

肿瘤流行病学专题

• 专家述评 •



[专家简介] 张文彬, 硕士研究生、研究员, 四川省癌症防治中心副主任、四川省有突出贡献的优秀专家、四川省卫生和计划生育委员会学术技术带头人; 中华预防医学会肿瘤预防与控制专业委员会常务委员、四川省免疫学会常务理事、四川省慢性非传染性疾病预防与控制专家委员会副主任委员; 《肿瘤预防与治疗》杂志副主编。主要从事肿瘤预防和肿瘤免疫基础研究工作, 作为课题负责人或参与者先后承担十余项国家及省市科研项目, 作为主编撰写专著两部。

探索大数据背景下肿瘤登记信息平台建设的新模式*

张文彬[△], 刘潇霞, 乔良, 李博

610041 成都, 四川省肿瘤医院·研究所, 四川省癌症防治中心, 电子科技大学医学院 癌症防治中心办公室

[摘要] 恶性肿瘤是严重危害我国公民健康的重大慢性疾病, 防控形势十分严峻。肿瘤登记是制定肿瘤防控策略, 评价防控效果的基础。我国肿瘤登记经过近 60 年的发展, 取得显著成绩的同时也存在覆盖面不足、数据深度广度不够、及时性欠佳等问题。近年来政府高度重视, 多次提出要充分发挥大数据平台优势和信息整合共享理念, 高质量发展肿瘤登记。因此, 如何充分利用现有医学信息资源, 依托大数据挖掘和人工智能技术等, 搭建一套及时、准确、全域覆盖的肿瘤监测平台具有重要现实意义, 并且是未来大势所趋。

[关键词] 肿瘤登记; 大数据; 监测; 平台建设; 信息化

[中图分类号] R73-31; R730.1 [文献标志码] A doi:10.3969/j.issn.1674-0904.2019.11.001

引文格式: Zhang WB, Liu XX, Qiao L, et al. Exploring a new model for the construction of cancer registries in the era of big data[J]. J Cancer Control Treat, 2019, 32(11):951-954. [张文彬, 刘潇霞, 乔良, 等. 探索大数据背景下肿瘤登记信息平台建设的新模式[J]. 肿瘤预防与治疗, 2019, 32(11):951-954.]

Exploring a New Model for the Construction of Cancer Registries in the Era of Big Data

Zhang Wenbin, Liu Xiaoxia, Qiao Liang, Li Bo

Department of Cancer Prevention and Treatment, Sichuan Cancer Hospital & Institute, Sichuan Cancer Center, School of Medicine, University of Electronic Science and Technology of China, Chengdu 610041, Sichuan, China

Corresponding author: Zhang Wenbin, E-mail: zhangwenbin855@163.com

This study was supported by grants from Science & Technology Department of Sichuan Province (NO. 2019JDPT0008).

[Abstract] Cancer is a major chronic disease that seriously endanger people's health in China. The situation of cancer prevention and control is serious. Cancer registration is the basis of developing the strategy of cancer prevention and control and evaluating the effect of cancer prevention and control. The development of our country's cancer registration in the past 60 years has achieved remarkable achievement, and also has the problems of insufficient coverage, insufficient data depth

and poor timeliness. In recent year, the government has attached great importance to the high quality development of cancer registration by making full use of the advantage of the big data platform and information integration and sharing. Therefore, how to make full use of the existing medical infor-

[收稿日期] 2019-10-12 [修回日期] 2019-11-12

[基金项目] *四川省科技基础条件平台重点项目(编号: 2019JDPT0008)

[通讯作者] [△]张文彬, E-mail: zhangwenbin855@163.com

mation resources, rely on big data mining and artificial intelligence technology and so on, set up a set of real-time, accurate and globally covered cancer monitoring platform is of great practical significance, and will be future trends.

