



分析数据模型（ADaM）

CDISC 分析数据模型团队 （ADaM Team）开发

中文版读者说明

本文档为分析数据模型（ADaM）2.1 版，包含了所做的修订，对应于分析数据模型实施指南（ADaMIG）1.0 版。

本中文版初稿由 Absolute Clinical Data Systems Co., Ltd 提供，C-STAR ADaM 团队负责修改、并校对和审订。翻译和审核流程详见第 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2日	终版	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	草案	待评论草案。

英文版修订历史

日期	版本	修订小结
2009年12月17日	2.1终版	发布的版本，反映了在评论时期确定的所有改动和更正。
2008年5月30日	2.1草案	待评论草案。
2006年8月11日	2.0终版	最终文档
2006年5月31日	2.0草案	包含来自公众审阅的评论
2006年2月15日	2.0草案	对于一般性考虑v1.0版的重新格式化，包含了受试者水平的模型，强调要求、命名、内容规则和指导原则。

注释：请参见附录G的陈述与保证；责任限制，及免责声明。

CDISC 分析数据模型 (ADaM) V2.1 翻译和审核 流程及致谢

翻译

由 Absolute Clinical Data Systems Co., Ltd (AbsCD) 公司提供 ADaM V2.1 翻译初稿。

审核

中国 CDISC 协调委员会 (C3C) 于 2011 年 5 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, 本中文版基于英文 V2.1 正式版, 该版本于 2009-12-17 正式发布。

校对审核遵循 C-STAR 项目组批准的流程。尽管所有校对人员都是有经验的志愿者, 但时间有限。ADaMIG 按照内容分成 3 大部分, 分别由 3 个小组来承担, 每个小组内任何部分内容均由两个独立的翻译校对。小组负责人、ADaM 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 ADaM 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 及与 CDISC 其它标准 (SDTM, ADaM IG) 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 所有流程审阅后的终稿即发布在 CDISC 网站征求意见。ADaM 翻译审阅团队根据收集到意见进行了修改定稿。最终定稿将发布在 CDISC 网站。

由于这是第一次尝试对 ADaM 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!

没有 AbsCD 和全体 ADaM 翻译审阅团队的努力, ADaM V2.1 中文稿的面世是难以想象的。核对的工作有时是很枯燥乏味的。对一些不易理解的句子, 难以翻译的术语和多个不同中文术语同时存在的, 翻译校对员需要经过多次反复推敲才能定夺。谢谢所有志愿者的奉献和 (审核时) 所属公司的支持!

C-STAR ADaM 翻译审阅组

目录

67		
68	1 引言/目的.....	4
69	2 背景/动机.....	5
70	3 分析数据模型概况.....	6
71	3.1 基本原则.....	6
72	3.1.1 可溯源性.....	6
73	3.2 分析数据流.....	6
74	3.3 元数据组成部分.....	7
75	4 分析数据集.....	9
76	4.1 实践方面的考虑.....	9
77	4.1.1 分析数据集的数量和内容.....	9
78	4.1.2 分析数据集和变量命名规则.....	9
79	4.1.3 变量排序.....	9
80	4.2 ADaM 数据结构.....	9
81	4.2.1 受试者水平分析数据集 (ADSL) 结构.....	9
82	4.2.2 基本数据结构 (BDS).....	10
83	4.2.3 未来的 ADaM 数据结构.....	10
84	5 ADaM 元数据.....	11
85	5.1 分析数据集元数据.....	11
86	5.1.1 分析数据集元数据举例.....	11
87	5.2 分析变量元数据.....	11
88	5.2.1 分析参数值水平的元数据.....	12
89	5.2.2 分析变量元数据举例 (包括分析参数值水平元数据).....	13
90	5.3 分析结果元数据.....	18
91	5.3.1 分析结果元数据举例.....	18
92	6 受试者水平分析数据集.....	23
93	6.1 不分析的受试者数据.....	23
94	附录.....	24
95	附录 A 文献.....	24
96	附录 B 定义.....	25
97	附录 C 缩写和简略词.....	27
98	附录 D 对即可分析的阐释.....	28
99	附录 E 复合终点的例子.....	29
100	附录 F 修订历史.....	32
101	附录 G 声明与保证; 责任范围, 及免责声明.....	33

103

1 引言/目的

分析数据模型(ADaM)文档具体说明在创建分析数据集和相关的元数据时要遵循的基本原则和标准。元数据是“关于数据的数据”或者“关于数据的信息”。分析数据模型支持分析结果有效率的生成、复制、和审阅。

分析数据集的设计通常由临床试验的科学和医学目标推动，一个基本原则是分析数据集的结构和内容必须支持临床试验的清晰无歧义的科学和统计方面的交流。

ADaM的目的是提供一个启动数据分析的框架，同时允许评审者和其它数据接收者清楚地理解数据从收集到分析结果的传承。ADaM是为了最优化地支持数据的衍生和分析；CDISC的研究数据列表模型(SDTM)是为了最优化地支持数据列表。

ADaM文档（也就是本文档）提供了核心要点并定义了ADaM概念和标准的精神及意图，它概述了创建分析数据集和相关的元数据时要遵循的基本原则。在本文档中描述了四种类型的ADaM元数据（即分析数据集元数据、分析变量元数据、分析参数值水平元数据和分析结果元数据），并且提供了范例。

本文档介绍了受试者水平的分析数据集(ADSL)（请参见4.2节）。一个基于CDISC的临床试验数据递交，即便没有其它分析数据集递交时，至少需要具备ADSL及其相关元数据。

本文档也介绍了ADaM基本数据结构(BDS)，不论治疗领域或者分析类型，此结构将被应用于大多数的ADaM数据集（请参见4.2节）。虽然一般而言BDS支持大多数的统计分析，但是一个研究也会包括特定标准化结构的分析数据集，用来表达其它的分析信息，比如受试者水平的分析数据集(ADSL)和不良事件分析数据集(AEAE)。

本文档是ADaM实施指南(ADaMIG)的基础，ADaMIG具体说明这些核心概念的标准化实施，具体说明ADaM标准数据集的结构和变量，包括命名规则。它还说明了对于实施问题的标准解决方法。

在采用ADaM的原则和标准构建数据集及其元数据时，要想充分获得分析数据集所能带来的益处，审阅者或其他数据接收者与申办者之间尽早而有效的交流是必不可少的，这方面是无论如何强调都不过分的。

在对于ADaM概念的说明中，范例引用了特定的编程语言。在ADaM文档从始至终，引用特定供应商的产品只是为了举例，不应理解为是对这些产品的支持。

注，本文档中的范例只是用于阐释，不应视为标准本身的内容。

2 背景/动机

监管的人类健康产品的上市批准过程通常包括临床试验数据的递交。在美国，数据是向美国食品药品监督管理局递交的必要部分。

在1997年，FDA通过出版使用电子记录代替纸质记录（21 CFR 第11章）的规章制度建立了对于电子递交数据的监管基础。在1999年，FDA标准化了临床试验中收集的电子递交数据的文件格式（SAS® 版本5 传输文件¹）。这在描述临床数据和数据定义文件(define.pdf)递交指南文档系列的第一个文档中进行了解释。

虽然1999年的指南在2006年撤销，但数据集仍然使用SAS传输文件格式递交，伴有一个定义文件“define.pdf”。这个定义文件是一个数据定义文档，它提供了一个包括在递交中的数据集清单，和对于每个数据集内容的一个详细描述（即递交数据集的元数据）。

从2005年开始，元数据可以使用可扩展的标记语言（XML）格式(define.xml)而不是PDF文档格式(define.pdf)递交，这在FDA的关于研究数据说明书文档中有所描述[7]。更多的关于define.xml的信息可以在CDISC网站找到[2]。

与制订临床数据递交指南并行，FDA已经采用ICH的人用药品注册技术要求作为注册递交的标准，发行了关于电子通用技术文档(eCTD)的指导文档作为医药产品申请的电子递交的框架。这个指南的第2次修订在2008年发布[6]。

根据FDA关于eCTD的指导文档，递交数据可分为四类：1) 数据列表，2) 数据清单，3) 分析数据集，4) 受试者概况。这些总集在一起称为病例报告制表(CRTs) [6]。递交的组织文件夹中数据集及其相关文件的说明总结于下图，来自“研究数据说明”。

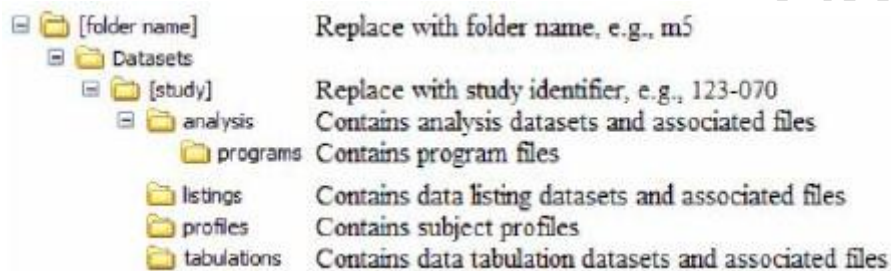


图2.1 组织文件夹中研究数据集及其相关文件的说明[7]

数据列表数据集和分析数据集定义为：

- **研究数据列表 (SDTM)** - 该数据集包含研究中收集的数据，按临床域组织。CDISC研究数据列表模型[5]和CDISC研究数据列表模型实施指南(SDTMIG) 中对这些数据集进行了描述[4]。
- **分析数据集 (ADaM)** - 该数据集用于统计分析和申办者报告，与SDTM域一起递交。ADaM数据集是用于统计分析的所有衍生数据的权威来源。这些数据集在本文档（CDISC分析数据模型文档）和CDISC ADaMIG[1]中描述。

标准化的分析数据集和元数据提供给数据接收者的好处不仅是清楚的交流和透明度。一旦接受了关于标准化数据集的原则的培训，审阅者和其它数据接收者能用较少的准备时间、更有效率的使用数据。另外，标准化的结构允许软件工具的开发，从而有益于使用、衍生、分析、重现、审阅分析结果。

SDTM的设计不是用来支持统计分析。ADaM数据集把衍生的和收集的数据（来自各种SDTM域、其它ADaM数据集、或任何它们的组合）合并成一个数据集，以便只经过很少的或不用编程就能进行分析。SDTM中不太容易处理的问题的实例是分析时间窗、复杂算法和缺失数据填补。

