

卫生健康信息数据元标准 2023 版和 2011 版比较研究*

孙 函¹ 杨啸林² 杨 晟² 周 伟³ 张胜发³ 孟凡红¹ 刘丽红¹ 朱 彦¹

(¹ 中国中医科学院中医药信息研究所 北京 100700 ² 中国医学科学院基础医学研究所 北京 100005

³ 中国医学科学院国家人口健康科学数据中心 北京 100005)

〔摘要〕 目的/意义 对比分析卫生健康信息数据元标准 2023 版和 2011 版, 探讨差异与改进之处, 以期为标准的更新与实施提供有益的参考和指导。方法/过程 整理对比 2023 版较 2011 版卫生健康信息数据元标准的更新内容, 分析总结修订对标准化程度、规范化水平以及数据完整性和准确性等方面的提升作用。结果/结论 2023 版在数据完整性和规范化等方面实现显著提升。分析 2023 版标准在实施过程中可能面临的挑战和问题, 提出针对性建议和对策。

〔关键词〕 卫生信息; 数据元; 值域代码; 数据标准; 标准化

〔中图分类号〕 R-058 〔文献标识码〕 A 〔DOI〕 10.3969/j.issn.1673-6036.2024.08.003

A Comparative Study on the Version 2023 and the Version 2011 of the Standards of Health Information Data Elements

SUN Han¹, YANG Xiaolin², YANG Sheng², ZHOU Wei³, ZHANG Shengfa³, MENG Fanhong¹, LIU Lihong¹, ZHU Yan¹

¹Institute of Information on Traditional Chinese Medicine, China Academy of Chinese Medical Sciences, Beijing 100700, China; ²Institute of Basic Medical Sciences, Chinese Academy of Medical Sciences, Beijing 100005, China; ³National Population Health Data Center, Chinese Academy of Medical Sciences, Beijing 100005, China

〔Abstract〕 **Purpose/Significance** To compare and analyze the 2023 edition and 2011 edition of the health information data element standards, and to discuss the differences and improvements, so as to provide useful references and guidance for the update and implementation of the standards. **Method/Process** The updated contents of the 2023 and 2011 editions of the health information data element standards are sorted out and compared, and the effects of the revision on the degree of standardization, the level of standardization, and the completeness and accuracy of data are analyzed and summarized. **Result/Conclusion** It is found that the 2023 edition has achieved significant improvements in data completeness and standardization. Additionally, targeted suggestions and strategies are proposed for the challenges and issues that might be faced during the implementation of the 2023 edition standards.

〔Keywords〕 health information; data element; range code; data standard; standardization

〔修回日期〕 2024-06-07

〔作者简介〕 孙函, 硕士研究生; 通信作者: 朱彦, 刘丽红。

〔基金项目〕 中国中医科学院科技创新工程 (项目编号: CI2021A05409); 中国医学科学与健康科技创新工程项目 (项目编号: 2021-I2M-1-057)。

1 前言

卫生健康信息标准是为医学信息产生、处理及管理 etc 制定的各类规范和行动准则, 涉及整个医学事务处理过程^[1]。在统一的卫生健康信息标准支持下, 数

据可以实现高效的互联互通和共享^[2]。数据元是数据记录的基本单元,由对象类、特性和值域组成^[3]。卫生健康信息数据元标准作为连接医疗信息、促进数据互操作的基石,在提升医院和区域信息化建设水平、推动医院高质量发展等方面发挥十分关键的作用^[4]。

2023 年 10 月国家卫生健康委员会发布《卫生健康信息数据元目录 第 1 部分:总则》等 34 项推荐性卫生行业标准,并于 2024 年 4 月 1 日起施行。2023 版标准对 2011 版内容进行修订和更新,规范了医疗数据的格式和内容,提高了数据的质量和可用性,为医疗决策、科研分析及公共卫生管理提供有力支持。本文对 2023 版和 2011 版卫生健康信息数据元目录和值域代码进行比较分析,系统梳理 2023 版修订情况及疑似问题,并提出针对性的解决建议,以期作为标准的更新与实施提供有益的参考和指导。

2 研究目的及方法

2.1 研究目的

通过比较两个版本卫生健康信息数据元目录及

值域代码修订情况,分析探讨两个版本的差异与改进之处,对存在的问题提出针对性建议,以期为卫生健康信息数据元标准的更新与实施提供有益的参考和指导,推动卫生健康信息化发展。

