人口健康科学数据标准体系框架构建研究*

王安然　吴思竹　李赞梅　胡拯涌
（中国医学科学院/北京协和医学院医学信息研究所，北京 100020）

摘要：标准规范是推动人口健康科学数据开放共享与分析利用的重要保障。全面调研国内人口健康科学数据相关标准规范的研制情况，对标准化工作的现状与局限性进行系统梳理。在现有通用领域科学数据标准体系框架和卫生健康信息标准体系框架的建设基础上，构建多维视角、层次清晰的人口健康科学数据生命周期管理标准体系框架，并进一步提出对人口健康科学数据标准化工作的建议。

关键词：标准体系框架；科学数据；数据生命周期；人口健康

中图分类号：R-05　DOI：10.3772/j.issn.1673-2286.2024.12.001

引文格式：王安然，吴思竹，李赞梅，等. 人口健康科学数据标准体系框架构建研究[J]. 数字图书馆论坛，2024，20（12）：1-8.

数据作为新型生产要素已成为促进国家科技创新和社会经济发展的关键资源，其开放共享程度是国家信息化、数字化发展水平的重要体现。2023年以来，国家陆续发布《中共中央 国务院关于构建数据基础制度更好发挥数据要素作用的意见》[1]、《数字中国建设整体布局规划》[2]和《“数据要素×”三年行动计划（2024—2026年）》[3]等政策文件，均强调了促进数据安全合规流通、有效利用和充分释放数据要素价值的重要性。人口健康科学数据涉及公共卫生、临床医学、生物医学、中医药学、药学、人口与生殖健康等多类学科领域，是赋能健康中国建设的重要基础性战略资源。在国家战略布局和政策制度的大力支持下，依托国家人口健康科学数据中心，我国人口健康科学数据实体汇交数量和体量增长快速，在基础设施建设和数据安全汇聚方面已取得阶段性成果。但由于涉及学科广泛、数据类型复杂，且不同数据生产机构的数据口径、数据标准不统一等[4-5]，人口健康科学数据管理与利用过程仍面临着数据资源质量参差、标准化程度低、流通阻力大、分析利用率不高等严峻挑战，极大地限制了科学数据开放共享进程的推进和数据价值的充分挖掘。

标准规范建设是实现多源异构数据资源标准化整合和共享利用的重要基础保障。面向人口健康科学数据共享管理的瓶颈问题，以及“数据要素×”行动计划和《“十四五”卫生健康标准化工作规划》[6]等对我国人口健康科学数据标准化工作和数据安全流通、融合创新等提出的新需求与挑战，亟须完善人口健康科学数据标准化顶层设计和协调机制，构建覆盖科学数据全生命周期、分类层次清晰的标准体系框架，从而指导人口健康科学数据共性标准和垂直领域标准规范的制修订工作。

本研究采用文献研究法和调查法，总结现有通用领域科学数据标准体系框架和卫生健康信息标准体系框架在指导人口健康科学数据标准化方面的不足，并系统梳理人口健康科学数据共享管理相关的各级标准规范的建设现状，分析当前人口健康科学数据标准化工作的研究情况与存在问题。围绕国家人口健康科学数

收稿日期：2024-09-27
*本研究得到中国医学科学院医学与健康科技创新工程项目“人口健康科学大数据智能管理与高效利用技术体系建设”（编号：2021-I2M-1-057）资助。

据管理和共享利用战略方针以及国家人口健康科学数据中心实际业务需求，提出面向科学数据生命周期、标准类型、标准类别、学科领域多维视角的人口健康科学数据标准体系框架，以期从标准维度为推动我国人口健康科学数据的规范化管理、提升国家人口健康科学数据中心数据资源建设质量和数据共享利用率奠定坚实基础。

1 相关研究

科技资源是从事科研活动的人力、物力、财力以及组织、管理、信息等软硬件要素的总称[7]。科学数据作为信息资源的一种，是科技资源的重要组成之一。2002年我国启动科学数据共享工程，人口健康是第一批9个试点领域之一。面向通用领域的科技资源及科学数据共享管理，多年来形成了3个较有代表性的标准体系框架。王志强等[8]围绕科技资源全生命周期管理共享流程提出了以定义与指南、资源描述、资源获取与处置、资源服务、评估评价和信息安全6个部分为主要内容的科技资源管理标准体系框架，并规范了科技资源的主要类型划分。依托国家科技基础条件平台专项课题，王卷乐等[9]综合考虑不同学科科学数据的共性特征和生命周期管理的业务需求，按照基础、通用、专用3个层次建立由定义与指南、科学数据描述、采集处理、汇交、保存维护、共享服务、评估评价以及安全8个部分组成的科学数据标准体系模型。胡良霖等[10]针对科学数据多源异构的复杂特点，制定中国科学院科学数据标准规范体系模型，重点包含为科学数据资源建设各生命周期阶段提供的专用标准和指导数据库建设与应用服务的指导标准。上述研究提出的标准体系框架均主要面向科技资源及科学数据标准化工作的共性需求，标准子类细分程度不足，且缺少针对不同类型、不同学科数据资源特征属性的针对性标准体系规划。

