

## 附件 2

# 体外诊断试剂临床试验数据递交要求 注册审查指导原则

## 一、前言

体外诊断试剂临床试验数据是评价产品安全有效性的重要支持性资料之一。规范地收集、整理、分析和递交临床试验数据有助于提高临床试验实施和管理质量，同时有利于监管机构快速、高效地掌握临床试验的开展情况，提高审评效率。

为指导注册申请人规范递交体外诊断试剂临床试验数据及相关资料，以便更好地开展临床试验资料审评相关工作，制定本指导原则。

本指导原则是在现行法规和标准体系以及当前认知水平下制定的，随着法规和标准的不断完善，以及科学技术的不断发展，本指导原则的相关内容也将进行相应的调整。

## 二、适用范围

本指导原则适用于以产品注册为目的开展的体外诊断试剂临床试验，包括在境外开展的体外诊断试剂临床试验。本指导原则仅涉及临床试验数据递交相关内容，不涉及临床试验过程中数据管理相关要求。

## 三、基本原则

### （一）真实原则

所递交的临床试验数据应与临床试验中所有的原始记录保持一致。

## （二）可追溯原则

应能够根据注册申请人所提交的数据、说明性文件和程序代码（如有），可从原始数据库重现形成分析数据库、临床试验报告中的统计结果，且形成的分析数据库和统计分析结果与注册申请人提交的内容一致。

注册申请人提交的临床试验数据库，应可溯源至临床试验中的原始试验记录、病例报告表，临床背景信息应可追溯至临床试验机构的病例相关信息管理系统或原始记录。原始试验记录可能包括样本的筛选入选表、样本编盲表、检测记录表等，如申报产品需要配套仪器进行检测，则还应包括配套仪器上的电子记录。以上原始试验记录不要求递交，应妥善保管备查。

## （三）数据全面且可读原则

所提交的临床试验数据根据申报产品的特点以及临床试验设计的不同，应保证数据全面、可读，且易于统计。数据库结构清晰、注释详尽，便于审阅。

## 四、临床试验数据库相关内容要求

一般来说，临床试验数据库相关内容应包括受试者相关信息以及临床试验样本的检测信息。不同的临床试验设计类型所应包含的数据信息亦不相同。根据产品特点和产品性能评价需要，体外诊断试剂临床试验包括不同的临床试验设计类型，根据在临床试验过程中试验体外诊断试剂检测结果对受试者的影响，一般分为观察性研究和干预性研究。其中观察性研究根据临床试验中所检测的不同时间点又可分为横断面研究和纵向研究。横断面研究是体外诊断试剂临床试验最常见的设计类型，其临床试验数据库内容的要求亦是体外诊断试剂临床试验数据库内容的通用要求。

对于纵向研究以及干预性研究应在通用要求的基础上再补充相应的数据。

### （一）来自于横断面研究的数据

临床试验中的入组人群、样本类型等均应与申报产品所声称的预期用途保持一致。因此数据信息中应包括受试者相关信息，包括临床诊断背景信息、样本类型、人口学信息（性别、年龄等）等。临床诊断背景信息包括临床诊断结果、相关的症状体征、以及诊疗信息（如需要）等。如申报产品的统计分析中需要进行亚组统计，则应包括与亚组划分相关的信息，如疾病的不同分期、不同进程等信息。

临床试验中关于临床样本的检测信息主要包括：试验体外诊断试剂的检测结果、对比方法检测结果等。对于根据所确定的阳性判断值来判定检测结果的产品，数据信息还应包括试验体外诊断试剂和对比方法的详细检测数值（如 Ct 值、S/CO 值等）。核酸检测类产品如涉及不同检测通道的，各通道的检测数值均应提供，包括内标的检测数值。