2.2 研究方法

利用文献调研和比较分析法,系统梳理两个版本数据元目录和值域代码结构、内容方面的修订情况及疑似问题,针对性地提出后续研究和实施建议。

3 数据元目录修订情况分析

3.1 总体情况

2023 版数据元总数由 1 399 个增加到 2 200 个,各部分修订情况,见表 1。2011 版数据元数量最多的 3 个部分是“第 7 部分”“第 12 部分”“第 9 部分”,2023 版新增最多的是“第 3 部分”,修订后数据元数量最多的 3 个部分是“第 3 部分”“第 12 部分”“第 7 部分”。

表 1 数据元目录修订情况(个)

分册	名称	修改数量	新增数量	删除数量	2011 版数据元数量	2023 版数据元数量
第 1 部分	总则	0	0	0	0	0
第 2 部分	标识	0	11	0	13	24
第 3 部分	人口学及社会经济学特征	16	285	0	62	347
第 4 部分	健康史	20	10	1	90	99
第 5 部分	健康危险因素	14	9	1	98	106
第 6 部分	主诉与症状	19	41	0	119	160
第 7 部分	体格检查	66	26	7	241	260
第 8 部分	临床辅助检查	14	32	1	51	82
第 9 部分	实验室检查	73	18	1	129	146
第 10 部分	医学诊断	60	97	0	73	170
第 11 部分	医学评估	65	51	5	127	173
第 12 部分	计划与干预	83	144	0	175	319
第 13 部分	卫生健康费用	10	42	0	10	52
第 14 部分	卫生健康机构	4	9	0	53	62
第 15 部分	卫生健康人员	4	3	0	31	34
第 16 部分	药品与医疗器械	19	25	0	25	50
第 17 部分	卫生健康管理	12	14	0	102	116
合计		479	817	16	1 399	2 200

3.2 内容修订

3.2.1 新增 2023 版新增数据元共 817 个,帮助

医生更全面地了解患者的健康状况,进行疾病预防和筛查,如“第 4 部分健康史”中新增了遗传性疾病史、个人史、既往史等数据元,“第 13 部分卫生

健康费用”中依据《住院病案首页》新增诊断类、治疗类、中医类、中药类等相关数据元 42 个。

3.2.2 删除 2023 版删除数据元 16 个。删除最多的是“第 7 部分体格检查”，共删除 7 个数据元。

3.2.3 数据元值的数据类型与允许值类型的修订 2023 版“第 1 部分总则”中，对数据元值的数据类型为字符型的数据元和允许值之间的对应关系给出了更明确的描述，见表 2。

3.2.4 数据元属性修订 2023 版数据元属性的修订共有 479 处。(1) 数据元名称修订。部分数据元名称增加“描述”后缀，对数据元允许值部分引用相关代码表的数据元增加“代码”后缀，还有部分数据元虽进行修订但未改变其内涵。(2) 文字描述

变化。如时间、日期型数据元在原有基础上增加“完整描述”，与“门诊”有关的数据元修订为“门（急）诊”。(3) 对不同类型数据元允许值做出修订。包括枚举型和非枚举型。(4) 数据元描述中更明确其所针对的对象。具体情况，见表 3。

表 2 数据元值的数据类型和允许值之间的关系

数据元值的 数据类型	数据元允许值类型	示例
S1	不可枚举且以字符描述	DE02. 01. 010. 00 电话号码
S2	枚举型，列举值不超过 3 个	DE02. 10. 091. 00 数据元住院情况代码
S3	代码表	DE04. 50. 001. 00 ABO 血型代码

表 3 数据元属性修订情况

修订项	修订内容	示例
数据元名称	增加“描述”	DE02. 10. 022. 00 “过敏史”——“过敏史描述”
	增加“代码”	DE06. 00. 289. 00 “医嘱项目类型”——“医嘱项目类型代码”
	未改变其内涵	DE06. 00. 163. 00 “治疗终止日期”——“治疗结束日期”
数据元描述	修改数据元的描述，未改变其内涵	DE02. 01. 002. 00 常住地类型代码中定义“长期……”——“经常……”
	时间、日期类数据元，增加“完整描述”	DE05. 01. 034. 00……纪年日期——……纪年日期的完整描述
	“门诊”修改为“门（急）诊”	DE07. 00. 002. 00 卫生健康费用中“……门诊费用……”——“……门（急）诊费用……”
数据元允许值	非枚举型，“在特定编码体系中的代码”修改为“在特定编码体系中的编码”	DE04. 10. 038. 00 体格检查浮肿程度代码定义中将“……在特定编码体系中的代码……”——“……在特定编码体系中的编码……”
数据元对象	定义中增加“患者”	DE04. 10. 118. 00 脉率“每分钟脉搏……”——“患者每分钟脉搏……”
	已有对象如“个体/受检者”修改为“患者”	DE04. 10. 093. 00 脊柱检查结果“受检者脊柱检查……”——“对患者脊柱检查……”
	特别的疾病/检查或有明确针对性的检查	DE04. 50. 115. 00 乙型肝炎病毒表面抗体检测结果代码“乙型肝炎病毒……”——“对乙肝患者进行乙型肝炎……”

