

基于德尔菲法的生物安全大数据标准体系框架构建^{*}

赵应辰 王松旺 李言飞 苏雪梅

(中国疾病预防控制中心传染病溯源预警与智能决策全国重点实验室 北京 102206)

〔摘要〕 目的/意义 构建生物安全大数据标准体系框架,为相关标准研制提供参考。方法/过程 采用文献综述法构建生物安全大数据标准体系框架下的标准类别条目池,采用三维结构法初步构建标准体系框架,采用德尔菲法优化完善标准体系框架。结果/结论 提出生物安全大数据标准体系框架,包括基础类标准、数据类标准、技术类标准等 7 个一级标准类别,26 个二级标准类别,强化了生物安全大数据标准顶层设计。

〔关键词〕 生物安全;大数据;德尔菲法;标准体系

〔中图分类号〕 R-058 〔文献标识码〕 A 〔DOI〕 10.3969/j.issn.1673-6036.2025.06.005

Construction of a Biosafety Big Data Standard System Framework Based on the Delphi Method

ZHAO Yingchen, WANG Songwang, LI Yanfei, SU Xuemei

National Key Laboratory of Intelligent Tracking and Forecasting for Infectious Diseases, Chinese Center for Disease Control and Prevention, Beijing 102206, China

〔Abstract〕 **Purpose/Significance** To establish a framework for the biosafety big data standard system, and to provide references for the development of related standards. **Method/Process** The standard category item pool of the biosafety big data standard system is constructed by the literature review method, the framework of the standard system is initially constructed by the three-dimensional structure method, and the Delphi method is used for optimization and refinement. **Result/Conclusion** The framework of the biosafety big data standard system is proposed, including 7 first-level standard categories such as basic standards, data standards, and technical standards, as well as 26 second-level standard categories, strengthening the top-level design of biosafety big data standards.

〔Keywords〕 Biosafety; Big Data; Delphi; Standard System

〔修回日期〕 2025-05-16

〔作者简介〕 赵应辰,硕士研究生,发表论文 1 篇;通信作者:苏雪梅。

〔基金项目〕 国家重点研发计划(项目编号:2022YFC2602301)。

1 引言

近年来突发公共卫生事件频发,生物安全风险日益复杂,对生物安全数据的有效管理和标准化建设成为国家生物安全治理的关键^[1]。生物安全数据涵盖疫情疫病发病和死亡数据、人员贸易往来数据、医疗卫生数据、外来入侵生物数据等^[2]。快速从这些数据中获取病原体基因数据、流行病学数据

和跨境传播路径数据对于疫情防控至关重要。

然而，当前我国生物安全数据分散于多个业务系统，如传染病监测系统、实验室管理系统等，存在数据格式不统一、流转受限、多部门间数据难以互联互通等问题^[3]，无法形成高效的早期预警和决策支持体系^[4]。此外，在生物武器防范、生物恐怖主义应对^[5]、高等级生物安全实验室（BSL-3/BSL-4）管理中^[6]，样本存储、实验操作、人员访问控制等环节均亟需标准化制度保障，以防止生物因子的泄漏和滥用。缺乏统一的数据采集、存储与分析技术标准，也会导致数据的完整性、真实性与可追溯性难以保障。

在国际上，大数据标准化已成为共识。国际标准化组织（International Organization for Standardization, ISO）和国际电工委员会（International Electrotechnical Commission, IEC）提出《信息技术 大数据参考架构》（ISO/IEC TR 20547—1: 2020）^[7]，美国国家标准与技术研究院制定大数据互操作性框架，欧盟的《通用数据保护条例》也为数据治理标准化提供了制度参考^[8]。我国在食品安全^[9]、健康医疗^[10]等领域已开展大数据标准体系建设实践，为标准研制和落地应用提供了基础。相比之下，我国现行生物安全标准系统性较弱，存在标准覆盖不全、交叉重叠、老化等问题^[11]，生物安全领域尚未建立大数据标准体系，缺乏顶层设计和统筹协调，导致生物安全数据的采集、存储、共享和应用缺乏统一标准，严重制约了生物安全大数据的标准化进程。为此，本文基于大数据标准化理论，结合我国生物安全管理体系的现实需求，初步构建生物安全大数据标准体系框架，以为促进标准有效衔接、标准研制，以及探索生物安全大数据标准化建设路径提供理论支撑。