临床试验数据应真实且可追溯，数据集中应有唯一可溯源的去标识化样本编号，该样本编号应能够溯源至该病例的所有背景信息，如病例编号、诊疗信息等。

如临床试验中涉及复测等情况，相应的数据集中应有初测及复测的数据，并备注复测原因。如有其他需要说明的信息，可增加一列“备注”，将信息增加至“备注”中。

### （二）来自于纵向研究的数据

除横断面研究外，有些产品需要进行纵向研究，对于纵向研究的数据，数据集应包括每个病例的每个时间点的具体数据，应

尽量将同一病例的每一时间点的数据汇总，采样时间点和检测结果应对应列出。对于此种临床试验设计，会涉及多个样本来自于同一个受试者的情形，则应同时提供去标识化的受试者编号和样本编号。

### （三）来自于干预性研究的数据

对于干预性研究的数据，数据集中除以上基本信息外，还包括病例的临床试验分组、具体的诊疗信息、病例的临床评价终点等信息。

## 五、递交形式

体外诊断试剂临床试验数据相关的申报资料通常包括原始数据库、分析数据库、说明性文件和程序代码（如有），以下对各申报资料具体格式和内容提出要求。鼓励注册申请人参照临床数据交换标准协会（**Clinical Data Interchange Standards Consortium, CDISC**）标准递交数据。

建议注册申请人综合考虑临床试验电子数据采集系统/数据管理系统（**Electronic Data Capture System, EDC**）在数据采集及管理方面的优势，逐步推进 EDC 系统的使用，尤其对于数据量较大、产品风险较高的临床试验。

外文资料提供中文翻译件时，需注意对于原始和分析数据集，至少应翻译数据集、变量标签、观测值中的描述性文本。

### （一）原始数据库

原始数据库通常包含从病例报告表和外部文件中直接收集的原始数据，应包括临床试验按照方案的要求入组的所有病例及样本信息，原始数据库中的缺失数据不应进行填补。按照临床试验方案的剔除标准进行剔除的病例亦应包括在内，同时应备注剔

除原因。

对于数据量较大的临床试验，原始数据库通常由多个不同的原始数据集组成。单个原始数据集是相同主题下多个变量的集合，这些变量的观测值共同形成该原始数据集。例如，人口学资料数据集可包括去标识化的受试者编号、年龄、性别、临床诊断等。不同临床试验涉及的原始数据库不完全相同。单个原始数据集应收集相同主题下的变量。各数据集需包括受试者唯一标识变量，以实现同一受试者不同数据集数值之间的关联。若临床试验采用了随机分组，原始数据库中应包含随机号等变量。

数据集和变量命名应遵循“可读性”的原则，建议在对其命名时参考数据集或变量的英文或拼音，通过名称可较容易联想到该命名的实际含义。数据集通常以两个字母组成的代码命名，如受试者人口学数据集（dm）、样本检测数据集（lb）、样本采集与处理数据集（cl）、不良事件（ae）等。将变量“性别”命名为“Sex”、“年龄”命名为“Age”、“受试者编号”命名为“SUBJID”、“样本编号”命名为“SAMID”、“临床诊断”命名为“DIAG”等。

## （二）分析数据库

分析数据库是为便于统计分析使用原始数据集形成的数据库，用于产生临床试验报告中的统计结果。分析数据库主要包括原始数据库中的变量数据和按照临床试验方案或统计分析计划（如有）中事先确定的方法使用原始数据库变量数据衍生的数据。

分析数据库通常由多个不同的数据集组成，其中的数据集形成应与临床试验报告中的统计结果相对应。如需进行亚组分析时，可针对亚组分析构建不同的分析数据集。对于定量检测类试

剂，需要对检测数据在线性范围内的样本纳入定量相关分析，则应针对线性范围内的所有样本生成分析数据集。为了便于统计分析复验，分析数据集中的变量应具有可溯性，变量结构应清晰，不需繁琐的数据前处理即可开展统计分析。