3.3 疑似问题及勘误

3.3.1 部分数据元修订描述有误 2023 版中存在修订描述与实际修订不一致的情况。如“第 7 部分”明确描述对数据元“齿列描述、呼吸频率（次/min）”的定义进行了修订，但经核对并未发现修订；其次是少量被提及数据元，实际并不存在，如“第 12 部分”中数据元“关后核对标志”在数据元名称修订描述中有提及，但是并未找到

该数据元。
3.3.2 数据元标识符变更带来的潜在问题 2023 版对“第 3 部分人口学及社会经济学特征”中数据元标识符进行了修订，包括出生日期、值班电话、地址－省（自治区、直辖市）等 25 个数据元。以“出生日期”为例，2011 版中标识符为 DE02. 01. 005. 01，2023 版修订为 DE02. 01. 005. 00，修订的目的应该是实现扩展场景数据元，但数据元标识符的改变，会造成已实施的数据集

或系统都需要做相应调整，才能实现与新版本保持统一。

3.3.3 数据元标识符删除带来的潜在问题 2023 版中存在删除的数据元标识符仍然被使用的情况，若系统同时使用两个版本的数据元可能会导致数据的不一致性。为了确保数据的正确解读与交换，应确保数据元标识符的唯一性，已删除的数据元其标识符也不应回收使用。

表 4 数据元值域代码修订情况

分册	名称	新增数量（个）	删除数量（个）	示例
第 3 部分	人口学及社会经济学特征	2	0	CV02. 01. 204 医疗保险类别代码表
第 5 部分	健康危险因素	1	0	CV03. 00. 305 传播途径代码表
第 6 部分	主诉与症状	1	0	CV04. 01. 015 肺结核患者症状及体征代码表
第 7 部分	体格检查	4	0	CV04. 10. 024 宫颈检查结果
第 8 部分	临床辅助检查	17	0	CV04. 30. 003 医学影像检查方法代码表
第 9 部分	实验室检查	1	0	CV04. 50. 020 输血品种代码表
第 10 部分	医学诊断	7	0	CV05. 01. 042 患病危险行为类别代码表
第 11 部分	医学评估	19	0	CV05. 10. 019 入院病情代码表
第 12 部分	计划与干预	17	0	CV06. 00. 235 中医药健康管理服务类型代码表
第 13 部分	卫生健康费用	1	0	CV07. 00. 002 住院费用分类代码表
第 16 部分	药品与医疗器械	0	1	
第 17 部分	卫生健康管理	3	0	CV09. 00. 403 入院途径代码表
总计		73	1	

4.2 内容修订

4.2.1 随数据元目录变化进行的修订 2023 版新增的数据元中，部分数据元允许值为非枚举型的数据元引用了相关的标准或者值域代码，为了确保数据元的正常使用，需要新增缺少的值域代码表，如医疗保险类别代码表。

4.2.2 适应新需求而新增的编码 科技的进步和医疗实践的创新不断带来新的诊断方法、治疗手段和疾病分类，数据元值域代码表也需要与时俱进，涵盖这些新的医疗信息和数据。如代码表“CV07. 10. 004 医疗费用结算方式”新增加代码值和值含义“09 移动支付”“10 数字化人民币”“11 绿色通道”，新增的结算方式可以满足互联网

4 值域代码修订情况分析

4.1 总体情况

2023 版数据元值域代码对 12 个部分进行修订和完善，新增代码表 73 个，删除代码表 1 个，“第 11 部分医学评估”新增最多。具体修订情况，见表 4。

时代支付方式多样化需求。

4.3 疑似问题及勘误

数据元目录与值域代码表不匹配。2023 版数据元值域代码表中删除了“第 16 部分药品与医疗器械”交通工具代码表，但是引用该代码表的数据元 DE08. 50. 003. 00 交通工具类别代码并未被删除且仍然引用该值域代码表，数据元目录应与数据元值域代码表统一，删除的代码表在数据元目录中也不应再出现。