ADaM元数据解释了ADaM数据集是如何从SDTM源数据创建的，因而在ADaM数据集中衍生和填补的变量不应再拷贝回SDTM源数据，试图这样做会在数据流中引入循环依赖，使变量间的重要关系失去关联。

为了简化本文档，分析数据集是在FDA电子递交背景中讨论的。从一开始，CDISC ADaM小组就得到了FDA统计和医学审阅者们的鼓励和知会，他们作为观察员参加ADaM会议，参与CDISC-FDA预试验。

ADaM基本原则的起源是与医学和统计审阅者交流的透明性和完整性的需求和医学和统计审阅者的评审的科学上正确性的要求。ADaM标准的制订是为了满足FDA和工业界的需求。除了适用于FDA的注册递交，ADaM还适用于更大范围的药物研究活动。ADaM为在申办者和CRO之间以及开发伙伴和独立数据监测委员会之间传输数据提供了标准。随着模型更加广泛的采用，申办者间分析数据集及相关元数据的共同模型就会推动买入授权、对外授权、联合投资及合并。

3 分析数据模型概况

3.1 基本原则

基本原则

分析数据集及相关元数据必须:

促进清晰无歧义的交流

提供分析数据和它的源数据（最终是SDTM）之间的可溯源性

能很容易地在常用软件工具上使用

分析数据集必须:

伴有元数据

即可分析

设计分析数据集和相关元数据的总体原则，是必须对于支持临床研究中进行的统计分析的数据集的内容和来源有清晰无歧义的交流。这个原则的内在要求就是可溯源性，可藉此理解分析值（分析结果或分析变量）从何而来，即一个分析值和它的前身之间的数据传承或关系。请参见3.1.1节中对于可溯源性更加详细的描述。

申办者应该力求递交“即可分析”的数据集，亦即其结构和内容允许只需很少量的编程就可进行统计分析。必要时仅需很少的额外处理，如排序或从分析集中选择适用的记录，统计分析软件即可直接使用即可分析数据集。不需要复杂的数据操作，如变换或转置，即可进行其支持的分析。这一方法的使用，使得分析员，如统计审阅者，不需要或最大限度地减少了对编程的需要。附录D给出在SAS中应用此原则的例子，而实际上这一理念适用于所有统计分析软件包。请注意，在ADaM背景下，分析数据集至少包括审阅与再现特定统计分析所需的数据。“即可分析数据集”并不要求包括仅用于支持数据清单或其它非分析性展示的数据，尽管也有人这么做。

分析数据集必须能很容易地在常用软件工具上使用，且必须伴有相应的元数据。理想的元数据是可机读的。元数据和其它文档，必须提供对分析的清晰、简明的交流，包括统计方法、假设、衍生和填补等。元数据、程序和其它文档用于系统化统计分析计划书（SAP）所述的分析及其它分析。这方面的详细讨论见第5章。

3.1.1 可溯源性

可溯源性的概念是分析数据模型的基石，这一特征使得可以理解元素与其前身之间在数据上的传承与关系。可溯源性促进了透明度，这是建立对结果或结论的信心的基本点。最终，ADaM的可溯源性，使得可以理解分析结果、分析数据集与SDTM域之间的关系。

可溯源性，是通过清晰地建立元素与它的直接前身之间的路径而实现的。整条路径，是通过从一个元素到它的前身，再到前身的前身等等来追踪，诸如此类，回追到SDTM域，最终到达数据采集工具。请注意CDISC数据采集标准协调会（CDASH）标准与SDTM是一致的，因此能够保证端到端的可溯源性。可溯源性建立了数据集间与数据集内的关联。例如，同一分析数据集中标识变量和其它变量的元数据，使得用户能够理解衍生的数据是如何（且，从某种程度来讲，为什么）生成的。

有两个水平的可溯源性:

- 元数据的可溯源性使得用户能够理解分析变量与其来源数据集及变量之间的关系，这是ADaM遵循的要求。这一可溯源性，通过描述（由元数据）使用的算法或者从其直接前身衍生或填充一个分析值采取的步骤建立起来。元数据的可溯源性也用来建立一个分析结果（如p值）和分析数据集之间的关系。
- 数据点可溯源性使得用户能够直接到达特定的前身记录，应尽可能实施。这个水平的可溯源性对审阅者试图追踪复杂的数据操作路径非常有用，这一可溯源性，通过提供数据（例如通过使用--SEQ变量）与用作分析值的输入的特定数据值之间的清晰链接而建立。需要注意，在有些情形下，数据点的可溯源性可能很困难、不切实际、甚至不可行，例如多功能睡眠记录仪研究中的脑电图记录，关键的结果可能是基于通过快速傅里叶转换从绝对电压结果衍生的光谱边参数。

当可溯源性成功实施时，审阅者能够确定:

- 存在于递交的SDTM研究列表数据中的信息
- ADaM分析数据集中衍生或填补的信息
- 创建衍生或填补数据的方法
- 用于分析的信息，相对于没有用于分析的信息，包括它们是为了支持可溯源性或将来的分析。

3.2 分析数据流

图3.2.1是一个典型的从起源到分析结果的一般数据流概念图，这个简略图只是阐述了一个合理的情形，并不是打算画出所有可能的关系和组成部分，也不是指出或者暗示这是实施此过程的唯一途径。例如，元数据可能实际上是彰显或者推动这个过程，而不是这个过程的输出。

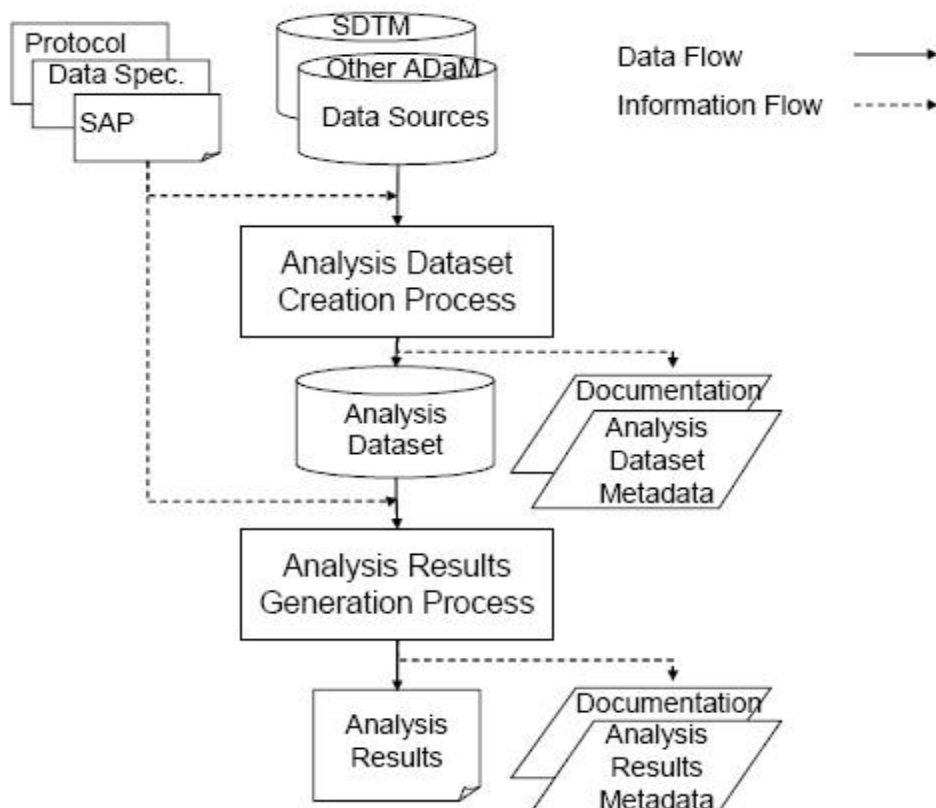


图3.2.1：分析数据流程图，显示数据流和信息流的一种情形

既然制订的ADaM标准是CDISC标准大家族的一部分，可假定它的来源是SDTM或其它分析数据集如受试者水平分析数据集（ADSL）。一个遵循CDISC标准的递交包括SDTM和ADaM数据集，因此，为了促进可溯源性，元数据需要描述这两套数据集之间的关系。

为了促进清晰的交流，对生成分析数据集和产生分析结果的过程进行了区分。这两个过程目的不同，因而需要不同类型的元数据，如下节简述。

分析数据集创建 — 用来创建分析数据集的处理和编程步骤。如在图3.2.1所示，分析数据集创建程序编写是根据分析计划和数据集说明。输入程序的数据是源数据（即SDTM及/或其它分析数据集），输出的是分析数据集。

分析结果的产生 — 用于产生分析结果（即图表表展示中的总结或统计推断）的编程步骤。如图3.2.1所示，分析结果产生程序是基于统计分析计划、数据衍生和数据集说明编写的，输入程序的数据是分析数据集，输出的是分析结果。

3.3 元数据组成部分

分析数据集和ADaM元数据易化临床试验数据的审评和进行的分析，本文档描述了四种类型的元数据，包括：

- 分析数据集元数据：据描述每个分析数据集，包括对数据集内容的简短描述（请参见5.1节）。这种类型的元数据对所有的ADaM数据集都是必需的。
- 分析变量元数据：描述分析数据集中的变量，包括分析变量来源及创建的信息，如对于分析数据集创建程序涉及的算法和参考资料的详细描述（请参见5.2节）。这种类型的元数据对所有的ADaM数据集都是必需的。
- 分析参数值水平元数据：描述的是一个分析参数内的检查或分析终点（即分析参数的每个唯一值）。这种形式的元数据在一个变量包含多种类型的检查或分析终点的数据结构中尤其需要，如ADaM基本数据结构（请参见5.2.1节）。这种类型的元数据对所有的ADaM BDS数据集都是必需的。
- 分析结果元数据：描述的是分析结果（由申办者规定），包括使用的数据集和进行的分析的信息（请参见5.3节）。这些元数据提供从统计显示结果到分析数据集中数据的可溯源性。分析结果元数据不