在人口健康科学数据共享管理领域，目前尚未提出具有针对性、指导性和协调性的标准体系框架，相关研究工作多参考国内卫生健康信息化建设过程中形成的标准框架。2009年原卫生部信息标准专业委员会提出包含基础、数据、管理和技术4个大类的卫生信息标准体系概念框架，该框架首次对我国卫生信息标准进行了归纳整理。2016年根据“十二五”期间的标准实践应用，国家卫生计生委统计信息中心与国家卫生标准委员会信息标准专业委员会对前概念框架进行了扩充完善，重新划分基础、数据、管理、技术、安全与隐私5个标准大类，同时结合业务需求提出由业务领域、标准类别、标准内容3个维度组成的医疗健康信息标准体系框架[11]。在此基础上，李赞梅等[12]以科学数据共享生命周期为主线，提出由基础、数据、管理、技术、服务与应用以及专用标准6个大类、23个小类组成的健康医疗科学数据共享标准体系框架。这些标准框架的研究为推动我国人口健康科学数据资源规范管理奠定了基础，但其研制的出发点多为医疗健康业务应用数据管理，如医疗服务、医疗保障等方面，并不完全适配人口健康科学数据，不足以支撑解决人口健康领域科学数据类型增多、规模快速增长，医学科技创新发展面临的资源建设质量参差、开放共享程度低、存在隐私安全风险等一系列问题。

本研究从全局视角进行人口健康科学数据标准化工作的顶层规划，旨在提出层次清晰、粒度适当的人口健康科学数据标准体系框架，从而推动人口健康科学数据在合法合规的基础上实现高效共享和深度应用。

2 人口健康科学数据标准化现状分析

2.1 标准规范研制现状

本研究从国家标准、行业标准、地方标准和团体标准4个层面进行国内人口健康科学数据相关标准规范的研制现状调研。标准采集基于全国标准信息公共服务平台、国家标准全文公开系统和国家标准馆等标准服务平台，截至2024年4月30日通过关键词检索和人工筛选共计得到863项标准，包括219项国家标准（含26项国标计划）、268项行业标准、64项地方标准和312项团体标准。纳入的标准涉及33项科学数据管理标准、45项科技平台和科技资源管理标准，其中除4项生殖健康专业领域科学数据团体标准和1项医学科学数据团体标准外，其他标准均为通用领域指导标准，目前尚缺少针对人口健康以及相关垂直学科领域的科学数据配套标准规范。此外，本研究还纳入了785项可为人口健康科学数据生命周期管理工作提供参考依据的标准规范。

总体而言，我国人口健康数据标准规范的制修订工作处于长期发展的过程。“十一五”期间的工作重点为基础类、数据类和以信息平台为核心的技术类、安全类

标准的研制，致力于推动信息共享[13-14]。“十二五”时期我国持续完善标准体系规划与建设，以业务需求为导向开展数据资源目录、数据元和值域代码、数据集标准建设，并开展标准测评工作[11,15]。在此阶段，全国科技平台标准化技术委员会（SAC/TC 486）首次出台9项科技平台领域国家标准，有力促进科技资源的规范管理和开放共享。“十三五”时期我国人口健康数据标准化建设进入快速发展阶段，不断拓展基础、数据、技术、管理、应用服务、安全保障等多类型标准的研制工作，同时促进相关标准的落地应用[16]。“十四五”时期人口健康数据标准化建设覆盖领域不断扩展。尤其针对科技平台和科学数据管理，全国科技平台标准化技术委员会发布20余项通用领域及人口健康相关专业领域国家标准及国标计划。此外，多家国家科学数据中心、中国标准化研究院等机构也着手科学数据团体标准制定工作[17]，截至2024年已发布实施的科学数据团体标准达30余项。