2 资料与方法

2.1 文献综述法

根据《中华人民共和国生物安全法》《大数据标准化白皮书（2020 版）》《标准体系构建原则和要求》（GB/T 13016—2018）和《信息技术 大数据技术参考模型》（GB/T 35589—2017）等相关指导性文件、法律法规和标准，以“数据标准”“信息

标准”“生物安全”“实验室管理”“外来物种入侵”“微生物耐药”“遗传资源”“传染病监测”为关键词，在万方数据、中国知网中查阅相关文献，经梳理分析后初步构建生物安全大数据标准体系框架标准类别条目池。

2.2 三维结构法

三维结构法是一种系统性分析方法，能够从多个维度梳理复杂信息，被广泛应用于不同领域标准体系构建^[12-14]。采用三维结构法对生物安全大数据标准进行分类和设计，初步构建生物安全大数据标准体系框架，其中 X 轴依据《大数据标准化白皮书（2020 版）》以标准类别划分，Y 轴依据《中华人民共和国生物安全法》适用的领域以专业领域划分，Z 轴依据《标准体系构建原则和要求》（GB/T 13016—2018）以标准层次划分，见图 1。

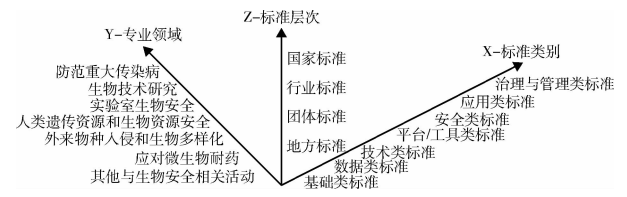


图 1 生物安全大数据标准体系三维模型

2.3 德尔菲法

德尔菲法是一种通过重复问卷调查获取专家意见，并整合专家组个人意见为群组共识意见的主观评价方法。通过德尔菲法获得标准体系中各标准类别的重要性、专家对标准类别的判断依据和熟悉程度。

2.3.1 德尔菲问卷 调查问卷主要包括 3 部分内容：“专家基本情况”“生物安全大数据标准体系框架”“生物安全大数据标准体系框架指标专家评价表”。“专家基本情况”中的“专家主要从事的专业领域”为多选题，允许专家根据自身研究和实践经验勾选多个相关领域。第 3 部分为调查问卷主体，包括标准类别评价、判断依据和熟悉程度。本研究进行了 2 轮德尔菲专家咨询，均按照重要性对标准类别进行评判。

2.3.2 专家咨询 根据研究目的，确定咨询专家

纳入标准：从事实验室管理、实验室检测、信息技术、大数据、传染病和突发公共卫生事件监测预警、标准管理等专业工作，熟悉生物安全及数据标准相关知识；具有高级及以上职称；从事相关工作年限 > 10 年；对本研究有较高积极性并自愿参与。

2.4 统计分析

使用 Microsoft Excel 2017 软件分析标准类别的重要性、专家背景、积极程度、权威程度等相关指标；使用 SPSS 23.0 软件计算各级标准类别的变异系数（coefficient of variation, CV）和肯德尔和谐系数（Kendall's W），采用 χ^2 检验对 Kendall's W 进行显著性检验，检验水准 $\alpha = 0.05$ 。

2.4.1 标准类别重要性 标准类别重要性评价采用 10 点量表，从“非常不重要”（1 分）到“非常重要”（10 分），评分数值越大，表明专家认为该标准类别越重要。各标准类别的重要性用算术平均值（ $\bar{X}$ ）和满分频率（ K ）来表示。其中， $\bar{X}$ 指所有专家对某个标准类别重要性评分的算术均数， K 指对某标准类别给出满分的专家数与参与某指标评分的专家总数的比值。 $\bar{X}$ 和 K 越大，说明该标准类别在标准体系中的重要性越高^[15]。

