

制药行业微生物数据电子化存储规范

1 范围

本规范定义了制药行业微生物数据电子化存储的总体架构,规定了电子数据主要以数据库的形式存储,及在内容组织、技术实现方面需要完成的工作和需要满足的要求,并提出了对数据库在运行维护和服务方面的要求。本规范适用于制药行业微生物数据电子化存储数据库的建设。

2 规范性引用文件

下列文件中的内容通过文中的规范性引用而构成文件必不可少的条款,其中,注日期的引用文件,仅该日期对应的版本适用于本文件,不注日期的引用文件,其最新版本(包括所有的修改单)适用于本文件。

GB/T 20273 数据库管理系统安全技术要求

GB/T 29859 生物信息学术语

GB/T 34798 核酸数据库序列格式规范

GB/Z 42344 制药机械(设备)计算机化系统验证指南

T/ZJPA 001 制药工业环境微生物数据库构建技术规范

3 术语与定义

下列术语和定义适用于本文件。

3.1 数据 (data)

在药品研制、生产、经营、使用活动中产生的反映活动执行情况的信息，包括：文字、数值、符号、影像、音频、图片、图谱、条码等；记录是指在上述活动中通过一个或多个数据记载形成的，反映相关活动执行过程与结果的凭证。

3.2 数据库 (database)

数据库技术是信息资源管理最有效的手段，数据库设计是指对一个给定的应用环境，构建最优的数据库模式，建立数据库以及应用系统，有效存储数据，满足用户对信息存储和处理的要求。保证全过程信息真实、准确、完整和可追溯。

3.3 微生物数据 (microbiological data)

记录被测微生物相关信息的数据，包括图形数据（如照片等）、文字数据（如名称、形状描述等）、实验记录及相关原始数据、生物信息学分析数据等。

3.4 微生物数据库 (microbiological database)

以电子化方式存储微生物数据的数据库及其数据库管理系统。制药行业微生物数据库是面向制药行业的微生物电子数据存储领域，由若干逻辑相关的数据资源按照统一的标准规范整合形成，具有系统性和完整性，并通过统一的系统提供一站式服务的数据库。

3.5 生物信息学分析数据 (bioinformatic analysis data)

利用生物信息学分析软件对制药行业微生物测序数据进行归纳、注释等分析后得到的数据。

3.6 完全备份 (full backup)

备份所有的数据。

3.7 增量备份 (incremental backup)

备份上一次备份 (无论是哪种备份) 后, 所有发生变化的数据。

4 数据准备

4.1 数据来源

在微生物测试过程中产生和记录的被测微生物的相关数据和信息, 主要包括以下几个来源:

- a) 测试产生的相关原始数据和对原始数据进行进一步分析得到的相关数据;
- b) 通过文献查询得到的相关物种注释信息。

4.2 数据处理

在对数据库建设时, 良好的数据管理规范会影响制造商生成和记录的所有数据的质量。这些做法应确保数据是可归属、清晰的、同步的、原始的、准确的、完整的、一致的、持久的和可用的。数据完整、一致、准确、可信和可靠的程度, 以及数据的这些特征在整个数据生命周期中得到维护的程度, 是有效的药品质量体系的基本要求, 可确保药品具有所需的质量。数据生命周期是指数据如何产生、处理、报告、检查、用于决策、存储和最终在保留期结束时丢弃。

应按照实验先后顺序和时间逻辑关系建立数据库再按照以下几种方式进行分类:

a) 按照数据对象，可分为样本信息、表型数据、基因型数据等；

b) 按照数据类别，可分为文字数据、图片数据、序列数据等；

c) 按照数据物种分类，可分为真菌/细菌/古菌，也可进一步细分为：门、纲、目、科、属、种、亚种、株。

4.2.1 样本信息

样本信息收录以下信息：采样点、采样人员、采样时间、采样方式、采样环境、批次信息、编号信息、采样材料信息、培养信息、设备信息等。

4.2.2 无菌检查数据

无菌检查数据可收录以下信息：无菌检查的环境参数、无菌检查隔离器设备信息和运行参数和报警信息、无菌检查前物料和设备清洁消毒信息、无菌检查方法信息、无菌检查装载材料信息、无菌检查实验操作信息、无菌检查环境监测材料和环境监测培养及结果判读信息、无菌检查卸载后设备清洁信息、培养基信息、培养设备/温度/时间信息条件、培养结果判定信息等。