为更清楚地描述当前阶段标准研制的覆盖范围，了解标准化工作的研究情况与局限，本研究对标准所属的标准类别、学科领域和数据生命周期管理阶段进行分类统计，见表1~表3。标准类别涉及基础类、数据类、技术类、管理类、安全类和应用服务类6个大类；除通用领域外，标准研制还覆盖公共卫生、临床医学、生物医学、药学、中医药、生殖健康等学科领域；相关标准适用于数据规划、数据采集生产、数据加工整理、数据汇交、数据保存处置、数据发布共享、数据再利用等一个或多个数据管理活动阶段。

调研发现，我国人口健康数据标准研制总体表现为国家标准、行业标准和团体标准的发展相对平衡，但地方标准建设有待完善。已发布的数据类标准多为行业标准（44.5%）和团体标准（37.9%），而基础类和安全类标准则表现为国家标准占比较大，分别为该类标准总数的69.0%和57.9%。通用领域的标准研制表现为以国家标准（55.3%）为主，以行业（11.9%）、地方（9.0%）和团体标准（23.8%）为辅。而在专业学科领域，除公共卫生领域中行业标准发布量占比最高（74.4%）外，其他学科的标准建设均以团体标准为主，如临床医学和中医药领域的团体标准占比分别为80.9%和70.0%。从标准类别角度看，数据类标准（58.6%）的发布量最为突出，面向居民健康档案、电

表1　人口健康科学数据标准规范的标准类别分类情况

单位：项

级别	基础类	数据类	技术类	管理类	应用服务类	安全类	总计
国家标准	49	56	45	23	2	44	219
行业标准	3	225	35	2	2	1	268
地方标准	3	33	14	6	3	5	64
团体标准	16	192	42	25	11	26	312
总计	71	506	136	56	18	76	863

表2　人口健康科学数据标准规范的学科分类情况

单位：项

级别	通用领域	公共卫生	临床医学	生物医学	药学	中医药	生殖健康	总计
国家标准	172	11	4	9	0	23	0	219
行业标准	37	212	17	0	2	0	0	268
地方标准	28	26	5	2	2	1	0	64
团体标准	74	36	110	18	14	56	4	312
总计	311	285	136	29	18	80	4	863

表3　人口健康科学数据标准规范的生命周期分类情况

单位：项

级别	数据规划	数据采集生产	数据加工整理	数据汇交	数据保存处置	数据发布共享	数据再利用	多阶段	总计
国家标准	49	63	5	5	18	18	1	60	219
行业标准	3	240	5	0	3	13	1	3	268
地方标准	2	46	0	0	0	3	3	10	64
团体标准	16	210	12	2	12	22	7	31	312
总计	70	559	22	7	33	56	12	104	863

子病历、卫生信息管理、医疗卫生服务、疾病控制、妇幼保健和临床检查等业务应用领域，以及公共卫生、临床医学、中医药和生物医学等专业学科领域，陆续发布一系列数据元目录、值域代码和数据集标准，为各业务应用领域数据的采集、存储、共享和系统平台开发等提供了标准依据。其次是技术类标准（15.8%），主要包括针对交互平台建设、数据加工处理、数据互联互通、人群队列技术规范等标准的研制。此外，虽然近年来我国数据标准化建设逐渐向数据安全、数据管理、测试评估、数据应用服务等方面扩展，但基础类、安全类、管理类和应用服务类标准的整体占比相对较低。

2.2 人口健康科学数据标准化现存问题

人口健康数据标准化建设覆盖范围不断扩大，其中由政府主导研制的国家标准和行业标准的管理相对规范，但地方标准和团体标准却缺少有效的监督管理，存在不同地区、部门、机构、平台的标准重复建设、分散建设以及多头管理等现象。近年来团体标准发布数量增长迅猛，但标准质量参差不齐，标准内容也存在交叉重复或相互矛盾之处，降低了标准的应用价值。此外，由于科学数据管理目的和应用范围与业务应用数据的不同，以及人口健康科学数据在专业性、隐私性、积累性、开放性和可用性等方面的特殊需求，人口健康科学数据的汇交管理、开放共享和安全保护有了更高要求，已发布的标准规范仅在当前没有针对性科学数据标准支撑的情况下可为科学数据生命周期管理工作提供参考，但并不能完全适配。

