ICS 07.080
CCS C 50

团 体 标 准

T/CEATEC XXX—2025

肠道菌群供体库数据溯源管理规范

Guidelines for data traceability management of gut microbiota donor databases

（征求意见稿）

2025-X-XX 发布 2025-X-XX 实施

中国欧洲经济技术合作协会 发 布

目　　次

前言

本文件按照GB/T 1.1—2020《标准化工作导则　第1部分：标准化文件的结构和起草规则》的规定起草。

请注意本文件的某些内容可能涉及专利。本文件的发布机构不承担识别专利的责任。

本文件由中国欧洲经济技术合作协会提出并归口。

本文件主要起草单位：。

本文件主要起草人：。

本文件为首次编制。

肠道菌群供体库数据溯源管理规范

1 范围

本文件规定了肠道菌群供体库数据溯源的数据溯源体系架构、数据溯源流程、数据溯源技术要求、数据溯源管理与安全保障。

本文件适用于开展肠道菌群采集、制备、检测、存储和分发的各类供体库机构，用于指导其建立和实施数据溯源管理系统。

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本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 22239 信息安全技术 网络安全等级保护基本要求

GB/T 32905 信息安全技术 SM3密码杂凑算法

WS/T 428 成人体重判定

3 术语和定义

下列术语和定义适用于本文件。

3.1

肠道菌群供体库 gut microbiota donor databases

通过规范化收集、处理、存储肠道菌群供体的生物样本及相关数据，为科研、临床等活动提供资源支持的机构或平台。

3.2

数据溯源 data traceability

追踪肠道菌群从供体到最终制品的所有相关数据和操作过程的能力。

4 数据溯源体系架构

4.1 总体原则

数据溯源管理应遵循以下原则：

a) 全程覆盖：覆盖从供体招募到菌群制品最终使用的所有环节；

b) 唯一标识：为供体、样本、批次等关键实体分配全局唯一的标识符；

c) 真实准确：确保记录的数据与实际情况一致，数据修改必须留有审计线索；

d) 及时完整：在操作发生时或发生后立即记录，确保记录项目的完整性；

e) 安全保密：保护供体隐私和敏感数据，防止未授权访问、篡改或泄露。

4.2 体系构成

数据溯源体系应包括：

a) 数据源：各环节产生的原始数据，如问卷、检验报告、仪器输出数据、人工记录等；

b）数据采集与记录系统：用于采集、录入和存储数据的工具，如电子数据采集系统、实验室信息管理系统、纸质记录表等；

c）数据存储与管理平台：集中存储和管理所有溯源数据的数据库或信息平台；

d）数据查询与追溯接口：供授权用户查询和追溯数据的界面或应用程序接口。

4.3 人员职责

应明确以下人员的数据管理职责：

a）数据管理员：负责数据系统的日常运维、数据备份、用户权限管理；

b）质量保证人员：负责对数据记录进行审计，确保符合规范要求；

c）实验室技术人员：负责准确、及时地记录实验操作和数据；

d）临床协调员：负责供体信息的采集、录入与核对；

e）供体库负责人：对数据溯源体系的完整性和有效性负总责。

5 数据溯源流程

5.1 供体筛查与入组阶段

5.1.1 供体基本信息

应至少包括唯一供体ID、姓名（可脱敏）、性别、出生日期、联系方式、身份证号（加密存储）等。

5.1.2 健康问卷

应至少包括个人史、家族史、旅行史、用药史、饮食习惯、生活方式等，问卷应采用结构化电子表单，选项明确。

5.1.3 临床检查数据

应至少包括身高、体重（用于计算BMI，应介于18.5kg/m^2～24.9kg/m^2，符合WS/T 428的规定）、生命体征等。

5.1.4 实验室检测数据

数据应至少包括：

a）基本项目：血常规、血生化（肝功能、肾功能、血糖、血脂等）；

b）传染病筛查：HIV抗体（阴性）、梅毒螺旋体抗体（阴性）、乙肝表面抗原（阴性）、乙肝核心抗体（如有要求）、丙肝抗体（阴性）、甲肝抗体IgM（阴性）、戊肝抗体IgM（阴性）；