4.2.3 微生物计数数据

微生物限度数据可收录以下信息：实验前实验设备清洁消毒信息、环境监测材料和环境监测结果、微生物计数方法信息、供试液制备、菌数测定操作信息、使用后设备清洁消毒信息、培养基信息、培养设备/温度/时间信息、菌落计数结果等。

4.2.4 控制菌数据

控制菌数据可收录以下信息：实验前实验设备清洁消毒信息、环境监测材料和环境监测结果、控制菌检查方法信息、供试液制备和增菌培养、选择和分离培养、鉴定等操作和结果信息、使用后设备清洁消毒信息、培养设备/时间信息、增菌所需设备和材料信息、培养设备/温度/时间信息、控制菌检查结果等。

4.2.5 内毒素数据

内毒素数据可收录以下信息：细菌内毒素测试方法信息、标准品配制信息、样品及标准品稀释信息、鲎试剂加样信息、设备信息、培养条件、凝胶法培养结果、浊度法检查数据、显色法检查数据及结果判读等。

4.2.6 微生物鉴定数据

根据 SOP 规定，表型数据可收录以下信息：培养基类型及培养基信息、培养条件、培养设备信息、菌落形态、菌落颜色、革兰氏染色特性、耐受性（耐酸、耐盐、耐高温等）、培养条件、特征生化反应等；基因型数据可收录以下信息：原始数据、设备信息、引物和试剂信息、序列信息、序列比对结果，其中记录 DNA 序列比对结果的同时应记录所使用的鉴定数据库名称及版本号。

4.3 数据检查

在数据采集与录入过程中，由于人为因素和技术限制，数据错误难以完全避免。为确保入库数据符合建库技术规范，应实施严格的数据检查流程。该流程主要涵盖以下方面，由数据库应用系统自动执行，并辅以人工校验。

4.3.1 自动化检查

4.3.1.1 字段类型匹配

系统应自动检测输入数据与相应字段定义的数据类型(如数值、日期、文本等)是否相符,防止类型不匹配导致的数据解析错误或运算异常。

4.3.1.2 字段长度限制

对每条记录的各个字段进行长度检查,确保输入数据长度未超出预设的最大长度限制,避免因超长数据导致存储溢出或显示问题。

4.3.1.3 值域验证

对具有特定取值范围或枚举值的字段,系统应核实输入数据是否落入规定的值域内,剔除无效或超出范围的数据。

4.3.1.4 文字乱码检测

对于包含文本信息的字段,应用系统应具备识别和纠正字符编码问题的能力,防止乱码现象影响数据的正确解读与使用。

4.3.2 人工校验内容

4.3.2.1 逻辑一致性审查

核查数据间的逻辑关系是否合理,如时间序列的连续性、关联数据的一致性等,确保数据内在逻辑的严密性。

4.3.2.2 异常值筛查

对自动检查未覆盖的潜在问题进行人工识别,如数据突变、极端值、模式异常等,排除数据采集过程中的误录或仪器故障导致的异常数据。

4.3.2.3 质量评估与修正

根据业务知识和经验,对数据的质量进行综合评估,对存在疑义或明显错误的数据进行修正,或退回至数据源进行核实。

经过上述自动化检查与人工校对双重保障,确保数据准确无误后,方可正式录入数据库,以维持数据库内容的高质量与可靠性。

5 数据库搭建

5.1 数据库应用系统要求

数据库应用平台应遵循《数据库管理系统安全技术要求》(GB/T 20273),具有切实可行的安全保护和保密措施,确保数据的安全。制药行业微生物数据库应用系统至少包括:数据分析和统计、数据展示、用户管理和数据管理四大类功能,且数据库的应用系统应满足以下要求:

- a) 数据库应用系统至少支持数据访问接口,接口应符合标准接口规范;
- b) 数据库应用系统应具备系统操作、数据维护和用户访问日志记录,并定期进行审计的功能;
- c) 数据库应用系统应开发精细化的用户权限模块,实现基于角色的访问控制,确保不同用户仅能访问其职责范围内所需的数据;
- d) 数据库应用系统应提供图形化界面,允许管理员针对用户进行细粒度的数据访问授权控制;
- e) 数据库架构应具备良好的可移植性和可扩展性;

f) 数据库逻辑结构和物理结构设计应对数据进行规范化并进行数据流分析,调整数据库服务器的系统参数,选择合适的数据类型,使数据库的存储结构和配置达到最佳;

g) 数据库中各种对象的命名和程序代码的编写应规范,数据库应用系统能够适应不同的数据库平台;

h) 合理设计表间关联,尽可能地降低数据的冗余,保证数据的一致性和完整性。

5.2 数据库 Web 服务器系统技术要求

数据库 Web 服务器系统需满足以下技术要求:

a) 数据库服务系统客户端应支持 Chrome、Firefox 等主流浏览器;

b) 数据库服务客户端应能适应网络环境变化,既能提供实时在线数据访问服务,也能在离线状态下实现本地数据缓存和有限功能操作;

c) 数据库 Web 服务系统支持数据查询、元数据查询和下载服务;

d) 数据库服务系统应支持开放元数据获取协议,遵循 XML、SOAP、WSDL、和 UDDI 等 Web Service 协议;

e) 数据库服务系统应有一定的并发数量,支持多人在线使用。

5.3 数据库计算机化验证与确认

数据库应按照第 3 类软件进行计算机化验证与确认,具体的验证方案及实施可参考《制药机械(设备)计算机化系统验证指南》

(GB/Z 42344)，同时还应关注相关法规动态和行业最佳实践，确保验证活动与最新要求保持一致。

6 数据库功能

6.1 信息展示

微生物信息展示页面应包含：

a) 样本信息，可包括成分信息、批号、样本照片、需检查项目等；

b) 样本检查项目信息，可包括检查方法、限值、培养条件、培养结果等；

c) 菌株的生物学分类信息，按照界-门-纲-目-科-属-种-株形式；

d) 菌落/细胞照片；

e) 菌株的拉丁名和中文名，以及曾用名；

f) 该菌种的常见模式菌株名称；

g) 生理生化特征，可包括革兰氏染色、菌落形态、细胞形态、是否产酸产气、过氧化氢酶活性、凝固酶、是否溶血等；

h) 序列信息，可包括细菌 16S 序列/真菌 ITS 等特征序列、基因组序列（文件链接）、蛋白质图谱（文件链接）等；

i) 相关文献记录。

6.2 搜索功能

可根据样本名称、检查项目名称（如无菌检查、控制菌检查、微生物限度、内毒素检查等）、物种名称（拉丁名/中文名）、核酸序列进行搜索。

6.3 数据比对

用户可直接输入序列或者上传序列文件或满足格式的质谱图谱文件，通过 BLAST 或质谱特征峰等比对软件进行比对，并给出满足相似性阈值的结果。

6.4 分析功能

- a) 用户可对选中的样本进行系统发育分析；
- b) 浏览基因组草图（如有基因组信息）；
- c) 功能基因查找；
- d) 同类样本项目检验方法查找；
- e) 根据录在系统中的各有效成分样本的设定限值，与项目检查结果比较，进行结果判定；
- f) 样本结果与系统中同类样本的相同项目比较分析；
- g) 趋势分析。

6.5 上传/下载

用户可将日常监测的样品及信息上传至数据库，提交时需提供样品的照片、批号、序列信息、物种学名、生理生化参数、有效成分、常用原辅料及添加剂、检验方法、限值等信息，用户提交的信息经过管理员审核无误后方可录入数据库。用户可以下载菌株的序

以上内容仅为本文档的试下载部分，为可阅读页数的一半内容。如要下载或阅读全文，请访问：<https://d.book118.com/797020100005006121>