针对科技资源及科学数据共享管理，我国已发布或计划发布的标准规范主要围绕通用性的数据资源描述、标识、汇交、收割、出版以及引用等方面，并重点布局数据安全保障标准建设，已推出科学数据权益保护、分级分类、安全管理、安全审计、安全传输等14项国标计划和团体标准。但这些已发布的通用领域科学数据标准并不能完全满足人口健康领域多学科、多类型科学数据共享管理的实际需求。一方面，数据类标准研制方面存在针对性标准不足的问题，缺少结合人口健康多学科领域特点的数据资源一致性描述、数据分类、数据编目等标准，以及可用于数据资源识别、发现、定位等的标识规则。另一方面，人口健康科学数据的质量控制与评价、产权归属、隐私伦理以及平台技术安全、应用服务安全等方面的管理过程与技术要求也亟待细化和明确。此外，面向人口健康科学数据生命周期管理视角，人口健康科学数据规划、采集生产、加工整理、汇交、保存处置、共享发布和数据再利用的方法、流程有待进一步规范，且缺少技术、管理和应用服务类标准，特别是针对不同专业学科领域，各生命周期管理节点均有待完善标准支撑。

3 人口健康科学数据标准体系框架构建

面向国家人口健康科学数据管理与共享利用的战略需求，基于国内通用领域科技资源和科学数据标准体系框架，以及当前人口健康信息化建设基础，结合国家人口健康科学数据中心的实际业务需求，本研究从科学数据生命周期管理、标准类型、标准类别、学科领域4个维度，归纳提出人口健康科学数据标准体系框架，不同维度相互交叉并派生标准子类，见图1。

3.1 科学数据生命周期

2018年国务院发布《科学数据管理办法》，首次从国家宏观层面明确加强科学数据全生命周期管理的重要意义[18]。标准规范的制定应贯穿人口健康科学数据的全生命周期流程，从而指导、规范科学数据管理的各环节业务活动。国外对科学数据生命周期管理的理论研究已较为成熟，面向通用领域[19]以及社会科学[20-22]、地球和环境科学[23]、地质科学[24]等不同学科领域形成一系列典型数据生命周期模型。国内相关研究机构和学者也以科技文献数据[25]、通用领域科学数据[9,26]、社会科学数据[27]和农业科学数据[28]等为研究对象提出了多种科学数据生命周期模型。本研究参考国内外相关科学数据生命周期管理模型的核心要素，结合国家人口健康科学数据中心开展人口健康科学数据汇交、保藏、共享管理以及挖掘分析等服务的实际业务需求，将人口健康科学数据生命周期管理流程划分为数据规划、数据采集生产、数据加工整理、数据汇交、数据保存处置、数据发布共享和数据再利用7个关键阶段。

