--ENRTPT	结束时间与参照时点对应关系	字符型	标识观测结束是在申办者定义参照时间点之前或之后，以变量--ENTPT 定义。
--ENTPT	结束时间的参照时点	字符型	--ENRTPT 代表申办者定义参照点的描述或 ISO 8601 字符格式的日期/时间。例如：“2003-12-15”或者“VISIT 2”。

## 2.2.6 临床的人口学数据域 (DM)

每个研究都必须包含一组有特定结构的标准观察数据，即人口学数据域（见表2.2.6）。人口学数据域是受试者所有其它观察数据的父域，其域名必须为“DM”。人口学数据域描述了研究受试者的基本特征，用于审阅者选择分析人群。人口学数据域和其它数据集一样，包含标识符变量、一个主题变量、时间变量以及修饰词变量。由于DM具有固定结构，只有以下时间变量允许并且可根据需要加入：VISITNUM、VISIT和VISITDY。记录修饰语变量DMXFN（外部文件名）是唯一可增添的附加变量，该变量源于发现类，也可用于指向一个外部文件，例如受试者的详细记述（narratives）。

表 2.2.6: 受试者的人口学数据域变量

变量名称	变量标签	类型	描述
标识符变量			
STUDYID	研究标识符	字符型	研究的唯一标识符。
DOMAIN	域名缩写	字符型	两个字符的域名缩写，必须是 DM。
USUBJID	受试者唯一标识符	字符型	同一产品的所有申报或递交的研究中，受试者的唯一标识符。

变量名称	变量标签	类型	描述
<b>主题变量</b>			
SUBJID	本研究受试者标识符	字符型	受试者标志，在研究中必须是唯一的，通常为采集在 CRF 中的为受试者 ID。
<b>修饰语变量</b>			
RFSTDTC	受试者参照开始日期/时间	字符型	受试者参照开始日期/时间，以 ISO 8601 格式表示。通常等于受试者第一次暴露于研究治疗的日期/时间。要求所有随机受试者具有。如果受试者没有符合日期要求的里程碑，如筛选失败或者受试者未分组，则值为空。
RFENDTC	受试者参照结束日期/时间	字符型	受试者参照结束日期/时间，以 ISO 8601 格式表示。通常等于受试者结束试验的日期/时间，也常等于受试者最后一次暴露于研究治疗的日期/时间。要求所有随机受试者具有。若受试者筛选失败或者未分组，则值为空。
SITEID	研究中心标识符	字符型	研究中研究中心的唯一标识符。
INVID	研究者标识符	字符型	研究中研究者的标志，可以用来补充 SITEID；若 SITEID 等于 INVID 则不需要此变量。
INVNAM	研究者姓名	字符型	研究中心的研究者姓名。
BRTHDTC	出生日期/时间	字符型	受试者的出生日期/时间，以 ISO 8601 格式表示。
AGE	年龄	数值型	以 AGEU 单位表示的年龄。可由 (RFSTDTC-BRTHDTC) 衍生得到，但 BRTHDTC 不是在所有情况下都可获得（出于受试者隐私的考虑）。
AGEU	年龄单位	字符型	年龄的单位。
SEX	性别	字符型	受试者性别。
RACE	人种	字符型	受试者的人种，申办方应参照 “Collection of Race and Ethnicity Data in Clinical Trials” (FDA, September 2005) ” 来采集人种 ( <a href="http://www.fda.gov/cder/guidance/5656fml.htm">http://www.fda.gov/cder/guidance/5656fml.htm</a> )
ETHNIC	族裔	字符型	受试者的族裔。申办方应参照 “Collection of Race and Ethnicity Data in Clinical Trials” (FDA, September 2005) 来采集族裔。 ( <a href="http://www.fda.gov/cder/guidance/5656fml.htm">http://www.fda.gov/cder/guidance/5656fml.htm</a> ) 。
ARMCD	计划分组编码	字符型	最长 20 个字符，并且不能含有特殊字符。
ARM	计划分组描述	字符型	受试者被分配到的分组名称。
COUNTRY	国家	字符型	受试者参加试验的研究中心所在的国家，以 ISO 3166 三字符格式表示。
<b>时间变量</b>			
DMDTC	采集日期/时间	字符型	采集人口统计信息的日期/时间，以 ISO 8601 格式表示。
DMDY	采集的研究日	数值型	采集数据的研究日，以整数天为单位度量。算法必须参照申办方定义的 RFSTDTC 变量。