是必需的，但推荐提供这些元数据，以便审评者确定关键的分析，提供结果、文档、数据集之间的链接，记录进行的分析。

前面的三种类型的元数据描述分析数据集，它们在分析数据集创建的过程中形成。分析数据集元数据对分析数据集进行了整体描述，而分析变量元数据和分析参数值水平元数据描述了数据集内的变量和观测值。

为了记录分析数据创建过程或者分析结果产生过程，元数据可以包括伪代码、代码片段、到程序的链接和/或到方案、统计分析计划、或其它文档的链接。

4 分析数据集

4.1 实践方面的考虑

分析数据集必须:

- 包括一个名为“ADSL”的受试者水平数据集（参见第6章）
- 由最佳数量的分析数据集组成，有足够的自足性，只需很少或不需要额外的编程或数据处理就能分析和审阅。
- 使用规则“ADxxxxxx”命名
- 使用ADaM标准变量名称，且遵循命名规则（适用时）。
- 如果是从SDTM变量直接拷贝且未改变名称的变量，则保留其值与属性（即遵循“同名、同义、同值”的一致性原则），保留SDTM变量的值和属性。
- 一次递交的多个研究及同一产品的多次递交中，使用相同的数据集和变量命名规则。

4.1.1 分析数据集的数量和内容

在临床研究报告或者递交中创建支持分析结果的分析数据集时，一个目标是产生进行各种分析所需要的最佳数量的分析数据集（最少有ADSL）。不要求不同的分析各有不同的数据集，一个单独的数据集可以同时支持多个分析。也不要求每个数据总结都一定要有分析数据集支持。另外，也不要求每个SDTM域都有一个相应的分析数据集，由申办者决定要创建的分析数据集。

创建一个分析数据集，可能需要多个数据集（如SDTM、其它分析数据集）。分析数据集必须包含所需的全部变量，以进行其设计用于支持的统计分析。例如，创建一个分析数据集的数据可能来自受试者水平分析数据集（ADSL）、处置（DS）、人口统计学（DM）、受试者特征（SC）、生命体征（VS）、调查问卷（QS）和暴露（EX）域。

设计分析数据集是为了促进最小程度的编程或数据处理就可以审阅和分析。分析数据集间的冗余（即在多个数据集内出现同一数据）经常是必要的，以便数据集是即可分析的（例如，不良事件分析数据集和疗效分析数据集都有年龄）。类似的，也可以包含一些变量和记录，实际上在任何的递交分析中都用到、但申办者或者审阅者仍然感兴趣（如发生临床关注事件的受试者的标识）或用于支持可溯源性。

附录E有一个具有复合终点的例子，需要复杂的算法和来自多个数据集的输入。

4.1.2 分析数据集和变量命名规则

分析数据集使用规则“ADxxxxxx”进行命名。受试者水平的分析数据集如在第6章所述，命名为“ADSL”。对于所有其它的分析数据集，名称的xxxxxx部分是申办者定义的，在给定的递交或者是一个产品的多个递交中使用共同的命名规则。在制订命名规则时，申办者应考虑eCTD指南文档中的要求[6]，也需要考虑遵循SAS传输格式的要求（如名称总长度不超过8位字符）。

分析数据集中创建的变量（不要与SDTM要求的任何标准变量混淆）的命名规则，应遵循标准变量名称和ADaM中定义的命名规则。如果分析变量名称是申办者定义的，只要有可能，也应在一个给定的递交和一个产品的多个递交中遵循一个共有的命名规则。

任何与SDTM变量同名的ADaM变量必需是SDTM变量的复制，它的标签、属性和值均不能有改动，ADaM遵循一致性原则，也就是所谓的“同名、同义、同值”。

更多的一般变量命名规则，请参见ADaMIG[1]。

4.1.3 变量排序

理想的来讲，分析数据集中变量的排序遵循逻辑顺序（不是简单的字母顺序）。关于分析数据集中变量顺序的更多信息，请参考FDA“研究数据说明”[7]。推荐申办者定义一个在数据集中排序变量的规则，然后对所有的数据集一致地应用该排序。数据集中变量的排序应该与定义文件中所述的顺序相匹配。

4.2 ADaM数据结构

本文档和ADaMIG中描述了二种ADaM标准数据结构：受试者水平分析数据集（ADSL），和基本数据结构(BDS)。

4.2.1 受试者水平分析数据集（ADSL）结构

ADSL数据集结构中一个受试者一条记录，包含受试者水平人群标识、计划的和实际的治疗组变量、人口统计、随机化因子、亚组变量和重要的日期。ADSL包含必需的变量（如ADaMIG所述）和其它用来描述试验中受试者经历的受试者水平重要变量。基于CDISC的临床试验数据递交中，即便没有其它分析数据集递交，ADSL及相关元数据也是必需的。关于ADSL的详细描述，请参见第6章。

虽然把研究中采集每个单独数据值拿来作为受试者水平的数据集如ADSL中的变量在技术上是可行的，

但那不是ADSL的意图或者目的。研究过程中随着时间面改变的关键终点或数据的正确位置是在一个BDS数据集中。

4.2.2 基本数据结构（BDS）

在一个BDS中，每个受试者、每个分析参数、每个分析时间点可有一条或多条记录。随分析的不同，某些情形下分析时间点是必需的。在没有分析时间点的情形下，其结构是每个受试者、每个分析参数一条或多条记录。这个结构包含描述分析参数（如PARAM及相关变量）的变量中心集，且包含分析值（如AVAL和AVALC及相关变量）。数据集中的其它变量提供了更多关于分析值的信息（例如受试者标识）或描述、追踪其衍生（如DTYPE），或使分析（如治疗组变量、协变量）成为可能。BDS支持参数和非参数分析，如ANOVA、ANCOVA、类别分析、逻辑回归、CMH检验、Wilcoxon秩和检验、到事件发生时间的分析等等。通常最发有多个的BDS分析数据集。BDS标准的细节，请参见ADaMIG [1]。

虽然BDS支持大多数的统计分析，但并不是所有的统计分析。例如，它不支持多个因变量（反应变量/结果变量）的同步分析或者一系列反应变量的相关分析。BDS的设计不支持不良事件或其它事件数据发生率的分析。

4.2.3 未来的ADaM数据结构

ADaM小组正在编写支持不良事件发生率分析的ADAE数据集说明文档。ADAE可能是更广泛地支持发生率分析的结构的一个示例，如不良事件、伴随用药等等。

5 ADaM元数据

ADaM元数据

- 分析数据集元数据
- 分析变量元数据
- 分析参数值水平元数据
- 分析结果元数据

临床试验分析基于的假设、统计方法、转换、衍生和填补应可清晰交流，这样可以很容易地重现值和结果。ADaM元数据通过提供分析的一般描述（方案数据分析章节、SAP、或者报告的分析方法）、分析结果、用于分析的数据和SDTM域的详细描述和之间的链接促进交流。下面的章节详细地描述了ADaM元数据的各组成部分。

所述分析数据集元数据和分析变量元数据的元数据结构是基于病例报告制表数据定义说明标准1.0.0版 (CRT-DDS) [2]，更多细节，请参考该文档。

本文档中的元数据示例只是为了解释说明，不是为了指定或者推荐报告的风格、格式或过程。此外，一些图解中的斜体字行只是作为引用，作为对字段定义的提醒，不是为了让申办者把这一行包括在元数据中。

5.1 分析数据集元数据

分析数据集元数据提供关于分析数据集的信息，包括对数据集内容的描述。最佳实践强烈推荐每个分析数据集都使用表5.1.1中列出的元数据字段描述。ADSL和BDS分析数据集（即遵循ADaM）必须用这些元数据字段描述。实践经验告诉我们，有时候也会用到非这些结构的分析数据集，建议也使用表5.1.1中的元数据字段描述它们。

表5.1.1 分析数据集元数据字段

分析数据集元数据字段	描述
数据集名	数据集的文件名，超链接到定义文件的相应分析数据集变量描述（即，数据定义表）。
数据集描述	对于数据集内容的简短描述性总结。
数据集位置	可以找到数据集的文件夹和文件名，最好超链接到实际的数据集（即XPT文件）。
数据集结构	数据集各记录代表的详细水平（例如，“每个受试者一条记录”，“每个受试者每个访视一条记录”，“每个受试者每个事件一条记录”）。
数据集的关键变量	与结构相应的变量名列表，最好能唯一标识和索引数据集中的每条记录。
数据集类型	确定数据集的总的类型，ADaM结构名（即“ADSL,”“BDS”）或“其它”（如果不是ADaM说明的结构）。
文档	关于数据集创建的起源数据、处理步骤和分析决定的描述。各种级别的功能性和复杂性的软件代码，可能提供如伪代码、或实际的代码片段。可能使用到外部文档的链接或者引用（如方案、统计分析计划、软件代码）。

5.1.1 分析数据集元数据举例

在本例中，分析的数据是阿尔茨海默病评估量表-认知部分（ADAS-Cog）11项分数的总和。评估的是调查问卷，所以调查问卷（QS）SDTM域用于收集数据。表5.1.1.1阐明了ADAS-Cog分析数据集ADQSADAS的分析数据集元数据。

表5.1.1.1 ADQSADAS分析数据集的分析数据集元数据

数据集名	数据集描述	数据集位置	数据集结构	数据集关键变量	数据集类型	文档
数据集的文件名	数据集内容的简短总结	可以找到数据集的地方	数据集的细节的水平	和结构平行的变量的名称	使用受控术语的数据集一般类型	链接或者引用关于数据集如何产生的文档
ADQSADAS	ADAS-Cog (11)分析的数据	adqsadas.xpt	每个受试者、每个参数、每个分析访视一条记录	USUBJID, PARAMCD, AVISIT	BDS	DSADQSADAS.SAS, 详细的ADAS-Cog的评分算法见SAP的14.11节

5.2 分析变量元数据

分析变量元数据描述分析数据集中的每个变量，包括变量的属性和定义。提供这些描述的元数据字段列在了表5.2.1中。最佳实践强烈的推荐每个分析变量都使用这些元数据字段进行描述。遵循ADaM的分析