2.4.2 专家积极程度 专家积极程度通过“专家积极系数（问卷回收率）”表示，反映专家的参与意愿和合作程度。积极系数越高，表示专家积极性越高。当积极系数 > 70% 时可认为专家积极性较高^[16]。

2.4.3 专家权威程度 专家权威程度以专家权威系数（Cr）表示，为判断依据（Ca）和熟悉程度（Cs）的算数平均值，当 $Cr > 0.7$ 时，可认为结果可靠^[17]。其中，判断系数根据专家评分的判断依据计算，赋值方法，见表 1。

$$Cr = \frac{Ca + Cs}{2}$$

(1)

表 1 判断依据及其影响程度量化（分）			
判断依据	对专家判断的影响程度		
	大	中	小
理论依据	0.3	0.2	0.1
工作经验	0.5	0.4	0.3
从国内外同行处了解	0.1	0.1	0.1
直觉判断	0.1	0.1	0.1

熟悉系数根据专家对条目的熟悉程度计算，熟悉程度按照李克特 5 级评分法分为 5 个等级：“熟悉” 1.0、“较熟悉” 0.8、“一般” 0.6、“不太熟悉” 0.4 和“不熟悉” 0.2^[18]。式中“ a ”“ b ”“ c ”“ d ”“ e ”分别代表各等级熟悉程度专家人数。

$$Cs = \frac{1.0 \times a + 0.8 \times b + 0.6 \times c + 0.4 \times d + 0.2 \times e}{a + b + c + d + e}$$

(2)

2.4.4 专家意见协调程度 专家意见协调程度指所有专家对全部标准类别作出判断的协调性，由变异系数和肯德尔和谐系数体现。当 $CV < 0.25$ 时，可认为专家意见趋于统一；Kendall's W 介于 0 ~ 1 之间，Kendall's W 越高则协调程度越高^[19]。

2.4.5 界值法 采用界值法^[20]对算数平均值、满分频率和变异系数进行筛选。 $\bar{X}$ 界值 = $\bar{X}$ 平均值 - $\bar{X}$ 标准差， K 界值 = K 平均值 - K 标准差，CV 界值 = CV 平均值 + CV 标准差。其中， $\bar{X}$ 和 K 如果得分大于界值，则该条目入选；CV 如果得分小于界值，则该条目入选。淘汰均不符合 3 个标准要求的标准类别；对不符合 1 个或 2 个标准要求的标准类别，根据合理性、系统性等原则进行修改，或者经讨论确定。

3 研究结果

3.1 专家基本情况（表 2）

表 2 专家基本情况			
类别	特征	数量 (人)	构成比 (%)
文化程度	硕士研究生	8	40.00
	博士研究生	12	60.00
工作年限 (年)	11 ~ 15	1	5.00
	16 ~ 20	2	10.00
	> 20	17	85.00
专业领域	实验室管理	9	45.00
	实验室检测	5	25.00
	信息技术	8	40.00
	大数据	7	35.00
	传染病和突发公共卫生事件监测预警	8	40.00
	标准管理	7	35.00
	生物安全	1	5.00

共纳入 20 名专家，主要来自实验室管理、信息技术、传染病和突发公共卫生事件监测预警等领域，职称均为正高级，工作年限以 20 年以上为主。

3.2 专家积极程度

第 1、2 轮专家咨询分别发出问卷 20、17 份，分别收回 19、16 份，全部问卷符合填写要求，问卷回收率分别为 95%、94%，均大于 70%，表明专家对本项研究的认同程度较高。

3.3 专家权威程度

两轮咨询中专家对标准类别的熟悉系数分别为 0.94 和 0.93，对标准类别的判断系数分别为 0.94 和 0.93，两轮咨询专家权威系数分别为 0.94 和 0.93，均大于 0.7，表明专家权威程度较高，可信性较强。

3.4 专家意见协调程度

两轮咨询各维度及标准类别均 $CV < 0.25$ ，Kendall's $W > 0.20$ ；两轮咨询 Kendall's W 均有统计学意义 ($P < 0.01$)，专家意见一致性较高，见表 3—表 4。