## 2.2.7 注释域 (C0)

很多研究采集注释信息。这些信息通常由一个主要研究者提供，但也可以由其他人提供，如中心审阅者。采集到的注释应在一个单独的注释域中递交，见表2.2.7。

有关额外的标识符变量和时间变量的使用，参见实施指南。

**表 2.2.7: 注释域变量**

变量名称	变量标签	类型	角色	描述
STUDYID	研究标识符	字符型	标识符	研究的唯一标识符。
DOMAIN	域名缩写	字符型	标识符	两个字符的域名缩写，必须是 C0。
RDOMAIN	关联域名缩写	字符型	记录修饰符	父记录所在域的域名缩写。若为一般性注释或未指明与哪个域相关，则值为空 (Null)。
USUBJID	受试者唯一标识符	字符型	标识符	同一产品的所有申报或递交的研究中，受试者的唯一标识符。
COSEQ	序号	数值型	标识符	序号，用于确保记录在域中的唯一性。应该按照一致的时间顺序。
IDVAR	标识变量	字符型	记录修饰语	父表中用于指明注释与哪个记录相关联的变量。例如 AESEQ 或 CMGRPID。仅在注释与个别记录相关时使用。若注释与整个域相关，即未指明哪个记录，则值为空 (Null)。
IDVARVAL	标识变量值	字符型	记录修饰语	父记录中标识变量的值。若注释与整个域相关，未指明哪个记录，则值为空 (Null)。
COREF	注释参考	字符型	记录修饰语	申办方指定的与注释相关的参考，可以是 CRF 页码（例如 650），或是一个模块的名字（例如 DEMOG），或是用于识别参考的信息的组合（例如 650-VITALS-VISIT 2）。
COVAL	注释	字符型	主题	注释的内容。超出 200 个字符的内容可记录在额外的列中 COVAL1-COVALn。
COEVAL	评估者	字符型	记录修饰语	用来描述注释的初创作者，例如：中心审阅者 (CENTRAL REVIEWER)。
CODTC	注释日期/时间	字符型	时间	如果采集了注释页，该变量表示专门注释表上的注释日期/时间，以 ISO 8601 格式表示。如果这是一个其它的域的子记录或注释日期没有采集，则值为空 (null)。

## 2.2.8 受试者元素表 (SE)

受试者元素表描述贯穿受试者的元素的实际顺序，以及每个元素开始日期/时间和结束日期/时间。此节对应于试验设计模型中试验元素（[第3.2.1节](#)）中的所讲的计划元素。因为实际数据和计划不一定一致，此模型允许添加受试者非计划元素的描述。

有关额外的标识符变量和时间变量的使用，参见实施指南。

**表 2.2.8 受试者元素—每一个受试者，每一个实际元素，一条记录**

变量名称	变量标签	类型	描述
STUDYID	研究标识符	字符型	研究的唯一标识符。
DOMAIN	域名缩写	字符型	两个字符的域名缩写，必须是 SE。
USUBJID	受试者唯一标识符	字符型	同一产品的所有申报或递交的研究中，受试者的唯一标识符。
SESEQ	序号	数值型	序号，用于确保记录在数据集内的唯一性。应该按照一致的时间顺序。
<b>主题变量</b>			
ETCD	元素编码	字符型	ETCD (ELEMENT 伴随变量) 限定于 8 个字符，并且无特殊字符限制。为了方便编程目的其值应短小，但是不期望 ETCD 必须做为一个变量名应用。