373 数据集必须用这些分析变量元数据字段进行描述。

374 表5.2.1分析变量元数据字段

分析变量元数据字段	描述
数据集名	分析数据集文件名
变量名	变量名
变量标签	变量的简短描述
变量类型	变量类型。有效值如“病历报告制表数据定义说明标准”所定义（如版本1.0.0，有效值包括“文本”、“整型”、和“浮点型”）。
显示格式	变量显示信息（即，在结果展示的表格或图形变量的格式）。建议语法和用于分析的软件包内嵌的格式术语相一致（如\$16或3.1，如果使用SAS）。
编码列表/受控术语	有效值或允许编码的列表以及对应的变量解码。这个字段可以包括对外部编码列表的引用（由名称和版本识别）或者超链接到定义文件的编码列表/受控术语章节。
来源/衍生	提供了关于变量传承的细节 –前身是什么，其中变量来自的源数据（SDTM或其它分析数据集）或者这些变量是怎样衍生的。这个字段用于识别直接前身来源和/或算法的简短的描述、或用于该来源的过程，能够包含超链接的文本，为读者提供更多参考信息。 来源/衍生可以简单到只有两个水平的名称(如ADSL.AGEGR)来识别变量来源的数据文件和变量（即拷贝的变量，没有任何改动）；可以是衍生的简单描述和用于衍生的变量（如“ADSL.BMI的归类”）；也可以是复杂的算法，其中元素包含完整的衍生算法的描述和/或到包含它的文档的链接和/或到分析数据集创建程序的链接。

375 参考5.2.2节的分析变量元数据的例子。

376

377 5.2.1 分析参数值水平的元数据

378 遵循ADaM基本数据结构（BDS）的分析数据集，即BDS类型的分析数据集，能包含多个分析参数。

379 BDS分析数据集中，变量PARAM包含对该数据集中每个分析参数的唯一描述。变量PARAMCD包含

380 PARAM中的分析参数的编码名称，两变量间是一一映射的。每个PARAM值标识数据集中一组的一或多行。

381 数据集的列（变量）的元数据通常依赖于PARAM/PARAMCD的值，这个概念与SDTM中的一个单独

382 变量值水平元数据类似，但在BDS中，几个变量的元数据的PARAM/PARAMCD不同很常见。为了描述这

383 些变量的PARAM/PARAMCD不同的情况，BDS分析数据集的变量水平元数据的元素必需有**参数标识符**。

384 参数标识符元数据元素确定随PARAM/PARAMCD不同而异的元数据，把变量元数据链接到PARAM/

385 PARAMCD的相应值。

386 受控术语减少了为PARAM/PARAMCD的多个值输入相同变量元数据的需求：

387 • 变量的参数标识符为“*ALL*”，说明分析数据集中那个变量的元数据对于PARAM/PARAMCD的所有值都是一样的。

389 • 变量的参数标识符为“*DEFAULT*”，除非另有说明，那个变量的特定的元数据应视为分析数据集中PARAM/PARAMCD所有值的元数据。

391 • 变量的参数标识符中PARAMCD的一个特定值说明那个变量指定的元数据应视为对特定的

392 PARAMCD适用的元数据，如果有的话，优先于指定的*DEFAULT*元数据。

393 参见5.2.2节中BDS数据集的分析变量元数据的例子，使用了参数标识符元数据元素应该注意的是，这

394 个元数据元素促进了元数据内容的输入和跟踪；在定义文件中如何展示元数据由申办者决定。

分析变量 元数据字段	描述
参数标识符	包含 1) PARAMCD的值，标识变量元数据适用的分析参数，或 2) 受控术语，指出分析参数的归类： *ALL* - 当变量元数据适用于数据集中所有的分析参数时使用 *DEFAULT* - 当变量元数据适用于数据集中除了那些在元数据特别列出的分析参数之外的所有的分析参数时使用。

395 参考PARAMCD的编码列表，数据集使用者就能够确定数据集中找到的所有分析参数值，能够确定分析参数特有的属性和当PARAMCD为某一特定值的时候变量的衍生算法。

397 注意对于PARAMCD变量，参数标识符是“PARAMCD”，变量存在的值的列表也作为分析数据集中分析参数和参数标识符的索引。

399

5.2.2 分析变量元数据举例（包括分析参数值水平元数据）

这个例子是在ADAS-Cog分析举例的扩展，分析的数据是ADAS-Cog (11) 总分，即ADAS-Cog11条目的子量表。评估是调查问卷，所以收集的数据在QS SDTM域中。表5.2.2.1 是ADAS-Cog分析数据集ADQSADAS分析变量元数据的举例，如5.1.1节中所述。

与任何其它BDS分析数据集一样，使用了分析参数值水平元数据。数据集包含各条目的分数，也包括总分，所以某些变量需要有不同的元数据，依赖于PARAM的值。本例用末次观测结转（LOCF）填补总分的缺失值；对单项的分数不进行填补。注意，PARAMCD的所有值必须列在编码列表元素中。

407

表5.2.2.1 ADQSADAS数据集的分析变量元数据

数据集名	参数标识符	变量名	变量标签	变量类型	显示格式	编码列表/受控术语	来源/衍生
分析数据集的文件名	PARAMCD 或 *ALL* 或 *DEFAULT*	名称	描述	类型	显示信息	有效值或编码和解码	变量从源数据的何处来或变量是怎样衍生的
ADQSADAS	*ALL*	STUDYID	研究标识符	text	\$12		ADSL.STUDYID
ADQSADAS	*ALL*	SITEID	研究中心标识符	text	\$3		ADSL.SITEID
ADQSADAS	*ALL*	SITEGR1	研究中心组1	text	\$3		ADSL.SITEGR1
ADQSADAS	*ALL*	USUBJID	受试者唯一标识符	text	\$11		ADSL.USUBJID
ADQSADAS	*ALL*	AVISIT	分析访视	text	\$19	基线, 第8周, 第16周, 第24周	如果ADQSADAS.ITTRFL='Y', 那么AVISIT是分析访视的名称, 如果ADQSADAS.ITTRFL=空, 那么AVISIT=空。基于ADQSADAS.ADY确定分析访视的时间窗算法的详细描述, 参见SAP的8.2节。
ADQSADAS	*ALL*	VISIT	访视名称	text	\$19		QS.VISIT
ADQSADAS	*ALL*	AVISITN	分析访视(N)	整型	3.0	3=基线, 8=第8周, 10=第16周, 12=第24周	如果ADQSADAS.ITTRFL='Y', AVISITN=AVISIT的数字编码, 如果ADQSADAS.ITTRFL=空, 则为空。
ADQSADAS	*ALL*	ADY	分析相对日	整型	3.0		如果ADQSADAS.ADT >= ADSL.TRTSDT, 则ADY=ADQSADAS.ADT - ADSL.TRTSDT + 1; 如果ADQSADAS.ADT < ADSL.TRTSDT, 则ADY=ADQSADAS.ADT - ADSL.TRTSDT。
ADQSADAS	*DEFAULT*	PARAM	参数	文本	\$16	ADAS-Cog第01项, ADAS-Cog第02项, ADAS-Cog第03项, ADAS-Cog第04项, ADAS-Cog第05项, ADAS-Cog第06项, ADAS-Cog第07项,	当ADQSADAS.PARAMCD指的是单项分数 (而不是总分) 时, PARAM是QS.QSTESTCD = ADQSADAS.PARAMCD 时QS.QSTEST的相应值 (对于受试者和访视)。

数据集名	参数标识符	变量名	变量标签	变量类型	显示格式	编码列表/受控术语	来源/衍生
						ADAS-Cog第08项, ADAS-Cog第09项, ADAS-Cog第10项, ADAS-Cog第11项, ADAS-Cog第12项, ADAS-Cog第13项, ADAS-Cog 第14项	
ADQSADAS	ACTOT11	PARAM	参数	文本	\$16	ADAS-Cog11 总分	'ADAS-Cog11 总分'被分配给总分记录
ADQSADAS	PARAMCD	PARAMCD	参数编码	文本	\$8	ACITM01, ACITM02, ACITM03, ACITM04, ACITM05, ACITM06, ACITM07, ACITM08, ACITM09, ACITM10, ACITM11, ACITM12, ACITM13, ACITM14, ACTOT11	与PARAM对应
ADQSADAS	*DEFAULT*	AVAL	分析值	浮点	3.0		当ADQSADAS.PARAMCD指的是单项分数（而不是总分）时，AVAL是当 QS.QSTESTCD = ADQSADAS.PARAMCD 时QS.QSSTRESN的相应值（对于受试者和访视）。
ADQSADAS	ACTOT11	AVAL	分析值	浮点	3.0		ADAS第1, 2, 4, 5, 6, 7, 8, 11, 12, 13 和14项分数之和，校正缺失值的详情请参见统计分析计划的14.2节。
ADQSADAS	*ALL*	BASE	基线值	浮点	3.0		ADQSADAS.ABLFL='Y'时的

CDISC 分析数据模型 V 2.1

数据集名	参数标识符	变量名	变量标签	变量类型	显示格式	编码列表/受控术语	来源/衍生
							ADQSADAS.AVAL
ADQSADAS	*ALL*	CHG	相对于基线变化值	浮点	3.0		ADQSADAS.AVAL - ADQSADAS.BASE
ADQSADAS	*ALL*	ABLFL	基线记录标识	文本	\$1	Y	如果记录包含基线值，即如果AVISITN=3时值为Y；否则为空。
ADQSADAS	*ALL*	TRTP	计划治疗	文本	\$20	安慰剂、咕诺美林低剂量、咕诺美林高剂量	ADSL.TRT01P
ADQSADAS	*ALL*	TRTPN	计划治疗(N)	整型	1.0	0=安慰剂、1=Xanomeline低剂量、2=Xanomeline高剂量	ADSL.TRT01PN
ADQSADAS	*ALL*	TRTDOSE	随机的每日剂量规格，毫克	整型	2.0	0=安慰剂、54= Xanomeline低剂量、81= Xanomeline高剂量	ADSL.TRTDOSE
ADQSADAS	*ALL*	AGE	年龄	整型	3.0		ADSL.AGE
ADQSADAS	*ALL*	AGEGR1	年龄组1	文本	\$5	<65, 65-80, >80	基于ADSL.AGEGR1，如果ADSL.AGE缺失，则为空
ADQSADAS	*ALL*	AGEGR1N	年龄组1(N)	整型	1.0	1= <65, 2= 65-80, 3=>80	基于ADSL.AGEGR1N，如果ADSL.AGE缺失，则为空
ADQSADAS	*ALL*	SEX	性别	文本	\$1	M, F	ADSL.SEX
ADQSADAS	*ALL*	SAFFL	安全集人群标识	文本	\$1	Y., N	ADSL.SAFFL
ADQSADAS	*ALL*	ITTFL	意向性治疗人群标识	文本	\$1	Y., N	ADSL.ITTFL
ADQSADAS	*ALL*	ITTRFL	意向性治疗记录水平标识	文本	\$1	Y	如果观察的数据适合分析（即QS.VISITNUM值在 3,8,10,12,201中）且如果QS.VISIT = 包含在ADQSADAS.ADY