[Key words] Cancer registration; Big data; Monitoring; Platform construction; Informatization

恶性肿瘤是严重影响我国公民健康的重大慢性疾病,近 20 年来我国恶性肿瘤发病率呈逐年上升趋势,成为我国居民的主要死因,严重危害居民健康,带来沉重的个人、家庭、社会经济负担^[1-2],已成为目前中国乃至全球最严重的公共卫生问题之一。获得高质量的肿瘤登记信息,及时掌握癌症的发病动态,是现阶段开展人群癌症防治研究的基础性工作。为此,本文通过回顾肿瘤登记历史、现状、问题等,结合新形势下的工作要求,借助信息化技术探索大数据背景下肿瘤登记的新模式。

1 我国恶性肿瘤形势严峻,国家高度重视

肿瘤登记是目前国际公认的肿瘤发病、死亡及生存信息收集的方法,是一项按一定的组织系统经常性地搜集、储存、整理、统计分析和评价肿瘤发病、死亡及生存资料的统计机制^[3]。开展肿瘤随访登记,掌握准确的人群肿瘤发病、死亡及生存数据,可以全面评估恶性肿瘤的流行特征,是制定卫生事业发展规划和肿瘤防控策略,评价防控效果的基础性工作。据国家癌症中心 2018 年年报最新肿瘤登记数据估计,全国恶性肿瘤发病率为 285.83/10 万,世标发病率为 186.39/10 万;恶性肿瘤死亡率为 170.05/10 万,世标死亡率为 105.84/10 万^[4]。随着我国经济的高速发展,工业化、城市化、老龄化进程的加剧以及人们生活方式的改变,我国的癌谱逐渐发生了变化,形成了以消化道癌症为主的发展中国家癌谱(胃癌、食管癌等,泛称“穷癌”)向发达国家癌谱(肺癌、乳腺癌、大肠癌等,泛称“富癌”)的过渡阶段,构成了“穷癌”、“富癌”并存的格局,防治形势复杂^[5]。随着恶性肿瘤的危害日益加重,我国对肿瘤防控愈加重视,对肿瘤登记的要求也逐渐提高^[2,6-7]。

近年来,国家先后出台的《健康中国行动(2019—2030 年)》、《“健康中国 2030”规划纲要》、《“十三五”卫生与健康规划》、《中国防治慢性病中长期规划(2017—2025 年)》、《中国癌症防治三年行动计划(2015—2017 年)》等一系列肿瘤防治相关政策性文件均将建立健全符合我国社会经济水平的肿瘤随访登记体系,开展全人群肿瘤发病、死亡及生存信息监测等,作为我国肿瘤防治工作的重点目标和任务。2019 年 2 月,国家卫健委副主任李斌介绍

我国在癌症防治方面重点工作时更是重点提出了要充分发挥大数据平台优势和整合共享理念。推进实现肿瘤登记工作在全国所有县区全覆盖,搭建国家癌症大数据平台,建成覆盖全国的癌症病例登记系统,推动与相关监测系统的数据交换和共享利用,及时掌握我国癌症发病动态,有针对性开展防治工作。

2 肿瘤登记发展状况

2.1 全球肿瘤登记组织众多,登记工作逐渐完善

全球肿瘤登记以欧洲、北美和澳洲开展较早,1929 年德国汉堡建立了世界第一个以人群为基础的肿瘤登记处;1942 年丹麦成立了第一个全国性的肿瘤登记处^[8],经过多年的积累及发展,逐渐形成了较为完善的登记系统。国际上涉及癌症诊治评价内容的肿瘤登记组织相继成立,如世界卫生组织下属国际癌症研究机构(International Agency for Research on Cancer, IARC)的 GLOBOCAN 项目、欧洲肿瘤登记网(European Network of Cancer Registries, ENCR)及澳洲肿瘤登记协会(Australian Association of Cancer Registries, AACR)等^[6]。一些登记体系的登记流程、登记资料、数据质量及分析日趋规范。

