

# 《中国肿瘤登记数据集标准》编制说明

中国肿瘤登记数据集标准编制项目组

中图分类号:R73-31 文献标识码:A  
文章编号:1004-0242(2021)08-0588-03  
doi:10.11735/j.issn.1004-0242.2021.08.A003

## 1 基本情况

### 1.1 概述

肿瘤登记是制定癌症防控政策及评价癌症防控效果的基础性工作。2008年,原卫生部设立“肿瘤登记随访项目”并纳入“国家重大公共卫生专项中央财政转移支付项目”,肿瘤登记工作在全国范围展开。目前,我国肿瘤登记处覆盖区县1 152个,覆盖人口约5.98亿,较为全面地掌握了我国癌症发病、死亡、生存状况及流行趋势。2015年原国家卫生和计划生育委员会、国家中医药管理局联合下发《肿瘤登记管理办法》,从制度上保证了全国肿瘤登记工作的顺利开展。

近20年来我国人群肿瘤登记工作进展迅速,肿瘤登记数据采集、上报及管理机制基本形成,数据监测网络体系逐步完善。肿瘤登记工作的高效开展离不开医疗卫生信息系统的支撑,大数据平台和人工智能等互联网技术的迅猛发展,为在全国范围内开展人群肿瘤登记工作提供了强大的技术支撑,进一步为肿瘤防治体系的信息化建设奠定坚实基础。制定口径统一、语义明确、结构规范的数据集标准是肿瘤登记信息化建设的重要环节。为此,国家癌症中心联合国内多所医疗卫生机构,集中专家力量编制了《中国肿瘤登记数据集标准》。

本标准规定了中国肿瘤登记工作相关元数据属性、数据元目录和数据元值域代码表。旨在为肿瘤登记信息的数据交换和共享提供一套术语规范、定义明确、语义语境无歧义的数据集数据元标准。

### 1.2 起草单位

#### 1.2.1 本标准起草牵头单位

国家癌症中心

#### 1.2.2 本标准主要参与起草单位

浙江省肿瘤医院、北京大学肿瘤医院暨北京市肿瘤防治研究所、湖北省肿瘤医院、空军军医大学、河北医科大学第四医院、上海市疾病预防控制中心、江苏省疾病预防控制中心、广东省疾病预防控制中心、四川省疾病预防控制中心、河南省肿瘤医院、安徽省疾病预防控制中心、福建省肿瘤医院、西藏自治区疾病预防控制中心、陕西省疾病预防控制中心、青海省疾病预防控制中心、深圳市慢性病防治中心、合肥市疾病预防控制中心、磁县肿瘤防治研究所、扬中市肿瘤防治研究所。

#### 1.2.3 本标准主要起草人员

赫捷、魏文强、张思维、孙可欣、郑荣寿、曾红梅、王少明、陈茹、杜灵彬、王宁、刘硕、张敏、庾吉妤、王霞、李辉章、贺宇彤、梁迪、顾凯、吴春晓、周金意、韩仁强、夏亮、成姝雯、刘曙正、陈琼、戴丹、周衍、白国霞、扎西宗吉、王艳平、周素霞、彭绩、雷林、张小鹏、宋国慧、华召来。

### 1.3 起草过程

2019年6月,该项目由国家癌症中心牵头负责,成立了由疾控中心、慢病所、医疗机构、医学高等院校等单位共同组成的标准编制项目组,项目组专家涵盖全国31个省(自治区、直辖市)的各级各类专家共79人。国家癌症中心向中国卫生信息与健康医疗大数据学会卫生信息标准专业委员会提交团体标准项目立项申请书。

#### 1.3.1 需求分析及技术路线设计

2018年10月,国家癌症中心根据多年工作经

验,充分分析肿瘤登记工作对专病数据规范化的信息需求,确定了数据集标准制订的技术路线。

(1)制订依据:本数据集制订的依据基础是国家《WS/T 303 卫生信息数据元标准化规则》《WS/T 305 卫生信息数据集元数据规范》等相关标准规范,同时根据近几年国内卫生信息标准化对专病数据的标准,制订一套术语规范、定义明确、语义语境无歧义的肿瘤登记数据集。

(2)现场调研:选取湖北省肿瘤医院等单位作为调研对象,了解肿瘤登记实际工作流程,作为数据集制订的基本依据。同时通过访谈,了解病案统计人员、医生及卫生信息技术人员对肿瘤登记数据集的信息需求。

(3)修订方案的确定:基于编制依据和现场调研结果,制定切实可行的数据集的设计方案,主要包括确定主要内容要求、术语定义等。

### 1.3.2 基本架构设计

根据上述技术路线,2018年10月至2019年6月,项目组召开多次全国肿瘤登记专家研讨会,并对北京大学肿瘤医院、浙江省肿瘤医院、深圳市慢性病防治中心等医疗卫生机构开展相关基础调研,收集大量数据内容需求,并对内容项进行了归类。在综合分析调研资料的基础上,最终形成了数据集标准的基础架构。

### 1.3.3 内容编制

本项目于2019年11月获得批准立项,国家癌症中心开始牵头筹备标准制订的各项准备工作。

2020年4月初,国家癌症中心组织召开标准撰写工作会议。项目组确定《中国肿瘤登记数据集标准》团体标准制订的技术路线,并对标准的封面、目次、前言、标准正文、索引、附录等部分进行撰写,形成标准初稿。