数据集名	参数标识符	变量名	变量标签	变量类型	显示格式	编码列表/受控术语	来源/衍生
							中的访视窗的名称，且如果ADQSADAS.ITTRFL='Y'，则ITTRFL='Y'；否则ITTRFL为空。
ADQSADAS	*DEFAULT*	DTYPE	衍生类型	文本	\$4		不适用，所以为空。
ADQSADAS	ACTOT11	DTYPE	衍生类型	文本	\$4	LOCF	如果ADQSADAS.AVAL（因而整条记录）为LOCF算法填补的值，则DTYPE = 'LOCF'，否则DTYPE为空。
ADQSADAS	*ALL*	ONTRTFL	治疗期记录标帜	文本	\$1	Y	如果ADQSADAS.TRSDT<=ADQSADAS.ADT<=ADQSADAS.TRTEDT，则ONTRTFL='Y'，否则ONTRTFL为空。
ADQSADAS	*ALL*	TRSDT	治疗开始日期	整型	yymmdd10.		ADSL.TRSDT
ADQSADAS	*ALL*	TRTEDT	治疗结束日期	整型	yymmdd10.		ADSL.TRTEDT
ADQSADAS	*ALL*	VISITDY	访视的计划研究日	整型	3.0		QS.VISITDY
ADQSADAS	*ALL*	VISITNUM	访视编号	浮点	4.1		QS.VISITNUM
ADQSADAS	*ALL*	ADT	分析日期	整型	yymmdd10.		与AVAL关联的QS.QSDTC，转变为SAS日期

5.3 分析结果元数据

这些元数据提供从用于统计展示的结果到分析数据集中数据的可溯源性。分析结果元数据不是必需的，但最好的实践是提供它们，通过确定关键分析、提供结果、文档和数据集之间的关联和记录进行的分析以给审阅者提供帮助。

分析结果包括统计展示（如结果的文本、表、图展示）或推断说明，比如p值或者治疗效应的估计。分析结果元数据以标准格式和可预见的位置提供了分析结果与用于产生它的数据之间的链接，这样审阅者可从一个分析结果链接到描述分析的重要信息，如进行分析的原因、用于产生分析的数据集和选择标准。

并不是临床研究报告或递交中的全部分析都需要分析结果元数据，甚至也不建议这么做。由申办者决定哪些分析应该具有分析结果元数据。例如，申办者出于市场声明的考虑，可能选择只对主要疗效分析和次要疗效分析提供分析结果元数据。

分析结果元数据描述临床研究报告或递交中特定的分析结果的主要属性。用于描述分析结果的元数据字段列在表 5.3.1中。使用单词“展示”而不是“表”，是因为“展示”更通用，可以指结果的表或图展示。

表5.3.1 分析结果元数据字段

分析结果元数据字段	描述
展示标识符	对于特定的分析展示的唯一标识符（如表或图的编号）。
展示名称	展示的标题，如果需要描述和确认展示（如分析人群），则包括更多信息。
结果标识符	确定展示内的特定分析结果。例如，如果一个展示有多个P值，分析结果元数据专门指其中的一个，本字段标识感兴趣的P值。当与展示标识符组合时，它提供了一个特定分析结果的唯一标识。
参数	BDS分析数据集中的分析参数，是分析结果的焦点。如果结果不是基于BDS分析数据集，则不适用。
参数编码	对应于BDS分析数据集中的PARAM变量。如果结果不是基于BDS分析数据集，则不适用。
分析变量	分析的分析变量。
原因	进行这个分析的基本原理。它指出分析是何时计划的（例如，“方案中提前规定的”、“统计分析计划中提前规定的”、“数据驱动的”、“监管机构要求的”）和证据体中分析的目的（例如，“主要疗效”、“关键次要疗效”、“安全性”）。原因的一个实例是“方案中规定的主要疗效分析”。
数据集	用于产生分析结果的数据集的名称。在大多数的情形下，这是一个单独的数据集。但是如果使用多个数据集，则都会列在这里。
选择标准	特定的和充分的分析子集及/或分子的选择标准 -用于确定为分析所选记录的变量及其值的完整列表。虽然语法不是ADaM特有的，期望信息可以很容易的包括在一个WHERE从句或类似的东西中，以保证选择适合分析的记录的精确集合。如果分析不包括分析数据集中的全部记录，则本信息是必需的。
文档	进行的分析的文字描述。本信息可能是文字描述、伪编码、或到另外文档的链接，如方案或者统计分析计划，或者到分析产生程序的链接（即用于产生分析结果的统计软件程序）。文档元数据元素的内容，取决于用于描述分析本身要求的详细水平，包含申办者是否提供相应的分析产生过程、申办者特有的要求及标准。本文档元数据元素将保持自由形式，意思是它不会变得遵循严格的结构或者受控术语。
编程说明	用于进行特定的分析的软件程序代码。这包括，例如模型说明（使用特定的变量名）和再现分析需要的所有的技术说明（例如协方差结构）。名称和适用的软件包的版本应该标明，可以作为此元数据元素的一部分，或者放在另一个文档中，如审阅者指南（关于审阅者指南的更多信息，参见附录B）。

5.3.1 分析结果元数据举例

图5.3.1.1和图5.3.1.2包含5.1.1和5.2.2节中描述的用ADAS-Cog分析数据集进行的分析的数据展示。就像在统计分析计划中描述的，第24周ADAS-Cog(11)总分的主要分析，使用疗效人群，对于24周时的缺失数据使用LOCF方法填补。应用协方差分析(ANCOVA)分析模型，纳入基线分数、中心群和治疗（作为一个连续变量）作为自变量，并且包含一个剂量反应的检验结果。治疗组二二比较使用ANCOVA模型进行，纳入基线分数、中心群和治疗（作为分类变量）作为自变量，提供治疗差异的结果。另外，进行了对于

ADAS-Cog的一个支持性的分析，使用混合效果模型重复测量(MMRM)分析。在本例中，疗效人群是意向性治疗人群。

图5.3.1.1 主要终点分析统计展示的示例

Protocol: CDISCPILOT01
Population: Efficacy

Page 1 of 1

Table 14-3.01
Primary Endpoint Analysis: ADAS Cog (11) - Change from Baseline to Week 24 - LOCF

	Placebo (N=79)	Xanomeline Low Dose (N=81)	Xanomeline High Dose (N=74)
Baseline			
n	79	81	74
Mean (SD)	24.1 (12.19)	24.4 (12.92)	21.3 (11.74)
Median (Range)	21.0 (5;61)	21.0 (5;57)	18.0 (3;57)
Week 24			
n	79	81	74
Mean (SD)	26.7 (13.79)	26.4 (13.18)	22.8 (12.48)
Median (Range)	24.0 (5;62)	25.0 (6;62)	20.0 (3;62)
Change from Baseline			
n	79	81	74
Mean (SD)	2.5 (5.80)	2.0 (5.55)	1.5 (4.26)
Median (Range)	2.0 (-11;16)	2.0 (-11;17)	1.0 (-7;13)
p-value(Dose Response) [1] [2]			0.245
p-value(Xan - Placebo) [1] [3]		0.569	0.233
Diff of LS Means (SE)		-0.5 (0.82)	-1.0 (0.84)
95% CI		(-2.1;1.1)	(-2.7;0.7)
p-value(Xan High - Xan Low) [1] [3]			0.520
Diff of LS Means (SE)			-0.5 (0.84)
95% CI			(-2.2;1.1)

[1] Based on Analysis of covariance (ANCOVA) model with treatment and site group as factors and baseline value as a covariate.
[2] Test for a non-zero coefficient for treatment (dose) as a continuous variable.
[3] Pairwise comparison with treatment as a categorical variable: p-values without adjustment for multiple comparisons.

Source: C:\cdisc_pilot\PROGRAMS\DRIFT\TFL\rtf_eff1.sas 21:05 Monday, June 26, 2006

表5.3.1.1和表5.3.1.2阐释了图5.3.1.1中显示的ADAS-Cog的特定元素的分析结果元数据。在图中有下划线的地方最好超链接到临床研究报告中的数据展示、定义文件中其它地方的元数据和SAP中的特定页。

表5.3.1.1举例说明了剂量反应分析的分析结果元数据，根据图5.3.1.1 (1) 确定。它也阐明了所做分析的描述，程序说明中没有提供的模型说明。

表5.3.1.1 图5.3.1.1中统计展示的剂量反应分析的分析结果元数据

元数据字段	字段的定义	元数据
展示标识符	特定分析展示的唯一标识符	Table 14-3.01
展示名称	展示标题	主要终点分析: ADAS Cog (11) – 从基线到第24周的变化 – LOCF
结果标识符	确定展示内的具体分析结果	剂量反应分析
参数	参数	ADAS-Cog (11)总分
参数编码	参数代码	ACTOT11
分析变量	分析的分析变量	CHG
原因	进行的分析的基本原理	方案中提前规定的主要的疗效分析
数据集	用于分析的数据集	ADQSADAS
选择条件	对于分析子集及/或分子的特定的和充分的选择条件	IITFL='Y' and AVISIT='第24周' 和 PARAMCD='ACTOT11'
文档	进行的分析的文本描述	SAP第10.1.1节。第24周ADAS-Cog(11)总分相对基线变化的剂量反应的线性模型分析 – 缺失值用LOCF方法填补，疗效人群。使用SAS中的PROC GLM产生P值（来自治疗剂量Type III SS）；模型中自变量是 TRTDOSE（随机化剂量：安慰剂是0；54是低剂量；81是高剂量），SITEGR1（中心群作为类变量）和BASE（基线ADAS-Cog分数）。
编程说明	用于进行分析的分析语法	