表 3 第 1 轮咨询专家意见协调程度

评测内容	CV	Kendall's W	χ^2	P
一级标准类别	0.052 ~ 0.180	0.278	26.370	<0.001
二级标准类别	0.040 ~ 0.223	0.283	86.000	<0.001
全部标准类别	0.040 ~ 0.223	0.312	130.486	<0.001

表 4 第 2 轮咨询专家意见协调程度

评测内容	CV	Kendall's W	χ^2	P
一级标准类别	0.052 ~ 0.150	0.333	32.000	<0.001
二级标准类别	0.065 ~ 0.200	0.247	98.606	<0.001
全部标准类别	0.052 ~ 0.200	0.267	136.487	<0.001

3.5 标准类别筛选和修改

根据第 1 轮咨询结果，二级标准类别“专题应用”等 3 个指标均未满足界值，直接删除；对其余标准类别，综合考虑专家意见，经课题组充分讨论，修改如下。一是删除“数据访问”“数据可视化”“数据质量”“计算平台”“管理平台”“服务平台”“系统安全”7 个二级标准类别。二是增加“技术类标准”1 个一级标准类别，“数据类标准”中增加“分类与编码”“共享文档规范”“大数据集描述”，“平台/工具类标准”中增加“大数据系统产品”“数据库产品”，“安全类标准”中增加“应用安全”“服务安全”“平台和技术安全”，“应用类标准”中增加“数据交易”，“治理与管理类标准”中增加“数据治理”，共计 10 个二级标准类别。三是将“基础标准”“数据标准”“平台标准”“安全标准”“应用标准”“管理标准”6 个一级标准类别修改为“基础类标准”“数据类标准”“平台/工具类标准”“安全类标准”“应用类标准”“治理与管理类标准”。将“数据共享”1 个二级标准类别修改为“数据开放共享”。

第 2 轮咨询结果显示，“大数据系统产品”和“数据库产品”两个二级标准类别 3 个指标均未满足界值，但经过课题组深入讨论与评估，认为这两个标准类别在生物安全领域数据管理和技术架构中具有重要支撑作用，予以保留。数据库产品确保生物安全大数据的高效存储、完整性和安全性，支持对多样化数据的管理和实时更新，为进行准确的数据分析和决策提供基础。大数据系统产品提供技术平台，支撑数据从采集、存储到处理、分析的全生命周期管理，确保生物安全数据能被高效处理，能够提高系统的稳定性、可扩展性。其余一级、二级标准类别经讨论后也予以保留。最终构建的生物安全大数据标准体系框架包括一级标准类别 7 个、二级标准类别 26 个，见表 5。

表 5 生物安全大数据标准体系框架（第 2 轮咨询结果）

一级标准类别	二级标准类别	$\bar{X}$	K	CV
A 基础类标准	A1 总则	9.158	0.147	0.579
	A2 术语与定义	9.158	0.105	0.474
	A3 参考架构	8.842	0.142	0.421
B 数据类标准	B1 元数据	9.421	0.089	0.579
	B2 数据元	9.158	0.105	0.474
	B3 分类与编码	9.000	0.153	0.526
	B4 共享文档规范	8.316	0.205	0.368
	B5 大数据集描述	8.842	0.115	0.316
C 技术类标准	C1 数据采集	9.000	0.123	0.421
	C2 数据存储	9.105	0.109	0.421
	C3 数据处理	9.105	0.109	0.474
	C4 数据传输	8.684	0.144	0.263
	C5 数据交换	9.105	0.109	0.368
	C6 数据销毁	8.947	0.151	0.526
D 平台/工具类标准	D1 大数据系统产品	8.158	0.179	0.105
	D2 数据库产品	8.158	0.165	0.158
E 安全类标准	E1 平台和技术安全	9.158	0.138	0.526
	E2 应用安全	9.212	0.100	0.474
	E3 数据安全	9.632	0.052	0.632
	E4 服务安全	9.211	0.086	0.421
F 应用类标准	F1 数据交易	8.527	0.132	0.158
	F2 数据开放共享	9.421	0.064	0.474
G 治理与管理类标准	G1 数据治理	8.421	0.155	0.211
	G2 数据管理	9.105	0.103	0.421
	G3 运维管理	8.474	0.164	0.316
	G4 管理评估	8.369	0.133	0.158