变量名称	变量标签	类型	描述
<b>修饰语变量</b>			
ELEMENT	元素描述	字符型	元素的名称。如果 ETCD 为“TCD 称。如果时, ELEMENT 应该为空(null)。
<b>时间变量</b>			
SESTDTC	元素开始日期/时间	字符型	以 ISO 8601 格式表示, 每个受试者的元素开始日期/时间
SEENDTC	元素结束日期/时间	字符型	以 ISO 8601 格式表示, 每个受试者的元素结束日期/时间
TAETORD	组内元素计划顺序	数值型	受试者所被分配组内元素计划序号
EPOCH	时期	字符型	与受试者所被分配组内的计划元素序列相关联元素的时期
<b>修饰语变量</b>			
SEUPDES	计划外元素描述	字符型	对于在非计划的元素内发生在受试者上的事件描述。只在 ETCD 为“UNPLAN”时使用。

### 2.2.9 受试者访视表 (SV)

受试者访视表描述受试者每个访视实际开始和结束日期/时间。此节对应于试验设计模型中试验元素(第 3.2.3 节)中描述的计划访视。因为实际数据和计划并不总是一致, 此模型允许添加受试者非计划访视的描述。

有关额外的标识符变量和时间变量的使用, 参见实施指南。

**表 2.2.9 受试者访视—每一个受试者, 每一个访视, 一条记录**

变量名称	变量标签	类型	描述
STUDYID	研究标识符	字符型	研究的唯一标识符。
DOMAIN	域名缩写	字符型	两个字符的域名缩写, 必须是 SV。
USUBJID	受试者唯一标识符	字符型	同一产品的所有申报或递交的研究中, 受试者的唯一标识符。
<b>主题变量</b>			
VISITNUM	访视编号	数值型	临床会见号码。(小数可用于插入非计划的访视) VISIT 的数字版本, 可用于排序。
<b>时间变量</b>			
VISIT	访视名称	字符型	试验方案中定义的临床会见描述或者对非计划的访视描述。可作为除 VISITNUM 和/或 VISITDY 之外关于临床会见的文字描述。
VISITDY	访视的计划研究日	数值型	根据人口学域中 RFSTDTC 变量计算的访视开始计划研究日。
SVSTDTC	访视开始日期/时间	字符型	以 ISO 8601 格式表示的受试者访视开始日期/时间
SVENDTC	访视结束日期/时间	字符型	以 ISO 8601 格式表示的受试者访视结束日期/时间
SVSTDY	访视开始的研究日	数值型	访视开始日期, 相对申办者定义 RFSTDTC 的研究日
SVENDY	访视结束的研究日	数值型	访视结束日期, 相对申办者定义 RFSTDTC 的研究日
SVUPDES	计划外访视描述	字符型	非计划访视内发生的受试者事件描述。若为方案中的计划访视, 则值为空 (Null)。

## 3 试验设计模型 (Trial Design Model)

### 3.1 介绍

试验设计模型定义了一个标准结构来表示试验计划的事件序列和治疗方案。此模型提供了一个标准方法，以定义受试者将要经历的治疗分组、计划访视和评估。

模型构建在元素、分组、时期和访视等概念上。与这些概念相对应的变量会在许多SDTM域中用到。相关实施指南定义了试验设计的具体细节和实例。

### 3.2 计划元素、分组和访视

在此模型下，计划的信息呈现在以下的三张表中：

- 试验元素表 (表 3.2.1) 描述了元素的代码（每个元素是唯一的）、元素描述以及一个元素开始和结束规则。规则可以描述为伪代码或者能决定从一个元素过渡到另一个元素的可执行代码。
- 试验分组表 (表 3.2.2) 描述了试验中每个计划分组，分组可以描述为一列有序元素，相同元素可以在一个分组中发生不止一次。为了包容试验设计的复杂性，在条件许可时该表允许从一个元素到另一个元素的分叉规则，也允许受试者跳过下一个元素而不必顺次向前的规则。
- 试验访视表 (表 3.2.3) 描述了试验中每一分组里访视的计划顺序和编号。当每个分组中访视不同时，设置每一个访视一个单独记录。它描述试验中访视 (VISIT)、访视编号 (VISITNUM) 和计划访视的研究日 (VISITDY) 的允许或计划值（这些值在后来采集的研究数据中作为时间变量），以及每次访视的开始和结束规则。在大多数设盲试验中，访视时间对于所有分组的所有受试者均相同。