2020年4月28日,国家癌症中心以在线视频方式组织进行第2次标准撰写专家组研讨会,针对第一次专家会上遗留的问题及对应的修改意见逐条进行讨论,完善标准初稿。本次会议上,标准撰写专家组完成了标准初稿的撰写工作,并在会后广泛征询标准顾问专家组专家对初稿的意见。

2020年8月21日,国家癌症中心组织召开第3次现场专家研讨会,来自17个单位的22名标准撰写组专家及4名顾问组专家参会。会上,专家再次论

证标准的适用性及可行性,针对标准初稿中的问题及有争议部分进行深入讨论,最终达成共识,形成《中国肿瘤登记数据集标准》。

2020年12月1日,《中国肿瘤登记数据集标准》在卫生健康信息团体标准预审会上接受专家审核,专家对文件提出6条修改意见。会后,国家癌症中心参考专家意见,对文件进行局部修改。

2021年5月13日,通过中国卫生信息与大数据学会卫生健康信息团体标准委员会会审,并于2021年7月15日正式发布。

### 1.4 主要起草单位及其承担工作

国家癌症中心全面负责本标准的制订工作。浙江省肿瘤医院、北京大学肿瘤医院、湖北省肿瘤医院、深圳市慢性病防治中心、空军军医大学(第四军医大学)等单位参与完成标准内容的整理、论证及修改工作。

## 2 与我国法律法规及其他标准的关系

本标准参照《GB 11643 公民身份号码》等国家标准,《WS/T 305 卫生信息数据集元数据规范》等卫生行业标准,以及国际上相关肿瘤登记标准与技术规范,并遵从《肿瘤登记管理办法》规定,修订编制而成。本标准不存在与我国的法律法规及其他标准相抵触的问题。

## 3 标准的制订与起草原则

本标准以我国肿瘤登记数据的科学采集方法及长期实践经验为依据,以满足我国肿瘤登记数据采集工作的实际需求为目的,同时确保与国际标准接轨。本数据集标准从元数据、数据元属性等方面对肿瘤登记主要业务环节数据结构的设计进行指导,促进数据的互联互通。本数据集的制订遵照以下原则:

(1)一致性原则:数据元素的选择以现行的业务表单为基础,遵从国家及行业的相关标准与规范,符合相关的法律法规。

(2)标准化原则:按照相关的卫生信息数据元标准,规范标准的应用范围、术语和定义、缩略语、数据集元数据和数据元属性。

(3)前瞻性原则:综合考虑肿瘤登记工作发展趋

势及相关数据需求,以期满足未来肿瘤登记数据集涵盖内容的丰富性。

#### 4 确定本标准各项内容的依据

本标准的规范性引用文件包括:《GB 11643 公民身份号码》《GB/T 2260 中华人民共和国行政区划代码》《GB/T 2261.1 个人基本信息分类与代码第1部分:人的性别代码》《GB/T 2261.2 个人基本信息分类与代码第2部分:婚姻状况代码》《GB/T 2261.4 个人基本信息分类与代码第4部分:从业状况(个人身份)代码》《GB/T 3304 中国各民族名称的罗马字母拼写法和代码》《GB/T 4658 学历代码》《GB/T 4761 家庭关系代码》《GB/T 14396 疾病分类与代码》《GB/T 18391.3 信息技术 数据元的规范与标准化-第3部分:数据元的基本属性》《WS/T 303 卫生信息数据元标准化规则》《WS/T 305 卫生信息数据集元数据规范》《WS/T 306 卫生信息数据集分类与编码规则》《WS 218 卫生机构(组织)分类与代码》《WS 363 卫生信息数据元目录》《WS 364 卫生信息数据元值域代码》《WS 370 卫生信息基本数据集编制规范》《WS 445.10 电子病历基本数据集第10部分:住院病案首页》,以及《中国肿瘤登记工作指导手册(2016)》等规范性书籍。

#### 5 征求意见和采纳意见情况

标准撰写专家组针对243条具体问题进行集体讨论并达成共识,形成标准初稿。顾问组专家及标准编制专家对初稿进行审阅,提出17条修改意见,项

目组对修改意见全部采纳并形成标准文件。国家癌症中心向省级肿瘤登记机构发函征询对标准文件的建议与意见,共收到21家单位、共计98条反馈意见,其中部分被采纳。卫生信息团体标准预审会、正式评审会上专家分别提出6条与7条评审意见,全部采纳。

#### 6 重大意见分歧的处理结果和依据

无重大意见分歧。

#### 7 根据需要提出实施标准的建议

《中国肿瘤登记数据集标准》团体标准对于肿瘤登记数据采集的规范统一具有重要的意义。因此,建议加强对本标准的宣贯工作,强化宣贯工作的组织领导,为参与此标准实施的相关单位及人员提供咨询服务,确保有关人员充分认识本标准的重要作用和意义,全面掌握、准确理解并贯彻实施本标准内容。同时,积累应用经验,为本标准的下一轮修订与完善奠定基础。

2019年国家癌症中心牵头建立中国肿瘤登记平台,为全国所有登记处提供网络报告数据的途径。本标准中所涉及的数据元是我国基于人群的肿瘤登记工作所采集的核心信息,同时也是构建中国肿瘤登记平台的核心数据元。标准的实施能够在肿瘤登记数据采集、数据质量控制和数据分享交换等多个环节中发挥重要的指导作用,从而有效满足我国肿瘤登记工作扩大覆盖范围、提升数据质量、增强使用效率的业务需求。