4 讨论

4.1 标准体系框架的科学性和适用性

在标准体系构建领域，德尔菲法的科学性与可靠性已在诸多实例中得到验证^[21-22]。本研究所选专家具有较高的专业水平和丰富的工作经验，均为研究生学历，博士研究生超过一半，85%的专家具有 20 年以上的工作经验，且专业领域与生物安全大数据范畴吻合。两轮专家积极程度均较好，权威程度均较高，专家意见具有较好的一致性，分别为 0.312 和 0.267，*P* 值均小于 0.01，保证了标准体系层次适当、划分清楚、可靠性较好。

4.2 标准体系框架在生物安全数据治理中的支撑作用

大数据标准体系构建不仅是数据管理的技术需求，更是行业发展的核心支撑^[9-10,23-24]。生物安全大数据标准体系框架构建具有广泛的实践价值，为数据管理、共享和应用提供了规范保障，有助于促进跨部门、跨行业、跨地域的数据流通，提升生物安全治理能力。本研究构建的生物安全大数据标准体系涵盖多个关键领域，从基础定义到数据管理和技术应用，可提升生物安全数据的规范化、共享性和安全性。基础类标准通过提供统一的总则、术语定义和参考架构，为整个标准体系奠定基础，确保各领域对生物安全大数据的理解一致；数据类标准专注于数据本身的标准化的，包括元数据、数据元、分类编码等，为数据采集和共享明确规则，确保数据的一致性和可追溯性；技术类标准规范数据采集、存储、处理、传输等技术环节，保障生物安全大数据生命周期中每阶段的操作可靠性；平台/工具类标准着眼于确保相关技术平台和工具的互操作性与安全性，为生物安全数据流通提供支撑；安全类标准同样贯穿于整个数据生命周期，确保数据和服务的安全性、完整性和保密性，防止数据泄漏和滥用；应用类标准推动大数据交易与共享标准化，为跨部门合作提供规范，促进数据的高效利用；治理与管理类标准涉及生物安全大数据的长期管理，确保数据治理、运维管理和评估工作的有效性，提升标准的执行力与持续性。

综上所述，该生物安全大数据标准体系框架为相关标准的研制提供了技术参考和方向，使相关单位能够有序开展标准研制，避免不同层级标准交叉冲突，加速了标准的研发进程。同时，该体系框架有助于推动跨领域、跨区域的生物安全数据共享和协作，打破数据孤岛、实现数据共享、促进数据应用、释放数据价值，提升对突发公共卫生事件和跨境病原体防控的响应能力，也进一步强化了国家生物安全防线。

5 结语

在大数据时代，生物安全数据的高效管理和合理应用已成为应对公共卫生风险、环境威胁和生物

安全事故的关键。研究基于德尔非法构建生物安全大数据标准体系框架,将标准划分为基础类标准、数据类标准、技术类标准等 7 个一级标准类别,26 个二级标准类别,强化了生物安全大数据标准的顶层设计,为生物安全大数据标准研制提供参考。

本研究尚存在一定局限性。一是所邀请的 20 名专家可能存在代表性不均、专业分布不全面等问题,后续研究将扩大专家遴选范围,对标准体系进行完善。二是该体系框架仍以大数据视角为主,在生物安全特定业务场景中的适配性和针对性仍有待进一步验证和细化,拟在未来研究中结合生物安全不同子领域(如实验室生物安全、外来物种入侵监测、微生物耐药性治理等)开展分领域标准体系构建研究,增强框架的领域适用性。