3.2 标准类型划分

标准类型的划分规范了人口健康科学数据生命周

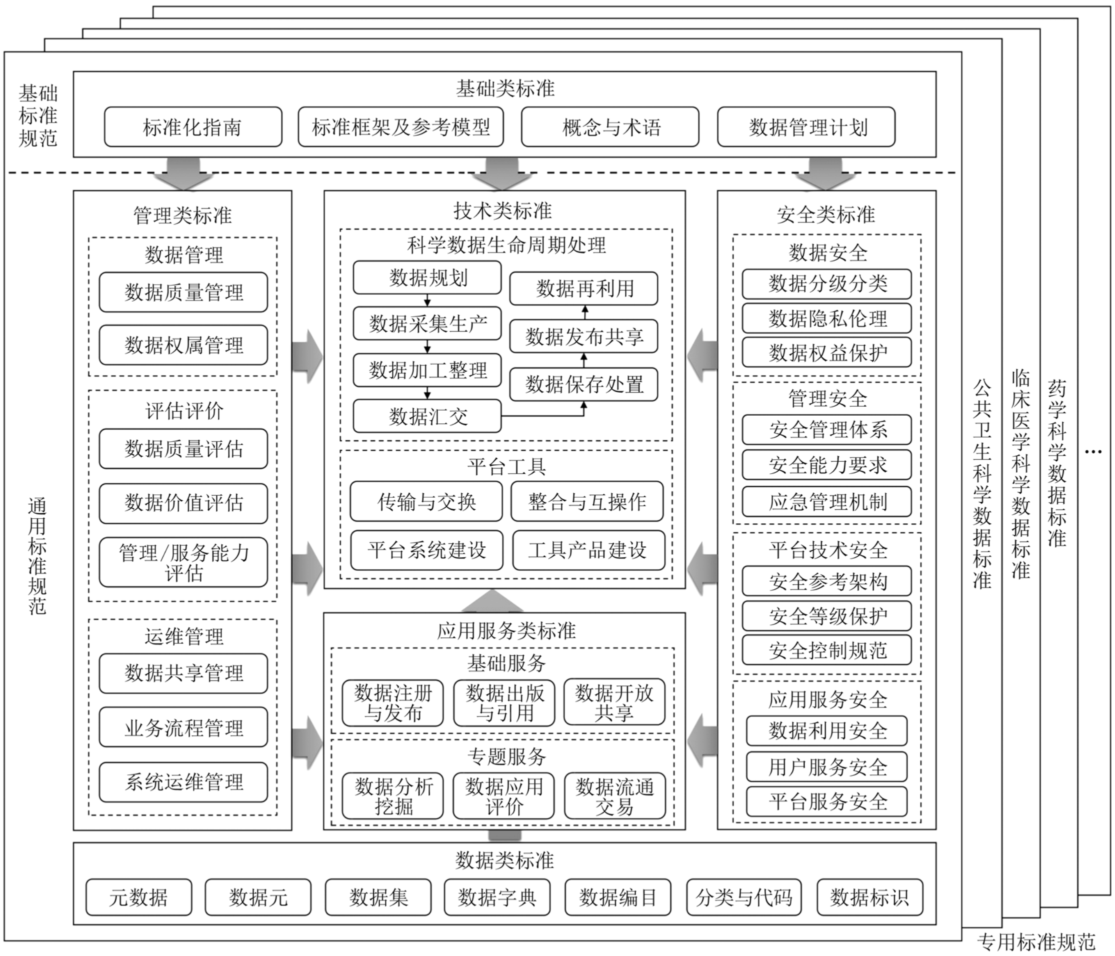

图1 人口健康科学数据标准体系框架

期管理标准体系的整体层次结构，一般为上层基础标准、中间层通用标准以及下层专用标准。基础标准通常是为人口健康科学数据标准的制定、协调、统一理解与应用等提供基础遵循的标准。通用标准是人口健康科学数据生命周期管理过程中的相关共性标准。专用标准是根据通用标准制定，符合人口健康领域特定学科数据特点和数据生命周期管理需求的标准。

3.3 学科领域划分

国家人口健康科学数据中心是科技部和财政部认定的20个国家科学数据中心之一，建设了国内人口健康领域数据规模最大、综合性最强的人口健康科学数据仓储（Population Health Data Archive，PHDA），已连续多年开展国家财政预算支持的各级科技项目数据、期刊论文关联数据和其他数据的汇交工作，积累了海量的科学数据资源。根据国家人口健康科学数据中心数据资源汇聚情况，标准体系应覆盖公共卫生、临床医学、生物医学、药学、中医药和生殖健康等学科领域。

3.4 标准类别划分

面向《“十四五”卫生健康标准化工作规划》对卫生健康信息标准体系的建设要求以及“数据要素×”行动规划对激发数据要素乘数作用的迫切需要，结合调查分析得出的人口健康科学数据标准化建设工作面临的实际问题，本研究将人口健康科学数据标准体系划分为基础类标准、数据类标准、技术类标准、管理类标准、安全类标准和应用服务类标准6个类别，各类别下细分不同标准子类，构成标准体系的主要内容。

（1）基础类标准。基础类标准在人口健康科学数据生命周期管理过程中发挥指导、协调作用，包括标准化指南、标准框架及参考模型、概念与术语以及数据管理计划4个标准子类。这些标准有助于从源头约束规

范人口健康科学数据的生命周期管理过程，形成统一的数据理解和解释，提高数据质量。尤其是概念与术语的标准化有助于实现人口健康科学数据在语义层面的一致性，这对于建设大规模知识图谱和人口健康领域专题数据库，实现从海量数据中发现有价值的信息至关重要。

（2）数据类标准。数据类标准是人口健康科学数据标准体系框架的核心部分，用于明确人口健康科学数据的资源描述与分类标识规则，以便通过科学的原则方法提升数据标准化程度，包括元数据、数据元、数据集、数据字典、数据编目、分类与代码、数据标识7个标准子类。人口健康科学数据资源多来源于专业细分领域，不同学科、不同病种的科学数据在组织格式、语义化程度等方面存在专业性差异，数据类标准的研制需要综合考虑科学数据共享管理的共性需求和数据资源的个性特征，为建设高质量数据库奠定坚实基础。

（3）技术类标准。技术类标准主要规定人口健康科学数据生命周期管理各阶段过程所需的数据处理技术和相关平台工具建设要求。其中：科学数据生命周期处理类标准涉及人口健康科学数据的规划、采集、加工、汇交、保藏、共享、再利用等环节的流程与技术要求；平台工具类标准包括不同数据平台、数据生产机构之间的数据传输与交换、数据整合与互操作、平台系统建设、工具产品建设4个标准子类，指导跨领域、跨机构的人口健康科学数据的整合分析、大数据技术与大数据分析技术研发等实践活动。