c）粪便常规与培养：寄生虫卵、艰难梭菌毒素、沙门氏菌、志贺氏菌、弯曲杆菌等（均为阴性）。

5.1.5 心理评估结果

由合格心理医师出具的评估报告。

5.2 样本采集与处理阶段

数据应至少包括：

a）样本信息：唯一样本ID、关联的供体ID、采集日期与精确时间、采集方法（如专用容器）、样本性状（布里斯托大便分型，1-7型）、样本量（精确到0.1g）；

b）采集环境数据：采集地点、环境温度（应在15℃～25℃之间）；

c）处理过程数据：接收时间、处理开始时间、处理人员ID、处理方法（如稀释、过滤、均质化条件）、使用的设备ID及校准状态、添加的溶液或保护剂的批号与体积、处理过程中的环境温度与时间控制；

d）中间产物信息：处理过程中产生的任何中间样本的唯一ID及其与原始样本的关联关系。

5.3 样本检测与表征阶段

5.3.1 微生物学检测

数据应至少包括：

a）需氧菌与厌氧菌培养计数（CFU/g）；

b）16S rRNA基因测序数据（原始数据存储路径、分析后的菌群Alpha多样性与Beta多样性指数、核心菌门/属的相对丰度）；

c）宏基因组测序数据（如有，存储路径及关键功能基因信息）。

5.3.2 安全性检测

应复核某些病原体（如CMV, EBV，根据需求）的PCR检测结果（阴性）。

5.3.3 活性与功能检测（如有）

数据应至少包括：

a）特定短链脂肪酸（SCFAs）浓度（如乙酸、丙酸、丁酸，单位μmol/g）；

b）样本pH值。

5.4 样本存储与入库阶段

数据应至少包括：

a）存储信息：最终制剂形式（如冷冻菌液、冻干粉）、分装规格（每管/瓶体积或重量）、分装日期、冷冻保护剂成分与批号、存储位置（超低温冰箱/液氮罐的唯一ID、架号、盒号、孔位号）；

b）存储环境监控数据：超低温冰箱温度（应稳定在-80℃±10℃），连续监控，数据自动记录或至少每日人工记录两次，液氮罐液位温度（应低于-150℃）及液位，定期记录；

c）入库确认：由库管员和QA人员双重确认后，在系统中更新样本状态为“已入库”。

5.5 样本出库与应用阶段

数据应至少包括：

a）出库申请：申请单号、申请单位、申请人、用途（临床研究/治疗、编号）、目标患者信息（如患者研究ID）、所需样本量、出库日期；

b）样本挑选与核对：被取出的样本ID、数量、取出时样本外观、取出人员ID、核对人员ID；

c）运输信息：运输载体ID、运输条件（如干冰维持温度低于-65℃）、启运时间、送达时间、承运方、运输过程温度监控记录（到达后接收方确认）；

d）接收确认：接收单位、接收人、接收日期时间、接收时样本状态与温度确认。

5.6 数据归档与长期保存阶段

所有与供体、样本相关的原始记录、电子数据、报告等，应在项目结束或样本销毁后，按照既定策略（如15年）进行归档。归档数据应置于安全的、只读的存储介质中，并建立索引以备查询。