这些数据集提供了对于不同研究是否具有可比性的重要信息。

#### 3.2.1 试验元素表 (TE)

表 3.2.1: 试验元素——每一个试验元素，一条记录

变量名称	变量标签	类型	描述
STUDYID	研究标识符	字符型	研究的唯一标识符。
DOMAIN	域名缩写	字符型	两个字符的域名缩写，必须是TE。
ETCD	元素编码	字符型	ETCD (ELEMENT伴随变量) 限定于8个字符并且无特殊字符限制。为了方便编程目的其值应该短小，但是不期望ETCD必须做为一个变量名应用。
ELEMENT	元素描述	字符型	元素的名称。
TESTRL	元素开始规则	字符型	说明元素的开始规则。
TEENRL	元素结束规则	字符型	说明元素的结束规则。每个元素必须具有TEENRL或TEDUR。
TEDUR	元素计划持续时间	字符型	元素计划持续时间，以ISO 8601格式表示。当元素的结束规则是在一个固定的持续时间之后结束时使用。

#### 3.2.2 试验分组表 (TA)

表 3.2.2: 试验分组——每一个分组、每一个计划的元素，一条记录

变量名称	变量标签	类型	描述
------	------	----	----

变量名称	变量标签	类型	描述
STUDYID	研究标识符	字符型	研究的唯一标识符。
DOMAIN	域名缩写	字符型	两个字符的域名缩写，必须是TA。
ARMCD	计划分组编码	字符型	ARMCD的长度限定为20个字符，并且无特殊字符限制。
ARM	计划分组描述	字符型	分组或治疗组名称。
TAETORD	分组中元素的顺序	数值型	分组中元素的序号。
ETCD	元素编码	字符型	ETCD (ELEMENT伴随变量) 限于8个字符并且无特殊字符限制。为了方便编程目的其值应该短小，但是不期望说ETCD必须作为一个变量名应用。
ELEMENT	元素描述	字符型	元素名称。
TABRANCH	分叉条件	字符型	在试验设计中，该元素结束时，受试者符合放入此治疗组的条件。例如：随机分配服用药物X。
TATRANS	过渡规则	字符型	如果试验设计允许受试者过渡到一个元素但不是序列中的下一个元素，在此具体说明过渡到那些其它元素或者另一个元素序列的条件（例如：有疗效者进入洗脱期）。
EPOCH	时期	字符型	与分组元素相关联的试验时期名称。

307 备注：同一个元素在一个治疗组中可发生多次，但是每次发生时的TAETORD和EPOCH的值都不同，而且TABRANCH和TATRANS的值也可能不同。

### 3.2.3 试验访视 (TV)

表 3.2.3: 试验访视——每一个计划的试验访视，一条记录。

变量名称	变量标签	类型	描述
STUDYID	研究标识符	字符型	研究的唯一标识符。
DOMAIN	域名缩写	字符型	两个字符的域名缩写，必须是TV。
VISITNUM	访视编号	数值型	临床会见号码。VISIT的数值型版本，可用于排序。
VISIT	访视名称	字符型	试验方案中定义的临床会见描述。可作为除变量VISITNUM和/或VISITDY之外关于临床会见的文字描述。
VISITDY	访视的计划研究日	数值型	计划访视的研究日。由于其序列本质，可用于排序。
ARMCD	计划分组编码	字符型	ARMCD长度限定为20个字符，并且无特殊字符限制。如果试验的访视时间不依赖于受试者在哪个分组，则ARMCD应为空(Null)。
ARM	计划分组描述	字符型	分组或治疗组的名称。
TVSTRL	访视开始规则	字符型	描述访视开始的规则及其与元素先后顺序的关系。
TVENRL	访视结束规则	字符型	描述访视结束的规则及其与元素先后顺序的关系。