（4）管理类标准。管理类标准贯穿人口健康科学数据生命周期管理的各阶段，为数据资源建设与应用服务等提供重要支撑，确保数据资源的有序组织。管理类标准包括3类。其中，数据管理类标准包括人口健康科学数据质量管理和权属管理：数据质量管理是确保人口健康科学数据准确性、可用性和可靠性的基础，而数据权属管理涉及数据多个流转环节和多个主体之间的责任、权利的转移与认证。评估评价类标准主要涉及人口健康科学数据质量评估、数据价值评估、服务和管理能力评估3个方面，制定科学合理的评估方法与可量化的评估指标体系有助于促进科学数据资源规范化管理。此外，运维管理类标准则包括人口健康科学数据共享管理、业务流程管理、系统运维管理3个标准子类。

（5）应用服务类标准。应用服务类标准主要分为基础服务标准和专题服务标准，其中：基础服务标准主要涉及人口健康科学数据的注册发布、出版引用、开放共享3个标准子类；专题服务标准主要为人口健康科学数据分析挖掘、应用评价、流通交易等活动提供参考依据。人口健康科学数据资源来源广泛，存在隐私和伦理风险，使得数据分析应用和流通交易形式更为复杂，需要建立标准以保障合法合规。

（6）安全类标准。安全类标准同样贯穿人口健康科学数据生命周期管理的各阶段，为数据资源、技术环境、运维管理和应用服务等方面提供安全保障。其中：数据安全类标准包括人口健康科学数据分级分类、隐私伦理、权益保护3个子类；平台技术安全类标准包括面向人口健康科学数据平台的安全参考架构、安全等级保护、安全控制规范3个子类；管理安全类标准包括人口健康科学数据安全管理体系、安全能力要求、应急管理机制3个子类；应用服务安全类标准包括平台服务安全、用户服务安全、数据利用安全3个子类。

4　人口健康科学数据标准化工作建议

（1）以顶层标准框架为指导，细化人口健康科学数据标准明细。我国科学数据管理的标准建设工作正处于起步阶段，尤其针对人口健康领域科学数据的共享利用缺少统筹性的基础指导标准。人口健康科学数据具有涵盖学科众多、数据类型复杂、包含隐私信息等特点，其规范组织、合规流转、融合利用等数据价值实现工作面临更多挑战。在后续工作中，应以本研究提出的人口健康科学数据标准体系框架为宏观指导，引领人口健康科学数据标准化工作的有序、深入开展，并进一步梳理形成人口健康科学数据标准体系明细表，持续推动人口健康科学数据领域相关标准规范的完善与统一，避免标准建设不一致、不兼容等问题，充分发挥标准规范对人口健康科学数据资源建设和开放共享服务的支撑作用。

（2）以实际业务需求为核心，完善科学数据生命周期标准覆盖。本研究从多个维度对人口健康科学数据标准体系框架的标准子类进行划分，以期充分满足人口健康科学数据生命周期管理过程中对标准规范的多元需求，但部分标准子类面临相关标准文件的编制空白及问题。结合“数据要素×”行动计划等国家政策的总体要求，当前亟须加快建设人口健康科学数据确权保护、数据安全与隐私保护、数据合规流通等急用的标

准，以保障人口健康科学数据的安全可控开放共享，使数据供得出、流得动。同时也要重点推进人口健康科学数据编目存储、数据质量控制、数据评估评价、数据共享服务、数据开发利用等方面的标准研制，提升人口健康科学数据资源建设质量，带动数据中心数据资源高质量供给与应用。

（3）协调多方参与加强标准供给，推进特色学科领域标准研制。我国科学数据相关标准规范研制工作主要采用政府主导与社会自主制定协同发展的新模式。国家标准具有全国范围的普适性，是整个标准体系的基石，但也存在着研制难度系数高、研制周期长、协调性不足等问题。相较而言，团体标准的研制具有流程灵活高效、参与群体广泛等特点，能更好更快地适应实际业务需求的创新变化[29]。近年来，国家层面政策也越来越鼓励发展团体标准[6,30]。人口健康科学数据涵盖公共卫生、临床医学、生物医学、中医药学、药学、生殖健康等众多学科，涉及大量来自人群队列、临床试验、生命组学研究的问卷、随访、药物、检验、测序和影像等多类型数据，数据资源异质性大。实现人口健康科学数据标准化还应充分发挥各级各类医疗卫生机构、科研院所、高等院校和行业协会等社会组织的作用，基于不同学科特色和实际场景应用需求，统筹协调多方共同参与团体标准研制，分阶段、有重点地推进各类标准建设，增加优质标准供给，将标准化工作从基础共性标准向专业领域标准延伸拓展。