表5.3.1.2举例说明了治疗组间二二比较的分析结果元数据，根据图5.3.1.1中的（2）确定。它也阐释了模型说明的纳入条件。

表5.3.1.2 图5.3.1.1的统计展示中成对治疗比较的分析结果元数据

元数据字段	字段定义	元数据
展示标识符	特定分析展示的唯一标识符	Table 14-3.01
展示名称	展示标题	主要终点分析：ADAS Cog (11) – 第24周相对基线变化 – LOCF
结果标识符	识别展示中的特定分析结果	治疗组二二比较
参数	分析参数	ADAS-Cog (11)总分
参数编码	分析参数代码	ACTOT11
分析变量	分析的分析变量	CHG
原因	分析编程的基本原理	方案中提前规定的主要疗效分析
数据集	用于分析的数据集	ADQSADAS
选择判据	分析子集及/或分子的特定的和充分的选择条件	ITTFL='Y' and AVISIT='第24周' and PARAMCD='ACTOT11'
文档	进行的分析的文字描述	ADAS-Cog(11)总分线性模型分析：第24周相对基线变化的治疗组二二比较和校正的均值；缺失值填补使用LOCF方法，疗效人群。使用随机化治疗作为分类变量、中心群作为分类变量、基线ADAS-Cog评分纳入模型。
编程说明	用于进行分析的分析语法	PROC GLM; CLASS SITEGR1 TRTP; MODEL CHG = TRTP SITEGR1 BASE; ESTIMATE 'H VS L' TRTP 0 1 -1; STIMATE 'H VS P' TRTP -1 1 0; ESTIMATE 'L VS P' TRTP -1 0 1; LSMEANS TRTP / OM STDERR PDIF CL; RUN;

Figure 5.3.1.2 支持性分析统计展示示例

Protocol: CDISCPILOT01
Population: Efficacy

Page 1 of 1

Table 14-3.11
ADAS Cog (11) - Repeated Measures Analysis of Change from Baseline to Week 24

	Placebo (N=79)	Exomeline Low Dose (N=81)	Exomeline High Dose (N=74)
LS Means (SE)	1.6 (0.49)	1.5 (0.52)	1.1 (0.56)
p-value(Xan - Placebo)		0.955	0.556
Diff of LS Means (SE)		-0.0 (0.70)	-0.4 (0.72)
95% CI		(-1.4;1.3)	(-1.9;1.0)
p-value(Xan High - Xan Low)			0.606
Diff of LS Means (SE)			-0.4 (0.75)
95% CI			(-1.9;1.1)

Note: The change from baseline is calculated as the post-baseline score minus the baseline score. The covariates included in the MMRM model are treatment, site group, time and treatment by time interaction, baseline ADAS-Cog (11) score, and baseline ADAS-Cog (11) score by time interaction.
Source: C:\cdisc_pilot\PROGRAMS\DRIFT\TFLS\rtf_eff_mmrn.sas 21:06 Monday, June 26, 2006

表5.3.1.3举例说明图5.3.1.2中的统计展示的分析结果元数据，阐明了使用元数据来描述一个展示。在示例中有下划线的项，最好超链接到临床研究报告中的数据展示、定义文件中其它地方的元数据和SAP的特定页。

表5.3.1.3 图5.3.1.2中统计展示的分析结果元数据

元数据字段	字段定义	元数据
展示标识符	特定分析展示的唯一标识符	<u>Table 14-3.11</u>
展示名称	展示标题	ADAS Cog (11) – 第24周相对基线变化的重复测量分析
结果标识符	识别展示内的特定分析结果	
参数	分析参数	ADAS-Cog (11)总分
参数编码	分析参数代码	ACTOT11
分析变量	分析的分析变量	CHG
原因	进行的分析的基本原理	SAP中提前规定的
数据集	用于分析的数据集	<u>ADQSadAS</u>
选择条件	分析子集和/或分子的特定和充分选择条件。	ITTFL='Y' AND AVISITN GT 0 AND DTYPE NE 'LOCF' AND PARAMCD='ACTOT11'
文档	进行的分析的文字描述	SAP第10.1.1节：第24周相对基线变化的校正均数和第24周治疗组间二二比较，所用的复测量模型纳入了治疗组（作为分类变量），中心（作为分类变量），时间，治疗*时间的交互作用，基线分数，基线*时间的交互作用，使用的是非结构化协方差矩阵。疗效数据，观察到的病例数据。
编程说明	用于进行分析的分析语法	PROC MIXED; CLASS USUBJID SITEGR1 AVISITN TRTP; MODEL CHG = TRTP SITEGR1

		AVISITN TRTP*AVISITN BASE BASE*AVISITN / OUTP=PRED DDFM=KR; REPEATED AVISITN / SUBJECT=USUBJID TYPE=UN; LSMEANS TRTP / DIFF CL; RUN;
--	--	---

450

451

6 受试者水平分析数据集

受试者水平分析数据集（ADSL）的数据结构不受临床试验设计类型的影响，即每个受试者有且只有一条记录。ADSL用于提供描述受试者特性的变量。这一结构允许其简单地与任何其它数据集合并，包括SDTM和分析数据集。

监管机构人员声称ADSL在临床试验的审阅中非常有帮助。即便没有其它分析数据集递交，ADSL和它的相关元数据在任何基于CDISC的临床试验数据递交时也是必需的。

ADSL的意图是提供关于受试者的描述信息，它被用于多种类型的分析，包括描述性、分类和建模。不应为了最小化分析数据集的数目，而强求用ADSL来支持所有的分析。虽然从技术的角度，把研究中的各个数据点拿出来作为受试者水平的数据集如ADSL中的变量是可行的，但那不是ADSL的意图和或目的。关键终点和在研究过程中随时间变化数据的正确位置，是在一个BDS数据集中。

ADSL是其它分析数据集中的受试者水平的变量的主要来源，比如人群标识和治疗组变量。当从ADSL合并数据到其它分析数据集时，只需要包括那些与这些分析数据集相关的字段。包括太多的无关变量（即支持分析不需要的变量）会使得用户更难找到重要变量，并妨碍清晰而简明的交流。

表6.1提供了一个ADSL分析数据集元数据的例子。

表6.1 ADSL分析数据集元数据的例子¹

数据集名称	数据集描述	数据集位置	数据集结构	数据集的关键变量	数据集类别	文档
ADSL	受试者分布、人口特征和基线特征	<u>adsl.xpt</u>	每个受试者一条记录	USUBJID	ADSL	SAP, DSADSL.SAS

包括在ADSL中的变量的最小集合依赖于疾病的特定性质和方案（更详细的清单参考ICH E3[8]；更多的描述，包括必需的变量，参考ADaMIG）。ADSL信息的例子包括（但不限于）：

- 人口统计学变量（例如年龄、性别、种族、其它相关因素）
- 疾病因素（例如疾病开始、疾病严重程度）
- 治疗代码/组
- 其它可能影响治疗反应的预后因素（例如吸烟、饮酒、女性的月经状况）
- 重要事件的日期（例如治疗开始和结束日期）
- 研究人群

ADSL包含描述临床试验受试者治疗前信息的变量，或者一些为了分析的目的以某种方式将受试者分组的变量。

总的来说，ADSL中的变量包括那些描述性的、重要的基线特征、用于随机的分层、用于识别受试者特定子群归属（如人群标识）或者识别重要事件的发生时间或重要事件是否发生（例如末次给药日期、死亡、中止）。例如，按年龄组分层随机时，受试者的年龄分类是研究受试者的重要描述变量，包含在ADSL中。

ICH指南（ICH E3，11.2节）[8]推荐“除了基线变量分组数据的表和图，相关的个体受试者人口统计学和基线数据...所有随机化受试者（按治疗和中心或者多中心研究分开）应该在受试者表格清单中报告”。通常FDA审阅者和申办者赞同受试者水平的数据递交满足这个要求。如果是这种情况，ADSL应该包括那些需要用来满足注册指南的变量。

6.1 不分析的受试者数据

分析数据集是否包括不分析受试者的数据（例如，筛选失败者），取决于申办者，并且应该与数据的审阅者或者用户交流。如果包括这些数据，则包括在适当的数据集中，如ADSL（而不是为不分析受试者提供单独的数据集）并使用适当的标识变量清楚地区分这些记录。元数据必须说明包括了这些数据以及如何区分它们。

¹元数据的陈述应当由申办者与接受者决定。这一例子仅用于描述内容而不是形式。

附录

附录A 文献

- 1) CDISC Analysis Data Model (ADaM) Team, 2009, "ADaM Implementation Guide," available on CDISC website at <<http://www.cdisc.org/standards>>
- 2) CDISC Define.xml Team, 2005, "Case Report Tabulation Data Definition Specification (define.xml)," available on CDISC website at <<http://www.cdisc.org/standards>>
- 3) CDISC SDS Metadata Team, 2007, "Metadata Submission Guidelines, Appendix to the Study Data Tabulation Model Implementation Guide 3.1.1," available on CDISC website at <<http://www.cdisc.org/standards>>
- 4) CDISC SDS Team, 2008, "Study Data Tabulation Model (SDTM) Implementation Guide Final Version 3.1.2," available on CDISC website at <<http://www.cdisc.org/standards>>
- 5) CDISC Submission Data Standards (SDS) Team, 2008, "Study Data Tabulation Model (SDTM) Final Version 1.2," available on CDISC website at <<http://www.cdisc.org/standards>>
- 6) FDA, 2008, "Guidance for Industry: Providing Regulatory Submissions in Electronic Format – Human Pharmaceutical Product Applications and Related Submissions Using the eCTD Specifications," available on FDA website at <<http://www.fda.gov/Drugs/GuidanceComplianceRegulatoryInformation/Guidances/ucm064994.htm>>
- 7) FDA Center for Drug Evaluation and Research (CDER), 2009, "Study Data Specifications, Version 1.5," available on FDA website at <<http://www.fda.gov/ForIndustry/DataStandards/StudyDataStandards/default.htm>> >
- 8) ICH Expert Working Group, 1995, "ICH Harmonised Tripartite Guideline: Structure And Content of Clinical Study Reports - E3," available at <<http://www.ich.org/LOB/media/MEDIA479.pdf>>