### 3.3 试验入选/排除标准 (TI)

试验入选/排除标准表 (TI) 对于试验的每一条入选和排除标准，都有一条相应的记录。

#### 3.3.1 试验入选/排除表

表 3.3.1: 试验入选/排除标准——每一个试验入选或排除标准，一条记录

变量名称	变量标签	类型	描述
STUDYID	研究标识符	字符型	研究的唯一标识符。



变量名称	变量标签	类型	描述
DOMAIN	域名缩写	字符型	两字符域名缩写，必须是TI。
IETESTCD	入选/排除标准简称	字符型	IETEST入选/排除标准的简称。当数据集由纵向转为横向时，可以用于做为列名册。IETESTCD值既不能长于8个字符，也不能以数字开头（如，1TEST）。IETESTCD不能包含字母、数字或者下划线以外的字符。前缀取IE是为了和IE域保持一致。
IETEST	入选/排除标准	字符型	完整的入选/排除标准描述。前缀IE用于和IE域保持一致性。
IECAT	入选/排除分类	字符型	用于对入选/排除标准进行分类：入选（INCLUSION），排除（EXCLUSION）。
IESCAT	入选/排除子类	字符型	用于对例外标准的进一步分类。可以用于子试验的区分或者用于划分主要和次要的例外。例如：主要（MAJOR），次要（MINOR）。
TIRL	入选/排除标准规则	字符型	描述入选/排除标准的规则。
TIVERS	方案版本	字符型	入选/排除标准的版本号。如果只有一个版本的话可以省略。

### 3.4 试验概要信息（TS）

试验概要信息域（TS）对于每一个试验概要特征，均包含一条相应记录。试验概要用来记录试验的基本信息，例如试验分期、方案标题和设计目标。

#### 3.4.1 试验概要表

表 3.4.1: 试验概要——每一个试验概要参数，一条记录

变量	变量标签	类型	描述
STUDYID	研究标识符	字符型	研究的唯一标识符。
DOMAIN	域名缩写	字符型	两个字符的域名缩写，必须是TS。
TSSEQ	序号	数值型	序号，确保在一个数据集中的唯一性。
TSGRPID	分组ID	字符型	用于把相关记录归于一组。
TSPARMCD	试验概要参数简称	字符型	TSPARMCD（TSPARM伴随变量）限定于8个字符以内，并且无特殊字符限制。为便于编程其值应短小，但不期望TSPARMCD必须作为一个变量名称应用。例如：AGEMIN，AGEMAX。
TSPARM	试验概要参数	字符型	试验概要参数的术语。TSPARM中值的长度不可多于40个字符，例如：受试者年龄计划的最小值，受试者年龄计划的最大值。
TSVAL	参数值	字符型	试验概要参数TSPARM的值，例如：若TSPARM的值是“试验适应症”，其值是“哮喘”。TSVAL不能为空（null），必须有值该记录才为有效。超过200字符的文字可以增加列TSVAL1-TSVALn。



## 4 数据集之间和记录之间的关系描述

在很多种情况下是需要或最好能对数据集之间或数据记录之间的关系进行描述。SDTM 对此定义了五种不同的关系类型：

- 同一数据集特定受试者的一组记录之间的关系。
- 受试者独立记录之间（通常在不同数据集中）的一种关系，例如治疗不良事件的伴随用药。
- 两个或两个以上数据集之间关系，其中一个（或多个）数据集的记录与另一个（或多个）数据集的记录有关联。
- 在使用通用观测数据类型中的标准变量无法表示的数据中，数据集里面的记录可以链接回到原先记录的这种依赖关系。
- 注释域中的某个注释记录与在其它数据集中父记录之间的一种依赖关系，例如一个不良事件的注释记录。

实施指南对以上每种关系类型均给出了具体而详细的定义并附实例说明。

### 4.1 关联记录数据集（RELREC）

表 4.1.1: RELREC数据集

变量名称	变量标签	类型	描述
STUDYID	研究标识符	字符型	域记录的研究识别符。
RDOMAIN	关联域名缩写	字符型	父记录两个字符的域名缩写。
USUBJID	受试者唯一标识符	字符型	可唯一识别某一产品的所有申请或递交的所有研究受试者的标识符。
IDVAR	标识变量	字符型	在通用观察数据类型数据集中，用于识别关联记录的变量名称。例如：--SEQ， --GRPID。
IDVARVAL	标识变量值	字符型	IDVAR 中描述标识变量值。如果--SEQ 用于描述记录，那么--SEQ 值录入在此。
RELTYPE	关系类型	字符型	标识出关联记录层次水平，值必须是 一对一（ONE） 或多对多（MANY）。然而其值只在标识数据集之间关系时才有必要。
RELID	关系标识符	字符型	某一 USUBJID 内唯一标示值用以标识关系。对同一个 USUBJID 的所有具有相同 RELID 值的记录被认为相关/关联。RELID 可以是申办者选择的任意值，且只在 RELREC 数据集中用来指出相关/关联的域记录才有意义。