分析数据集可基于其产生的相应统计结果进行命名，建议在分析数据集名称中加入“AD”（analysis data）前缀，以标识该数据集为分析数据集。

### （三）说明性文件

#### 1. 数据说明文件

数据说明文件用于描述原始数据库和分析数据库的内容和结构，是助于审阅者快速了解数据库中各数据集、变量及其相互间的结构关系，准确理解递交的数据内容。建议采用 Excel 文件，以表格的形式列明原始数据库和分析数据库中所含的数据集、变量、变量类型（如字符型、数值型）、标签、赋值及其对应关系。为了便于审阅，数据集和变量应具有相应的中文标签，标签长度不宜过长。若使用了外部词典（如 MedDRA），应明确使用的外部词典名称和版本号。

分析数据库的说明文件需列明衍生变量的生成规则，明确涉及到的变量和计算方法。如注册申请人使用程序代码生成分析数据库，建议以表格的形式列明生成各分析数据集所用到的程序代码文件和原始数据集名称。

#### 2. 统计分析说明文件

注册申请人应将从分析数据库到最终生成临床试验报告中的统计结果的过程或计算方法详细列明。

如统计分析的过程直接在 Excel 中实现的，应在说明性文件

中将所用的函数公式列明。如使用统计工具进行统计分析但未使用程序代码的，可以文字加图示的形式详细阐明统计分析步骤，应注明所用统计工具的版本号。

统计分析中使用程序代码的，应在统计分析说明文件中解释说明程序代码文件的使用方法、系统及软件环境，包括使用代码文件时是否需修改以及如何修改程序代码。同时，以表格形式逐一列明生成各统计结果图表所使用到的程序代码文件和数据集文件名称。注册申请人应说明原始数据集和分析数据集所用编码（如 UTF-8、EUC-CN 等），以避免所递交的数据集出现乱码的情形。同时应参照以下第（四）条要求提供相应的程序代码。

### 3. 注释病例报告表（如有）

相对于空白 CRF，注释 CRF 增加了注释内容，反映了数据库中变量与 CRF 表信息收集的对应关系。例如，在性别空白处注释变量名称 Sex。利用注释 CRF，审阅者可直观地查阅各变量在 CRF 表中的位置，CRF 中可能收集了一些与临床试验结果分析无关的冗余数据，这些数据可不包含在递交的数据库中，但应在注释 CRF 上明确标注为“不递交”，并阐明理由。

### 4. 其他说明性文件

除以上说明性文件外，鼓励注册申请人提交其他有利于审评人员快速了解临床试验数据库内容和结构的说明性文件（如概览性文件、其他特殊情况说明文件等）。

### （四）程序代码（如有）

如数据库的管理或统计分析中使用程序代码的，应提供程序代码，需递交的代码主要包括：用于原始数据库生成分析数据库的代码、分析数据库生成统计结果的代码等，用于调整格式或生

成表格的相关代码不需递交。递交的代码应符合通常的编程格式和编程规范，结构清晰明了，易于阅读。代码中以中文注释的形式描述不同程序模块的目的以及其他需解释的内容。如提交的代码中引用了宏程序，则需要提供相应的宏程序代码，并说明可以运行该程序的软件版本和系统环境。

### (五) 形式要求

原始数据库、分析数据库、说明性文件和程序代码（如有）分别放置于四个文件夹中。

临床试验数据库可以 Excel 形式进行递交，如使用 EDC 系统进行数据采集及管理的，原始数据库和分析数据库建议采用 XPT<sup>1</sup> 数据传输格式进行递交，建议全部原始数据集形成一个 XPT 文件，全部分析数据集形成一个 XPT 文件。建议采用 XPT 第 5 版本（简称 XPT V5）或以上版本作为数据递交格式。

数据说明文件及统计分析说明文件采用 PDF、Word、Excel 文件，其中变量词典建议采用 Excel 文件，注释病例报告表建议采用 PDF 文件。

程序代码建议采用 TXT 文件。

## 六、起草单位

国家药品监督管理局医疗器械技术审评中心

<sup>1</sup>XPT (XPOR)



格式文件的功能。