附录B 定义

ADaM – CDISC分析数据模型

ADaM基本数据结构(BDS) – 为了方便分析和审阅而设计的一种数据集结构，数据结构为每个受试者每个分析参数每个分析时间点一条或多条记录。根据分析的需要，有些情况下要求有分析时间点。BDS支持大多数的分析，ADaMIG对其进行了描述。

ADaM实施指南(ADaMIG) – 具体说明ADaM标准数据集结构和变量的文档，包括命名规则。它也具体指出了对于实施问题的标准解决办法。ADaM文档和ADaMIG应当一起使用。

分析数据集 – 用于统计分析和报告的数据集。

分析数据集生成程序 – 用于生成分析数据集的计算机程序。

分析数据集元数据 – 描述分析数据集结构、内容和衍生的信息。

分析结果产生程序 – 用于产生分析结果（例如，包含在表或图中的总结和统计推断）的计算机程序。

分析参数(PARAM) – 一种行标识符，以唯一标识共用相同定义的一组值。例如：主要疗效分析参数是“3分钟坐位收缩压（mmHg）”。注意ADaM分析参数包含唯一标识一组相关分析值所需全部信息。相反，为了确定一组相关值，SDTM --TEST列可能需要与修饰语列，如--POS, --LOC, --SPEC等等结合起来。在本文档，“参数”是“分析参数”的同义词。

分析参数值水平元数据 – 描述一个或一组分析参数中分析值的信息。

分析结果元数据 – 描述包含在一个临床研究报告或递交中特定分析结果的信息。

分析时间点 – 一种行标识符，用于把分析参数中的值分类到用于分析的时间或概念组中。这些分组可能是观测的、计划的或衍生的。例如：主要的疗效分析在第2周、第6周和终点分析时间点进行。

分析值 – (1) 是分析参数描述的字符(AVALC)或数字(AVAL)。分析值可能是输入数据的值、输入数据值的分类或衍生的。例如：参数“平均心率（次/分钟）”的分析值是根据每次访视三次测量的平均值衍生。(2) 另外，特定功能的值视为分析值。例如：基线值（BASE）、相对基线的变化(CHG)。

分析变量元数据 – 描述分析数据集中变量的信息。

定义文件 – 如在病例报告制表数据定义说明中所述[2]，1999年FDA电子递交(eSub)指南和电子通用技术文档(eCTD)的文档指出，递交应提供所包括数据内容和结构描述的文档。该文档名为数据定义文档（例如在1999的指南中叫作“define.pdf”）。数据定义文档提供递交的数据集清单以及各数据集内容的详细描述（即元数据）。为了增加自动化的水平，改进监管审阅过程的效率，define.xml能被用于提供机器可读格式的数据定义文档，其正式名称是病例报告制表数据定义(CRT DD)说明。SDTM和ADaM数据集各有相应的定义文件。

元数据 – 关于数据的信息或数据。

记录 – 数据集中的行。

审阅者指南 – 可以包含在递交中的文档，指引审阅者到递交材料的各个方面。根据参与项目的FDA的审阅者建议，审阅者指南包含在CDISC SDTM/ADaM预试验项目递交材料中。（项目的报告可以在www.cdisc.org得到。）本文档对于提供太复杂或太冗长以至于不容易在其它来源的元数据中描述的信息非常有用。它可用于描述不容易在变量水平交流的问题，或者应用到多个分析的数据集，例如申办者命名规则、不完整日期的填补规则。它不应重复制大量的能在其它来源的元数据中找到的信息。

SDTM – 研究数据制表模型 – 由CDISC递交数据标准(SDS)撰写的文档，描述递交到监管机构的用于表达临床研究数据的一般概念模型。SDTM为描述从临床试验收集并递交到监管机构的信息的组织提供了一般性的框架[5]。

SDTM实施指南(SDTMIG) –由CDISC SDS小组编写的文档，意图是指导递交到注册机构如FDA的标准临床试验制表数据集的组织、结构和格式。它提供了特定的域模型、假设、商业规则和准备基于SDTM的标准制表数据集的例子[4]。

可溯源性 – ADaM的特性，允许一个分析数据集的用户理解数据的传承或元素和它的前身之间的关系。可溯源性促进了透明度，这对于建立对结果或结论的信心方面是必不可少的组成部分。最终ADaM的可溯源性允许理解分析结果、分析数据集与SDTM域之间的关系。可溯源性是通过建立元素和它的直接前身之间的清晰路径来建立。完整的路径通过从一个元素到它的前身来追踪，然后再到前身的前身等等，回溯到SDTM域，最终到达数据收集工具。注意CDISC临床数据采集标准协调(CDASH)标准和SDTM是相协调的，因此有助于保证端到端的可溯源性。

例如：基于元数据和分析数据集的内容，审阅者能够追踪主要和次要疗效分析值是如何从每个受试者的SDTM数据衍生的。

变量 –数据集列中的列。

附录C 缩写和简略词

下表是在本文档中多次使用的缩写词和简略词，这里没有包括各SDTM域的解释（例如QS, DM），以及对所引用的变量的描述。

表C.1 文档中所用的缩写和简略词

缩略语	定义（英文）	定义（中文）
ADAE	ADaM Adverse Event Analysis Dataset	ADaM不良事件分析数据集
ADaM	CDISC Analysis Data Model	CDISC分析数据模型
ADaMIG	Analysis Data Model Implementation Guide	分析数据模型实施指南
ADAS-Cog	Alzheimer's Disease Assessment Scale - Cognitive Subscale	阿尔茨海默病评定量表-认知部分
ADSL	ADaM Subject-Level Analysis Dataset	ADaM受试者水平分析数据集
ANCOVA	Analysis of Covariance	协方差分析
BDS	ADaM Basic Data Structure	ADaM基本数据结构
BMI	Body Mass Index	体重指数
CDASH	Clinical Data Acquisition Standards Harmonization	临床数据获取协调标准
CDISC	Clinical Data Interchange Standards Consortium	临床数据交换标准协会
CFR	Code of Federal Regulations	美国联邦法规
CRT	Case Report Tabulation	病例报告列表
CRT-DDS	Case Report Tabulation Data Definition Specification	病例报告列表- 数据定义说明
eCTD	electronic Common Technical Document	电子通用技术文档
FDA	United States Food and Drug Administration	美国食品药品监督管理局
ICH	International Conference on Harmonisation	人用药品注册技术要求国际协调会
ITT	Intent-to-Treat	意向性治疗
LOCF	Last Observation Carried Forward	末次观测结转
MMRM	Mixed Effects Models Repeated Measures	重复测量混合效应模型
PDF	Portable Document Format	便携文件格式
SAP	Statistical Analysis Plan	统计分析计划
SDS	Submission Data Standards	递交数据标准
SDTM	Study Data Tabulation Model	研究数据列表模型
SDTMIG	Study Data Tabulation Model Implementation Guide	研究数据列表模型实施指南
XML	Extensible Markup Language	可扩展标记语言
XPT	Filename extension for a SAS transport file	SAS传输文件的文件扩展名

附录D 对即可分析的阐释

为了阐释即可分析的概念，再次考虑5节描述的例题。在图5.3.1.1，剂量反应分析和剂量二二比较都包含在展示中。即可分析的意思不是指在一个单独的统计程序能够产生这一格式化的表格，而是指使用相应的分析数据集作为输入，通过运行一个标准的统计程序（例如SAS过程步、S-PLUS函数等等），能够重现表中的所有统计量。这意味着审阅者只用最小编程努力就能够重现和探究这些结果，允许审阅者把注意力集中于结果而不是程序。

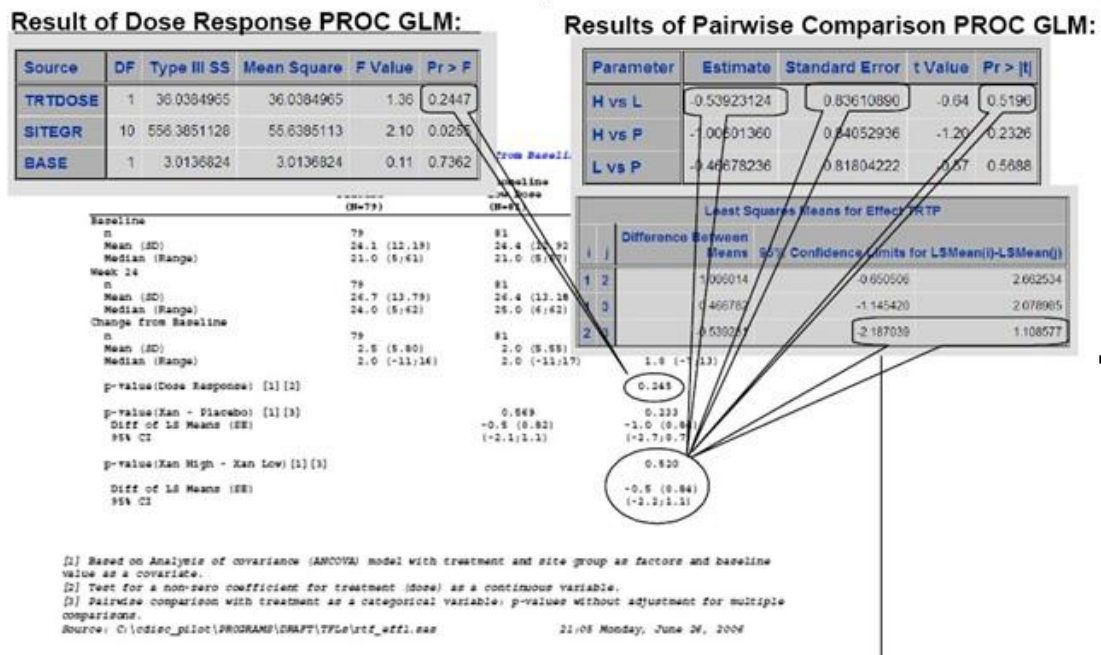
例如，下面的SAS代码使用包含相应变量的分析数据集重现了表14-3.01（图5.3.1.1）剂量反应分析结果。请注意使用where从句选择了适当的记录进行分析。

```
*** DOSE RESPONSE ANALYSIS ***;
PROC GLM DATA=A.ADQSADAS(WHERE=(ITTFL='Y' AND AVISIT='Week 24' AND
PARAMCD='ACTOT11'));
CLASS SITEGR1;
MODEL CHG = TRTDOSE SITEGR1 BASE;
RUN;
```