2.2 美国肿瘤登记系统完善、贡献突出

美国是肿瘤登记开展较好的国家,美国国家癌症研究所(National Cancer Institute, NCI)、疾病预防与控制中心和北美中心癌症登记处协会(North American Association of Central Cancer Registries, NAACCR)等机构均有参与肿瘤登记,其开展登记时间长,质量高,是被《五大洲癌症发病率》收录登记处资料最多的国家。其中,以 NCI 建立的 SEER 数据库最为突出,也是北美最具代表性的大型肿瘤登记注册数据库。SEER 项目始于 1973 年,目前共有 18 个地区的登记处加入,是全球最准确、最完整的以人口为基础的癌症登记处。SEER 数据库覆盖全美 28% 人口的癌症登记,并且广泛覆盖了美国的不同人群^[1],是美国唯一以人群为基础,高质量地收集了包括癌症患者人口学特征、原发部位、形态学、分期、肿瘤标志物、首次治疗和随访等资料的癌症报告系统,提供及时、准确和连续性的癌症资料和数据,监测北美癌症趋势。与此同时,SEER 数据库资料允许相关机构及人员通过多种方式获取,用于深入研究,并且为了便于肿瘤数据的分析,SE-

ER 开发了一系列软件实现频数和率分析、生存分析、风险分析、趋势分析等。目前可获得的 SEER 数据库包括:SEER 数据、标准人口数据、美国人口数据、美国死亡数据、县特性数据以及一些与 SEER 链接的数据库资料^[3]。值得注意的是,SEER 的肿瘤原始数据主要来源于医院、临床医生、病理医生等,并在后期与医院保持紧密联系,补充完善核实患者的病历、分期和治疗等临床信息,这是 SEER 数据长期保持高质量并被作为世界癌症登记处的质量标准的重要原因。收集内容广泛的高质量数据、积极与各机构合作共享并开展多样化研究的登记模式为全球的登记工作开启了新的思路。

2.3 中国肿瘤登记快速发展,取得可喜成绩

中国肿瘤登记起步于 20 世纪 50 年代,经过近 60 年的发展,登记点数量逐步扩增,登记内容逐步规范,经费投入逐年增加,并且分别在 1973 年、1990 年和 2006 年开展了 3 次全国死因回顾性调查,明确了恶性肿瘤是我国居民主要死因以及我国各时期的癌谱分布^[9-11],至今全国的肿瘤登记及监测随访网络初步建成。目前,全国登记处数量已经达到 574 个,登记地区覆盖人口 4.38 亿,占 2017 年全国人口数的 31.51%^[12]。为进一步加强肿瘤登记工作,2015 年国家卫生健康委员会(原国家卫生和计划生育委员会)和国家中医药管理局联合发布《肿瘤登记管理办法》,这也是我国第一个关于肿瘤登记工作的法规性文件^[12]。2017 年发布的《五大洲癌症发病率》第 11 卷中,我国已有 36 个肿瘤登记点的发病资料被收录,分布在我国 15 个省、自治区和直辖市,我国的肿瘤登记在国际上已具备一定的影响力。

3 肿瘤登记存在的问题

随着信息化发展和大数据时代的到来,传统的肿瘤登记存在的问题也逐步凸显出来,国家癌症中心魏文强主任提出我国肿瘤登记的不足主要体现在肿瘤登记点数量不足和分布不均衡、数据深度和广度不足、监测数据时效性不高以及现行的传统监测手段对数据可获得性较差等方面^[12]。各省因肿瘤登记起步时间、信息化程度不同,其问题的严重程度有所不同。目前全国各地肿瘤登记存在一些共性问题:

3.1 覆盖面不足

由于受经费限制,大部分地区实际工作中按照肿瘤登记技术流程和规范开展信息收集的地区覆盖有限,有些省份仅局限于有财政经费支持登记点,覆盖比例不足,省级代表性不够。

3.2 数据质量欠佳

《五大洲癌症发病率》数据纳入标准对登记资料中诊断依据不明比例、病理诊断比例、仅有死亡医学证明书比例、原发部位不详或未特指病例比例、癌种变化趋势稳定等指标做出了要求^[13],目前全国各肿瘤登记处数据质量达标的并不多,这与传统的信息采集经历了太多人工环节密切相关,因人为因素导致数据失实,数据质量和可信度不高。

3.3 数据时效性差

现有的肿瘤登记模式在数据收集、整理、分析方面均需要较长时间。通常情况,当年的肿瘤新发病例数据需要经过约 3 年的补充漏报才能稳定,导致数据支持循证决策的时效性不足。

3.4 登记信息深度不足

以人群为基础的肿瘤登记只能获取发病、死亡的基本信息,信息量少,普遍缺乏肿瘤病理、分期、诊治和生存转归等诊疗信息,对深入研究带来较大困难。

3.5 随访效率低下

医务人员主动与患者接触开展主动随访的难度大、成本高,有些地区甚至没有开展随访;全国尚无健全的信息化系统监测患者就诊轨迹,主动随访、被动随访双薄弱。

综上,现有的肿瘤随访登记模式已无法完全满足日益增长的肿瘤防控政策需求和科研需求,模式转换迫在眉睫。

4 探索大数据背景下构建省级肿瘤随访登记大数据服务平台

由于各级医疗机构信息化建设的飞速发展,电子病历得到广泛使用,诊疗数据的高度信息化、结构化,所以肿瘤登记资料越来越多的来源于医疗机构病案首页。与此同时,越来越多的研究也提示借助信息平台开展肿瘤登记工作是未来的发展趋势,国内学者分别提出了以医保人群为基础、以医院为基础开展肿瘤登记的设计思路^[14-15]。利用现有医学信息资源,依托大数据技术,实现肿瘤登记资源整合、信息共享,借助人工智能,深入挖掘登记信息是解决目前肿瘤登记诸多问题的关键手段^[16-17]。

4.1 整合医学信息资源,搭建肿瘤监测大数据平台

探索搭建省级肿瘤随访登记大数据服务平台,整合现有的全省住院病案首页、死因监测等信息资源,结合信息化技术深度采集诊疗信息,构建肿瘤患者发病、死亡、随访及全生命周期数据库,实现多途

径数据源统筹管理与交换。同时平台与基层医疗机构卫生信息管理系统对接,在基层开展肿瘤患者主动随访,信息实时反馈。最终构建集肿瘤发病、死亡、生存监测,患者随访管理为一体的标准化基础资源平台,满足服务患者、服务科研、服务政府决策三位一体的肿瘤防治体系信息化目标。

4.2 借助信息化技术,保证监测时效性

充分利用省级卫生计生统计数据综合采集平台,借助覆盖全域的病案首页信息,实现资源整合和数据交汇,是搭建省级“平台”的基础与关键,解决全省肿瘤登记数据采集时效性不足问题,实现省域内全人群肿瘤发病、死亡及生存信息监测,推进数据开放共享、深度挖掘应用,满足卫生行政管理者、医疗卫生工作者、科研工作者及肿瘤患者等多方需求。

4.3 运用人工智能技术,实现监测数据的精准化

基于病案首页构建的省级“平台”,数据资源立足于医院诊疗信息,运用文字抓取技术、人工智能等方式,准确及时采集肿瘤详细的诊断信息(如肿瘤 TNM 分期、诊断依据、病理学类型、分化程度等)、治疗信息(如手术、放化疗、介入治疗、生物治疗等)、随访信息(如生存质量、生存状态、死亡原因、死亡时间)以及危险因素信息(如既往病史、家族史、饮食行为习惯等),累积一套高精度监测数据,便于开展深入的癌症防控方面研究。