6 数据溯源技术要求

6.1 数据标识与编码

应采用统一的编码规则生成唯一标识符：

a）供体ID：DBK（供体库代码）-D（供体类型）-XXXXXX（6位数字序列），如DBK-H-000123；

b）样本ID：关联供体ID-YYYYMMDD-SSS（当天样本序列号），如DBK-H-000123-20251112-001；

c）批次ID：用于同一处理过程产出的多个样本，如LOT-20251112-P001。

6.2 数据采集与记录

应满足以下要求：

a）宜优先采用电子数据采集系统，系统应具备数据校验功能（如范围检查、格式检查）；

b）手工记录应使用耐久的墨水笔填写，字迹清晰。修改时应划线更正，注明修改人、日期及原因，原信息仍清晰可辨；

c）所有记录均应有记录人签名和记录日期。

6.3 数据存储与安全

应满足以下要求：

a）电子数据应定期备份（建议每日增量备份，每周全量备份），备份数据异地存放；

b）存储系统应具备高可用性和容灾能力，应符合GB/T 22239中第三级及以上安全要求；

c）数据访问应实行严格的权限控制，遵循最小权限原则；

d）敏感数据（如供体身份信息）应进行加密存储；

e）数据存储期限应符合相关法规要求，原则上不少于菌群制品使用后15年。

6.4 数据传输与接口

系统间数据传输应采用安全协议（如HTTPS，SFTP），与外部系统（如医院HIS/LIS）的接口应定义清晰的数据格式和交换协议。

6.5 数据查询与追溯

系统应能支持以下查询：

a）通过供体ID，查询其所有样本的全流程信息；

b）通过样本ID，反向追溯至供体所有筛查信息及该样本的所有处理步骤；

c）通过批次ID，查询该批次下所有样本的信息及处理记录；

d）应能生成完整的追溯报告。

7 数据溯源管理与安全保障

7.1 组织与人员

7.1.1 应设立数据管理与数据安全岗位，明确其职责。数据管理岗位负责数据质量标准执行与监督；数据安全岗位负责安全策略制定与实施。

7.1.2 所有相关人员必须接受数据管理规范、标准操作规程、保密及安全培训，并考核合格后方可上岗。

7.2 文件与记录控制

应建立文件控制程序，对所有标准操作规程（SOP）和记录表格进行版本控制。所有数据记录均作为受控文件管理，确保其填写、修改、归档和销毁过程可追溯。

7.3 数据全过程质量控制

7.3.1 事前控制

制定明确的数据采集规范；对人员和设备进行双重管控，人员应培训授权，设备应定期校准维护。

7.3.2 事中控制

利用信息系统对数据录入进行实时校验（格式、逻辑、值域）；建立关键数据双人审核机制，确保数据准确无误后方可入库。

7.3.3 事后控制

定期开展数据质量检查（每月全面检查，每周抽样检查），对发现的问题进行分析与整改，并持续优化数据管理流程。

7.4 信息系统与安全控制

7.4.1 系统管理

用于数据管理的计算机化系统应经过验证，确保其符合预定用途，并制定系统维护和应急预案。

7.4.2 访问控制

应满足以下要求：

a）采用基于角色的访问控制（RBAC）模型，遵循“最小权限原则”分配权限；

b）实施严格的身份认证机制，鼓励采用双因素认证，并制定强密码策略；

c）权限的分配与变更必须经过审批，并定期进行权限审计与清理。

7.4.3 数据加密

应满足以下要求：

a）传输加密：数据在网络传输过程中应采用加密传输协议（如SSL/TLS、IPsec等），加密算法符合GB/T 32905的要求；

b）存储加密：对敏感数据（如供体身份信息）进行加密存储，并对加密密钥进行安全管理，定期更换。

7.4.4 安全审计

建立安全审计日志系统，记录所有关键操作（如数据访问、修改），日志应包括操作人、时间、内容和结果等信息，并长期保存。应定期审查审计日志以发现异常。

7.4.5 应急处置

应制定数据安全事件应急预案，明确流程与职责，并定期组织演练。发生安全事件时，应立即启动预案进行处置并上报。用于数据管理的计算机化系统应进行验证，确保其符合预定用途，并制定系统应急预案。

7.5 审计与改进

质量保证部门应定期（如每年）对数据溯源体系进行内部审计，并对审计发现、偏差、投诉等进行分析，采取纠正和预防措施，持续改进数据管理体系。