5 讨论

5.1 2023 版带来的提升

在对卫生健康信息数据元标准两个版本进行

比较研究后,发现 2023 版在数据标准化和规范性方面取得显著提升。主要体现在以下 3 方面。一是适应了新需求和新形式。数据元标准需要随着实际需求变化而不断修订和更新,2023 版中新增了疾病诊断、治疗方式、药品信息等 817 个数据元,修订了表示格式 156 个,以适应医疗健康领域的新需求,促进医疗健康服务的改进和提升。二是提升了一致性与规范性。2011 版中部分数据元表示不够规范统一,如枚举不超过 3 个的数据元值的数据类型有用 S1 表示的,也有用 S3 表示的。2023 版对数据元值的数据类型和数据元允许值的对应关系给出了更明确的描述规则,规范化后的数据类型表示有助于理解和表达医学领域的数据和信息,促进信息共享。三是增强了互操作性和共享性。修订后的标准有助于在不同系统、不同领域之间建立统一编码规则,保证数据的一致性,减少代码不规范或不一致导致的数据错误,从而提高数据的准确性;通过统一的标准,不同医疗机构之间的信息系统将能够更好地互联互通,有助于优化医疗资源配置,提高医疗服务的协同性和连续性。

5.2 实施挑战与应对建议

新旧版本之间的转换和兼容面临诸多挑战,如数据元文件描述有误、数据元标识符的一致性、数据元目录与值域代码及领域数据集的统一等问题。对此提出以下建议,以期能够推动卫生健康信息数据元标准新旧版本的转换与过渡,更好地支持现代医疗体系下信息互通共享。一是发布贴近实际应用场景的多样化版本。数据元相关信息标准的制定和发布,应适应数字化实施场景的实际需求。在标准制定过程中,建议使用数据库及信息系统进行管理维护,以确保数据的完整性和一致性,减少手动编辑可能出现的笔误或疏漏。标准发布时,除了采用传统的纸质和非结构化版本,建议发布电子版甚至结构化的数据库版,方便推广和使用。二是构建统一的数据元资源平台。以实现数据元的注册、审批、发布和版本等全生命周期的管理和服务,以及不同领域的数据开放

共享和协同应用^[5]。国外已有多较成熟的元数据注册系统投入使用,如澳大利亚 METEOR 平台^[6]、美国 caDSR II 平台^[7]、德国 DZL 系统^[8]等。建议参考国外已经建立的管理机制和平台系统的优点,结合国内实际情况,制定和开发符合国内需求的管理平台。三是建立持续更新机制。数据元标准的更新、完善、维护是一项常态化的工作,有必要建立健全更新维护机制,实现定期评估标准的适用性、收集用户反馈、修订完善标准等,以确保数据元标准与时俱进。另外,要满足新旧版本之间的兼容性,保证使用端数据的延续性和一致性,旧版本标准已作废的数据元及编号,原则上保留,不能在新标准中回收利用。四是加强领域内数据元标准复用的规范和协调。由于各种原因,各子领域标准的制定经常出现冲突和不一致的情况。例如,团体标准《T/CIATCM002—2019 中医药信息数据元目录》中的“穿刺部位”数据元占用 2023 版卫生信息数据元目录“本次随访形式类别代码(DE06.00.410.00)”的标识符。类似冲突可能导致数据在跨领域、跨标准交换时出现混淆。因此,有必要对这类复用情况进行规范,加强子领域间的协调和沟通,确保数据元在不同数据集和应用场景中的一致性和准确性。五是实现更高层面的语义一致性和互操作性。当前我国卫生信息数据标准化方法已经从整理数据元步入构建语义模型阶段^[9],现有数据元通过场景修饰和编码无法完美实现多粒度的语义层级描述。已有相关研究采用信息建模法进行卫生信息数据标准开发^[10],建立初级卫生信息标准体系^[11]和国家健康概念数据模型^[12],为多源数据的标准化和管理提供元数据参考。此外还要基于信息模型或本体技术进行更复杂和深入的语义建模与表示,以期满足更高层面的语义一致性和互操作性需求。

