**eCTD中临床试验数据库及相关资料**

**的申报要求**

**（征求意见稿）**

**国家药品监督管理局药品审评中心**

**2019年9月**

**目 录**

[一、临床试验数据库 3](#_Toc20489136)

[二、数据说明文件 3](#_Toc20489137)

[三、其他 4](#_Toc20489138)

[附件1 STF文件标签 6](#_Toc20489139)

[附件2 文件夹结构 7](#_Toc20489140)

**eCTD中临床试验数据库及相关资料的申报要求**

在eCTD申报资料中应包含临床试验数据库及相关资料，在骨架文件中应位于相应的临床总结报告之后，并使用适当的标签进行标识。

**一、临床试验数据库**

申请人应递交临床试验的原始数据库及分析数据库。原始数据库通常应包含从病例报告表（CRF）和外部文件中直接收集的所有原始数据。分析数据库为统计分析中衍生新建的数据库，应包含所有分析时所需的变量，包括衍生变量，且所有衍生变量均应能通过递交的原始数据库生成。递交的数据库应为SAS XPORT（Version 5或以上版本）传输格式（XPT格式）。

如果申请人以临床数据交换标准协会（CDISC）标准递交数据，则可将研究数据表格模型（SDTM）数据库视为原始数据库，分析数据模型（ADaM）数据库视为分析数据库。

如果数据库中单个数据集文件大小超过4G，则需将其进行拆分，并就拆分规则与结果在数据说明文件与数据审阅说明中分别阐述说明。

在eCTD申报时，原始数据库使用“data-tabulation-dataset-legacy”或“data-tabulation-dataset-sdtm”标签进行标记，分析数据库使用“analysis-dataset-adam”或“analysis-dataset-legacy”标签进行标记。标签的使用见附件1。

**二、数据说明文件**

申请人应递交原始数据库和分析数据库的数据说明文件。数据说明文件中应说明数据库中所包含数据集的基本信息，包括数据集的名称、内容简述、结构等。数据说明文件中还应说明各数据集中所包含变量的变量名、标签、数据格式、取值范围、数据来源等。对于衍生变量，需要明确具体的衍生方法，包括对衍生规则的描述及相应编程程序代码。

数据说明文件一般为XML（包括对应的XSL文件）或PDF格式。除此之外，还建议递交PDF格式的数据审阅说明。数据审阅说明是通过另一独立文档描述的方式来进一步帮助审评人员更准确、高效的理解与使用所递交数据。

在eCTD申报时，原始数据库数据说明文件使用“data-tabulation-data-definition”标签进行标记，分析数据库数据说明文件使用“analysis-data-definition”标签进行标记。

**三、其他**

申请人应递交临床试验报告相关的注释CRF（aCRF），格式为PDF。注释CRF是对空白CRF的标注，记录CRF各数据项的位置及其在相对应的数据库中的变量名和编码。

申请人应递交关键编程程序代码（如分析数据库的衍生过程和主要分析结果的生成过程等），所递交的程序代码应可读性强，不包含复杂的外部程序调用，采用TXT纯文本格式。

在eCTD申报时，注释CRF使用“annotated-crf”标签进行标记，编程程序代码使用“analysis-program”标签进行标记。

**附件1 STF文件标签**

|  |  |
| --- | --- |
| **标签** | **说明** |
| data-tabulation-dataset-legacy | 原始数据库（非CDISC标准） |
| data-tabulation-dataset-sdtm | 原始数据库（CDISC标准） |
| data-tabulation-data-definition | 原始数据库数据说明文件、数据审阅说明（若有） |
| analysis-dataset-adam | 分析数据库（CDISC标准） |
| analysis-dataset-legacy | 分析数据库（非CDISC标准） |
| analysis-data-definition | 分析数据库数据说明文件、数据审阅说明（若有） |
| annotated-crf | 注释CRF |
| analysis-program | 编程程序代码 |

**附件2 文件夹结构**