（4）加强标准质量与应用效果评价，做好标准宣贯与贯彻落实工作。标准质量和实施效果是影响标准作用发挥的重要因素[31]。基于地方标准和团体标准质量参差、建设冲突等问题，国家层面应加强对地方和团体标准研制与管理过程的监管，建立标准规范质量评价体系，保障标准在规范性、协调性、适用性等方面符合业务应用要求。近年来，各领域国家科学数据中心为促进科学数据汇交管理与开放共享也陆续建立了相应标准规范，并推动团体标准建设，但部分标准仅在中心内部使用，有待在行业内进一步宣贯推广。有必要制定科学高效的标准宣贯实施、监管与应用反馈机制，发挥国家级科学数据中心的带头作用，推动各级各类数据中心和相关部门、机构做好标准规范的试点应用工作，在确保标准规范有效落地实施的同时，基于实际应用过程的意见反馈不断进行标准的完善与修订。

5 结语

人口健康科学数据是赋能健康中国建设、推动数据驱动的医学科技创新的重要基础资源。本研究在全面调研、梳理国内人口健康科学数据标准化工作现状的基础上，深入分析了人口健康科学数据标准研制工作的弱势领域和存在问题，同时从新时期数据要素创新发展的战略需求和人口健康科学数据全生命周期管理的实际业务需求出发，对人口健康科学数据标准体系框架的层次结构、应涵盖的标准类别与细化内容进行规划设计。本研究提出的人口健康科学数据标准体系框架填补了当前阶段国内人口健康科学数据领域标准化指导框架的空白，其应用实施将有助于指导不同机构、不同学科、不同层次类别标准规范的制修订工作，提高不同标准规范间的协调应用能力，以科学合理的标准体系布局为高质量人口健康科学数据资源建设和数据要素价值充分释放提供有力支撑。

参考文献

[1] 中共中央 国务院关于构建数据基础制度更好发挥数据要素作用的意见[EB/OL]. [2024-08-09]. http://www.gov.cn/zhengce/2022-12/19/content_5732695.htm.

[2] 中共中央 国务院印发《数字中国建设整体布局规划》[EB/OL]. [2024-08-09]. http://www.gov.cn/xinwen/2023-02/27/content_5743484.htm.

[3] 国家数据局等17部门近日联合印发《“数据要素×”三年行动计划（2024—2026年）》[EB/OL]. [2024-08-09]. https://www.ndrc.gov.cn/hdjl/yjzq/202312/P020231215685140119139.pdf.

[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] 盛小平，田婧，向桂林. 科学数据开放共享中的数据质量治理研究[J]. 图书情报工作，2020，64（22）：11-24.

[6] 国家卫生健康委关于印发“十四五”卫生健康标准化工作规划的通知[EB/OL]. [2024-08-12]. http://www.gov.cn/zhengce/zhengceku/2022-01/27/content_5670684.htm.

[7] 唐仁华，伍莺莺，吴承春. 对促进科技资源共享问题的几点思考[J]. 科技创业月刊，2005，18（7）：8-9.

[8] 王志强，杨青海. 科技资源管理标准体系研究[J]. 标准科学，2019（3）：6-11.