相似的，下面SAS代码重现了剂量二二比较结果。

```
*** PAIRWISE DOSE COMPARISON ANALYSIS ***;
PROC GLM DATA=A.ADQSADAS (WHERE=(ITTFL='Y' AND AVISIT='Week 24' AND
PARAMCD='ACTOT11'));
CLASS SITEGR1 TRTP;
MODEL CHG = TRTP SITEGR1 BASE;
ESTIMATE 'H VS L' TRTP 0 1 -1;
ESTIMATE 'H VS P' TRTP -1 1 0;
ESTIMATE 'L VS P' TRTP -1 0 1;
LSMEANS TRTP / OM STDERR PDIF CL;
RUN;
```

图D.1阐释如上SAS代码的结果与相应的结果展示元素之间的关系（图5.3.1.1中表14-3.01）。



图D.1 阐释SAS输出 vs 来自即可分析数据集的结果展示

附录E 复合终点的例子

如在4.1.1节所言，分析举例只需要ADSL和SDTM，无法覆盖所有的分析，特别是关于疗效分析的时候。甚至全范围的安全分析，通常会包含的更复杂的例子，仅基于ADSL和SDTM的分析数据是不能覆盖的。这个例子阐释了具有复杂支撑的一个明显的简单的二分类结果变量（一次头痛发作的治疗结果），及从不同的来源数据集提取数据元素。它描述了一个复合终点，需要来自从疗效数据集（不同时间点的头痛严重性）以及不良事件和合并用药数据集的数据。终点是“持续无偏头痛症状”。

终点（持续无偏头痛和症状，基于国际头痛协会指导原则）定义为：

1. 基线时头痛严重程度为中、重度，且
2. 给药后（即首剂试验药物后）2小时头痛严重程度为“无疼痛”，且
3. 给药后48小时内无头痛复发，且
4. 从首剂到给药后48小时期间未使用急救性镇痛或止吐药，且
5. 给药后2到48小时没有相关的症状（恶心、呕吐、畏光、畏声）。

上面文本中的**AND**说明，必须符合所有的条件才能考虑一个受试者已经经历了反应。

对于这个例题，使用如下定义和说明书：

头痛严重程度

头痛严重性由受试者在既定时间点的主观评估（给药后基线、0.5、1、1.5、2、3和4小时），等级是0（无疼痛）到3（最严重的疼痛）。

伴随症状

受试者记录是否如下相关症状在相应的时间点存在或消失（给药后的基线、2和4小时）：畏光、畏声、恶心、呕吐。另外，受试者被教导列出任何如上症状，作为日记卡中的“不良症状”，如果（1）在他们接受测试药物之后强度显示了不同寻常的增强，或者（2）否则是在他们接受测试药物之后，和他们平常的偏头痛症状相比，在特征方面显示了一个重要的变化。因此，研究者将记录所有这样的症状作为不良经历。因此，相关症状消失的完全评价将包括一个对于不良事件数据集的审查。

头痛复发

头痛复发被定义为在接受试验药物后两小时宣布头痛减轻（轻度或没有疼痛）在接受测试药后48小时返回到严重程度2或3（中度或严重）。第一次接受药物之后2到24小时之间和第一次接受药物之后24到48小时之间受试者记录最大的头痛严重程度。

急救用药

受试者记录接受测试药之后加用的任何镇痛药/镇吐剂，记载日期、时钟时间（上午/下午）、药物名称（例如可待因）、片剂/胶囊的数量、和每个片剂/胶囊中的剂量。急救用药也定义为在接受测试药物之后48小时内接受任何附加的测试药物的剂量。急救用药的使用是根据合并用药和暴露数据集来确定的。

为了确定是否一个受试者满足持续无头痛和症状的标准，必须确定五条标准每一条的答案。终点的衍生，需要头痛严重程度和伴随症状数据（一个或多个SDTM域）、AE域、CM域、EX域、和ADSL的输入。

在表E.1阐释的BDS分析数据集中，假定对于所有标准内含问题的答案都保留在分析数据集中。因为目的是为了阐释复杂性而不是一个完整的分析数据集，在图解中只列出了几个分析数据集变量(PARAMCD、AVAL和AVALC)。另外，不是试图为此例描述特定的SDTM域和变量，而是为起源/衍生字段提供一个简单的文字描述。在“真实的”元数据中，这一元数据元素应该指向特定的域和变量，且应该包括如何确定域中的记录是数据。（例如，当对于这个USUBJID，QSCAT=xxx）。

这个例子说明了起源/衍生可能很冗长复杂。对于复杂的衍生变量，起源/衍生字段应当提供一个到外部文档的链接，解释各种数据的起源和创建变量涉及的算法。

671 表E.1 复合终点的例子 – 图解选择的变量的分析变量元数据的例子

数据集名称	参数标识符	变量名称	变量标签	变量类型	显示格式	编码列表/受控术语	起源/衍生
ADSYMFR	PARAMCD	PARAMCD	参数代码	文本	\$8	HASPNFR HASEVBL HASEV2 HARECUR HARESCUE HASYPMPD HASYPMAE	HASPNFR, “ADSYMFR.PARAM=给药后2-48小时持续无偏头痛和症状” 时 HASEVBL, “ADSYMFR.PARAM=基线头痛严重程度” 时 HASEV2, “ADSYMFR.PARAM=给药后2小时头痛严重程度” 时 HARECUR, “ADSYMFR.PARAM=当给药后48小时内头痛复发” 时 HARESCUE, “ADSYMFR.PARAM=从第一次给药到给药后48小时急救用药” 时 HASYPMPD, “ADSYMFR.PARAM=给药后2-48小时日记卡上指出的伴随症状” 时 HASYPMAE, “ADSYMFR.PARAM=给药后2-48小时不良事件数据集中指出的伴随症状” 时
ADSYMFR	*DEFAULT*	AVAL	分析值	整数	1.0	0=N 1=Y	基于 ADSYMFR.AVALC衍生, 如果 ADSYMFR.AVALC缺失则为空
ADSYMFR	HASEVBL	AVAL	分析值	整数	1.0	0=无疼痛 1=轻度疼痛 2=中度疼痛 3=严重疼痛	基线时的头痛严重程度是来自日记卡数据: 在给药时记录的头痛严重程度, 如果缺失则为空
ADSYMFR	HASEV2	AVAL	分析值	整数	1.0	0=无疼痛 1=轻度疼痛 2=中度疼痛 3=严重疼痛	给药后两小时的头痛严重程度是来自日记卡数据: 记录的给药后两小时头痛严重程度, 如果缺失则为空
ADSYMFR	*DEFAULT*	AVALC	分析值(C)	文本	\$1		空
ADSYMFR	HASPNFR	AVALC	分析值(C)	文本	\$1	N=否 Y=是	给药后2-48小时持续的无偏头痛和症状的值是基于分析数据集中其它终点。如果 (ADSYMFR.HASEVBL=2 或 3) 且 ADSYMFR.HASEV2=0 且

数据集名称	参数标识符	变量名称	变量标签	变量类型	显示格式	编码列表/受控术语	起源/衍生
							ADSYMFR.HARECUR=N 且 ADSYMFR.HARESCUE=N 且 (ADSYMFR.HASYMPD=N 且 ADSYMFR.HASYMPAE=N) 则值为Y, 否则为N。
ADSYMFR	HARECUR	AVALC	分析值(C)	文本	\$1	N=没有头痛复发 Y=头痛缺失复发了	在给药后48小时内头痛复发的值是基于日记卡。如果受试者记录的初次给药后2-24小时和24-48小时的时间段里他们的最厉害头痛严重程度是0（无疼痛）则值为N；如果两个时间段任何一个的头痛严重程度>0，则值为Y；缺失则为空。
ADSYMFR	HARESCUE	AVALC	分析值(C)	文本	\$1	N=未使用急救用药 Y=使用了急救用药	初次给药到给药后48小时内使用的急救用药是基于CM域、ADSL和EX域。如果在初次给药到给药后48小时内没有使用镇痛药或镇吐剂（CM域），且如果在这个时间段内没有使用其它的研究药物（ADSL和EX域）则为N，否则为Y。
ADSYMFR	HASYMPD	AVALC	分析值(C)	文本	\$1	N=无伴随症状 Y=有伴随症状	日记卡上指出的给药后2-48小时的伴随症状是基于给药后2到48小时畏光、畏声、恶心或呕吐是否出现。如果出现为Y，如果没出现为N。
ADSYMFR	HASYMPAE	AVALC	分析值(C)	文本	\$1	N=无伴随症状 Y=有伴随症状	给药后2-48小时AE数据集中指出的伴随症状是基于是否在AE域中发现了这些症状。如果给药后2-48小时没有畏光、畏声、恶心或呕吐则为N，如果他们在AE中报告，则为Y。

附录F 修订历史

版本2.1是分析数据模型的第二次正式发布，最初的版本是2006年8月发布的分析数据模型版本2.0。版本2.1包括修订的内容，以便对应分析数据模型实施指南（ADaMIG）的版本1.0文档。

由于进行了重大修订和格式修改，对文档的修改未能全部列出。文档中的例子进行了实质性的改动。介绍了ADaM基本数据结构。现在文档中包括了可溯源性的概念。文档的很多部分（如分析数据集变量）已经被挪到ADaMIG中。

一个重要的变化是澄清了期望SDTM作为ADaM的输入来源。

澄清了ADaM元数据，而且为了清晰性进行了扩展：

- 分析数据集元数据：删除了“目的”元数据字段。删除了一个分析数据集应该在数据集标签中具有“分析”或者“统计量”的要求。

- 分析变量元数据：起源字段的澄清。“参数值水平的元数据定义”代替值水平的元数据。

- 分析结果元数据：是分析结果元数据中的字段的重大扩展，以促进组成部分的清晰性。

当文本挪到ADaMIG中时，对分析数据集变量进行了重大扩展和修订。

附录G 声明与保证;责任范围, 及免责声明

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