6 结语

本文通过比较分析卫生健康信息数据元标准 2011 版与 2023 版,评估修订带来的改进,同时指

出可能存在的问题，例如文件描述的不准确、数据元目录与值域代码及数据集的一致性等问题。基于这些问题，建议可以发布多版本标准，构建数据元管理平台；同时利用现有的信息模型或本体进行更深层次的语义建模与表示，以期简化数据元的使用并提高管理的规范性。随着信息技术和医疗卫生领域的持续进步，数据元标准需要不断更新，融入新技术元素，以满足数字化和智能化的卫生服务需求。此外，还应密切关注国际标准和国内实践的最新动态，并加强与国际标准化组织、行业协会以及医疗机构的合作与交流，共同推进数据元标准的制定、修订和实施，为提升医疗卫生服务的质量、促进健康事业的发展提供坚实的支撑。

作者贡献：孙函负责数据收集与整理、论文撰写；杨啸林、杨晟负责协助数据收集与整理、提供指导；周伟、张胜发、孟凡红负责参与论文撰写；朱彦、刘丽红负责研究设计、论文审核与修订。

利益声明：所有作者均声明不存在利益冲突。

参考文献

1 孟群. 我国卫生信息标准体系建设 [J]. 中国卫生标准管理, 2013 (9): 9.

2 姚克勤, 赵菲, 邱五七. 我国公共卫生信息化建设发展历程、问题与建议 [J]. 医学信息学杂志, 2021, 42 (11): 42 - 45.

3 张文丽, 姚瑞, 曹广, 等. 中医药经验传承服务信息基本数据集标准研制 [J]. 医学信息学杂志, 2022, 43 (2): 2 - 7.

4 李岳峰, 胡建平, 吴士勇. 国家医疗健康信息互联互通标准与测评体系建设 [J]. 中国卫生信息管理杂志, 2023, 20 (1): 7 - 12.

5 张胜发, 罗葳, 马玉环, 等. 关于加快健康医疗大数据向科学数据转化的思考与建议 [J]. 医学信息学杂志, 2022, 43 (11): 8 - 13.

6 Australian institute of health and welfare hospital resources [EB/OL]. [2024 - 05 - 21]. <https://meteor.aihw.gov.au/>.

7 CADSR (Cancer data standards registry and repository) [EB/OL]. [2024 - 05 - 21]. <https://cadsr.cancer.gov/>.

8 CoMetaR - a collaborative metadata repository [EB/OL]. [2024 - 05 - 21]. <https://data.dzl.de/>.

9 张黎黎, 汤学军, 刘丹红, 等. 卫生信息互联互通数据标准化方法研究 [J]. 中国卫生信息管理杂志, 2016, 13 (5): 467 - 472.

10 赵霞, 刘丹红, 李小华, 等. 卫生信息数据标准开发方法研究 [J]. 中国数字医学, 2019, 14 (8): 22 - 25.

11 XIA Z, XIAOHUA L, WEI Y, et al. Primary health information standard system based on semantic interoperability [J]. BMC medical informatics and decision making, 2018, 18 (S5): 112.

12 YANG Z, JIANG K, LOU M, et al. Defining health data elements under the HL7 development framework for metadata management [J]. Journal of biomedical semantics, 2022, 13 (1): 10.

(上接第 13 页)

14 徐伟. 公共数据权属：从宪法国家所有到民法国家所有权 [J]. 当代法学, 2024, 38 (1): 121 - 133.

15 关于加强行政事业单位数据资产管理的通知 [EB/OL]. [2024 - 05 - 08]. https://www.gov.cn/zhengce/zhengceku/202402/content_6931055.htm.

16 数据安全法 [EB/OL]. [2024 - 03 - 22]. http://www.npc.gov.cn/npc/c2/c30834/202106/t20210610_311888.html.

17 MILLER H G, MORK P. From data to decisions; a value chain for big data [J]. IT professional, 2013, 15 (1): 57 - 59.

18 李海舰, 赵丽. 数据成为生产要素：特征、机制与价值形态演进 [J]. 上海经济研究, 2021 (8): 48 - 59.

19 马费成, 熊思玥, 孙玉姣, 等. 数据分类分级确权对数据要素价值实现的影响 [J]. 信息资源管理学报, 2024, 14 (1): 4 - 12.

20 盛小平, 吴红. 科学数据开放共享活动中不同利益相关者动力分析 [J]. 图书情报工作, 2019, 63 (17): 40 - 50.

21 袁昊. 数据的财产权构建与归属路径 [J]. 晋阳学刊, 2020 (1): 103 - 111.