### 4.2 补充修饰语数据集（SUPQUAL）

表 4.1.2: SUPQUAL数据集

变量名称	变量标签	类型	描述
STUDYID	研究标识符	字符型	父记录的研究标识符。
RDOMAIN	关联域名缩写	字符型	父记录两个字符的域名缩写。
USUBJID	受试者唯一标识符	字符型	父记录的唯一受试者标识符。
IDVAR	标识变量	字符型	数据集中的标识变量确认相关记录。例如：--SEQ， --GRPID。
IDVARVAL	标识变量值	字符型	父记录的标识变量值。

变量名称	变量标签	类型	描述
QNAM	修饰语变量名称	字符型	修饰语变量的简称，用作父域的数据记录列名称。此变量取值不能长于 8 个字符，也不能以数字开头（如 1TEST）。QNAM 值不能包含除字母、数字和下划线以外的任意字符。经常是申办者的原始或运营数据集 <sup>2</sup> 的列名称。
QLABEL	修饰语变量标签	字符型	QNAM 全称或标签。QLABEL 值不能超过 40 个字符。经常是申办者的原始或运营数据集 <sup>3</sup> 的列标签。
QVAL	数据值	字符型	QNAM 的结果值。此变量必须赋值，SUPP--中不能含有 QVAL 值为空（Null）的记录。
QORIG	来源	字符型	因为 QVAL 可描述由采集（来自 CRF 表）、衍生和分配的混合数据，QORIG 可用来标识这些数据的出处。例如病历报告表（CRF）、分配（ASSIGNED）或者衍生（DERIVED）。
QEVAL	评估者	字符型	仅用于主观的结果（例如由某人或某小组判断赋予）。客观采集或衍生的记录，其值应为空（Null）。评估人员可包括仲裁委员会、统计师、数据库管理员、临床协调员等。

338

<sup>2</sup>【译者注】原文为 operational dataset。这里翻译成原始或运营数据集。

<sup>3</sup>【译者注】原文为 original dataset。这里翻译成原始或运营数据集，和 QNAM 描述部分保持一致。

## 5 利用此模型进行数据监管递交

SDTM采用标准化设计，以便最大程度地覆盖来自于人类或动物研究的数据。本文档描述了模型的基本概念及通用结构。多个实施指南分别涵盖了对来自于临床研究与动物研究采集数据关于域定义的具体建议，明确了每个域所包含来自某个通用观测数据类中的变量。目前公布的实施指南包括用于人类临床试验的CDISC SDTM实施指南（SDTMIG），及用于动物实验的非临床数据交换标准实施指南（SEND）。这些实施指南亦描述了实施过程中的一些基本假定与业务规则，并提供了大量实例来阐明如何从原始数据到标准数据的转换。每个希望递交标准数据的申办者在准备递交材料前，均应首先参考上述实施指南（SDTMIG和SEND）。

## 6 SDTM 版本历史（暂未翻译）

本节暂未翻译，请自行参考原文（<http://www.cdisc.org/standards>）。

### 6.1 从 SDTM V1.1 到 SDTM V1.2 的改动

本节暂未翻译，请自行参考原文。

## 7 附录 A 陈述和保证、责任限制和免责声明（暂未翻译）

译者注：未翻译，请参考原文。

注：CDISC 知识产权政策可在以下网站找到：

[http://www.cdisc.org/stuff/contentmgr/files/0/4c011269a1cb48cb0c8eab2c3e172694/misc/cdisc\\_20ip\\_20policy\\_final.pdf](http://www.cdisc.org/stuff/contentmgr/files/0/4c011269a1cb48cb0c8eab2c3e172694/misc/cdisc_20ip_20policy_final.pdf)