作者贡献: 赵应辰负责调研实施、数据分析、论文撰写;王松旺负责问卷设计、论文撰写;李言飞负责数据整理、论文修订;苏雪梅负责研究设计、提供指导。

利益声明: 所有作者均声明不存在利益冲突。

参考文献

- 1 徐建国,舒跃龙,易在炯,等.我国国境生物安全防控体系现代化建设研究[J].中国工程科学,2023,25(5):21-29.
- 2 韩辉,伍波,侯颖,等.大数据在国门生物安全风险防控中的研究和应用[J].口岸卫生控制,2024,29(1):12-14.
- 3 张婷,李辉.生物安全情报工作研究[J].科技情报研究,2025,7(1):109-117.
- 4 石晶金,于广军.健康医疗大数据共享关键问题及对策[J].中国卫生资源,2021,24(3):223-227,237.
- 5 王爱新.疾控机构微生物实验室生物安全管理探讨[J].中国卫生标准管理,2024,15(9):70-73.
- 6 贾延芳,郑健,高英堂.针对新发传染病生物安全实验室的管理与展望[J].世界华人消化杂志,2020,28(21):1059-1067.
- 7 Information technology—big data reference architecture—part 1: framework and application process: ISO/IEC TR 20547—1: 2020 [EB/OL]. [2025-06-11]. <https://www.iso.org/standard/71275.html>.
- 8 CRUTZEN R, YGRAM PETERS G J, MONDSCHIEIN C. Why and how we should care about the general data protec-

- tion regulation [J]. Psychology & health, 2019, 34 (11): 1347-1357.
- 9 张飞,景亚萍,朱平,等.食品安全大数据标准体系建设研究[J].食品科学,2020,41(13):318-325.
- 10 陈敏,牟海燕,秦健.健康医疗大数据标准体系框架研究[J].中国数字医学,2018,13(4):14-16,33.
- 11 赵应辰,王松旺,李言飞,等.1993—2023 年我国生物安全标准现状与分析[J].预防医学情报杂志,2024,40(12):1599-1605.
- 12 许彦鑫,苏伟.基于三维架构的综合标准体系构建与应用[J].航天标准化,2017,(3):8-10,19.
- 13 翟运开,张瑞霞,路薇,等.基于三维结构的精准医疗示范应用标准体系构建[J].中国卫生资源,2020,23(1):33-37.
- 14 刘宇,张蕾,王伟,等.自动化集装箱码头标准体系构建与评价研究[J].交通运输研究,2023,9(2):91-99.
- 15 郝秀炜,陆翼,陈向心,等.基于德尔非法的辽宁省中医智慧医疗大数据标准化体系建设的指标构建[J].中国医药导报,2024,21(8):201-205.
- 16 朱媛媛,曹承建,张磊,等.基于德尔非法和层次分析法构建健康企业评价指标体系的研究[J].中华劳动卫生职业病杂志,2024,42(2):112-117.
- 17 叶雨昕,唐文熙,李帅龙,等.基于德尔菲-熵权法的智慧医保发展水平评估体系构建研究[J].中国医院管理,2024,44(2):1-5.
- 18 GUAN C Y, LIU Y L, HUANG H, et al. Outpatient experience of human caring scale (OEHC - Scale): improvement by Delphi method [J]. Current medical science, 2018, 38 (2): 360-371.
- 19 李洪兴,罗庆,张荣,等.肯德尔和谐系数 W 检验及程序实现[J].中国医院统计,2013,20(3):170-173.
- 20 曾强,魏文彬,刘静,等.基于德尔非法构建天津市职业病防治能力评估体系[J].中华劳动卫生职业病杂志,2023,41(10):871-875.
- 21 张永泽,丁选胜,杨涓,等.基于德尔非法构建我国临床药师知识体系标准[J].中国医院药学杂志 2020,40(7):795-799.
- 22 吴群华,章媛,俞峰,等.基于德尔非法构建基层医院急诊分诊标准体系[J].重庆医学,2020,49(5):853-858.
- 23 胡良霖,朱艳华,高瑜蔚,等.烟草科研大数据标准体系的构建[J].烟草科技,2020,53(4):100-106.
- 24 姚艳敏,白玉琪.农业大数据标准体系框架研究[J].农业大数据学报,2019,1(4):76-85.