[9] 王卷乐，石蕾，徐波，等. 我国科学数据标准体系研究[J]. 中国

科技资源导刊，2020，52（5）：45-51，77.
[10] 胡良霖，郑晓欢，朱艳华，等. 中国科学院科学数据标准体系研究与实践[J]. 图书馆，2019（11）：6-10.
[11] 汤学军，董方杰，张黎黎，等. 我国医疗健康信息标准体系建设实践与思考[J]. 中国卫生信息管理杂志，2016，13（1）：31-36.
[12] 李赞梅，钱庆，李姣，等. 健康医疗科学数据共享标准体系框架构建[J]. 医学信息学杂志，2018，39（11）：49-53.
[13] 李岳峰，胡建平，周光华，等. 我国卫生信息化建设：现状与发展[J]. 中国卫生信息管理杂志，2012，9（5）：7-10.
[14] 孟群. 我国卫生信息标准体系建设[J]. 中国卫生标准管理，2012（12）：24-31.
[15] 陈礼潮，孙金海，陈立富. 我国人口健康信息数据标准化进展[J]. 职业与健康，2018，34（16）：2287-2291.
[16] 李岳峰，胡建平，庹兵兵，等. 我国卫生健康信息标准建设成效与思考[J]. 中国卫生信息管理杂志，2021，18（3）：324-329.
[17] 廖方宇，朱艳华，胡良霖，等. 科学数据安全技术及基础技术标准研究[J]. 科技成果管理与研究，2023，18（5）：88-90.
[18] 张洋，肖燕珠. 生命周期视角下《科学数据管理办法》解读及其启示[J]. 图书馆学研究，2019（15）：37-43，13.
[19] HIGGINS S. The DCC curation lifecycle model[J]. International Journal of Digital Curation，2008，3（1）：134-140.
[20] Data Documentation Initiative Alliance. DDI lifecycle version 3.3[EB/OL]. [2024-08-12]. https://ddialliance.org/ddi-lifecycle.
[21] UK Data Service. Research data management[EB/OL]. [2024-08-12]. https://ukdataservice.ac.uk/learning-hub/research-data-management/.
[22] Inter-University Consortium for Political and Social Research. Guide to social science data preparation and archiving：best practice throughout the data life cycle·6th edition[EB/OL]. [2024-08-12]. https://www.icpsr.umich.edu/files/deposit/dataprep.pdf.
[23] MICHENER W K，JONES M B. Ecoinformatics：supporting ecology as a data-intensive science[J]. Trends in Ecology & Evolution，2012，27（2）：85-93.
[24] FAUNDEEN J L，BURLEY T E，CARLINO J A，et al. The United States Geological Survey Science Data Lifecycle Model：U.S. geological survey open-file report 2013—1265[EB/OL]. [2024-08-12]. http://dx.doi.org/10.3133/ofr20131265.
[25] 常志军，许丽媛，于倩倩，等. 基于生命周期模型的科技文献数据管理体系研究[J]. 农业图书情报学报，2022，34（6）：36-49.
[26] 储节旺，夏莉. 嵌入生命周期理论的科学数据管理体系构建研究：牛津大学为例[J]. 现代情报，2020，40（10）：34-42.
[27] 孟祥保，钱鹏. 数据生命周期视角下人文社会科学数据特征研究[J]. 图书情报知识，2017，34（1）：76-88.
[28] 高飞，周国民，满芮. 基于生命周期理论的农业科学数据中心化管理模式[J]. 大数据，2022，8（1）：24-36.
[29] 曹俐莉，靳宗振. 中国团体标准化发展历程与治理策略[J]. 科技导报，2022，40（11）：111-116.
[30] 中共中央 国务院印发《国家标准化发展纲要》[EB/OL]. [2024-08-12]. https://www.ndrc.gov.cn/fggz/fzzlgh/gjjzxgh/202112/t20211201_1306575_ext.html.
[31] 高艳玲，王志民，隋媛，等. 标准质量与实施效果评价方法及应用研究[J]. 标准科学，2020（5）：60-64，74.

作者简介

王安然，女，硕士，助理研究员，研究方向：医学科学数据标准化。
吴思竹，女，博士，研究员，通信作者，研究方向：医学科学数据管理和应用，E-mail：wu.sizhu@imicams.ac.cn。
李赞梅，女，硕士，助理研究员，研究方向：医学科学数据管理。
胡拯涌，男，硕士研究生，研究方向：医学科学数据标准化。

Construction of Population Health Scientific Data Standard System Framework

WANG AnRan WU SiZhu LI ZanMei HU ZhengYong
(Institute of Medical Information, Chinese Academy of Medical Sciences/Peking Union Medical College, Beijing 100020, P. R. China)

Abstract: Standards and specifications are pivotal in facilitating the open sharing, analysis, and utilization of population health scientific data. This research comprehensively investigates the development of standards and specifications related to population health scientific data within the country, providing a systematic analysis of the current status of standardization efforts and the limitations they face. Building upon the existing general scientific data standard system frameworks and the population health information standard system frameworks, a multi-dimensional perspective and clear hierarchical population health scientific data standard system framework covering the lifecycle management is proposed. Furthermore, the study also presents thoughtful suggestions on the standardization work of population health scientific data for moving forward.

Keywords: Standard System Framework; Scientific Data; Data Life Cycle; Population Health

（责任编辑：管清潆）