4.4 开发质控模块,保证监测数据质量

全国肿瘤登记数据连续多年被《五大洲癌症发病率》收录的登记点,大多起步早、依托医疗机构采集数据,如四川省盐亭县、上海市、浙江嘉兴市、浙江海宁市等,深耕多年,取得了可喜的成绩。“平台”建设初期,需充分依托大数据挖掘、人工智能以及信息抓取技术,深度提取病案信息,构建具备更详细诊疗信息的发病、死亡数据库,并参照 IARC《五大洲癌症发病率》对数据质量的相关要求,开发数据智能审核模块,评估资料完整性和可靠性。将割裂的人群为基础的肿瘤登记和医院为基础的肿瘤登记进行重塑、联系并融合互补。

4.5 主动、被动随访密切结合,提升随访效率。

肿瘤随访是肿瘤登记工作的重要组成部分,对临床科研具有较大指导作用。至今尚无成熟的肿瘤患者随访信息平台,构建省级大数据平台,通过主动抓取患者再次就医记录,自动整合生成被动随访信息,对无法通过被动随访获得生存资料的患者,依靠属地化管理由基层卫生服务机构开展主动随访,并在适当时机与疾控、医保、公安、民政等多部门进行

发病、死亡及生存数据交换,实现高效随访。

恶性肿瘤的精准监测数据越来越受到广泛的重视,业内专家提出精准监测数据是现阶段乃至未来很长一段时间内我国制定肿瘤防治政策、评价防治工作效果重要基础^[12]。探索通过大数据、信息化新技术建立共享的省域内肿瘤随访登记大数据服务平台,并与传统肿瘤登记数据互为补充。利用信息化手段推进肿瘤登记工作从质和量上更上一个新的台阶,同时也为广泛开展医院为基础的肿瘤登记奠定基础。

[参考文献]

- [1] Torre LA, Bray F, Siegel RL, et al. Global cancer statistics, 2012 [J]. CA Cancer J Clin, 2015, 65(2): 87-108.
- [2] Chen W, Zheng R, Baade PD, et al. Cancer statistics in China, 2015 [J]. CA Cancer J Clin, 2016, 66(2): 115-132.
- [3] 国家癌症中心. 中国肿瘤登记工作指导手册(2016) [M]. 北京: 人民卫生出版社, 2016.
- [4] 孙可欣, 郑荣寿, 张思维, 等. 2015 年中国分地区恶性肿瘤发病和死亡分析 [J]. 中国肿瘤, 2019, 28(1): 1-11.
- [5] 陈万青, 李贺. 人群肿瘤监测在癌症防控中的作用 [J]. 肿瘤预防与治疗, 2018, 31(1): 1-4.
- [6] 国家卫生计生委, 国家中医药管理局. 肿瘤登记管理办法 [EB/OL]. http://www.gov.cn/gongbao/content/2015/content_2868879.htm. [2015-02-04]/[2019-10-15].
- [7] 国家卫生计生委, 国家发展改革委, 教育部, 等. 中国癌症防治三年行动计划(2015-2017 年) [EB/OL]. <http://www.caca.org.cn/system/2015/10/28/011243026.shtml>. [2015-09-09]/[2019-10-15].
- [8] 杨念念, 严亚琼, 李卫平. 我国肿瘤登记发展方向探讨 [J]. 中国肿瘤, 2009, 18(5): 347-349.
- [9] 全国肿瘤防治研究办公室. 中国恶性肿瘤死亡调查研究(1990-1992) [M]. 北京: 人民卫生出版社, 2008: 1-629.
- [10] 陈竺. 全国第三次死因回顾抽样调查报告 [M]. 北京: 中国协和医科大学出版社, 2008: 14-17.
- [11] 孔灵芝, 赵平. 中国肿瘤死亡报告: 全国第三次死因回顾抽样调查 [M]. 北京: 人民卫生出版社, 2010: 10-22.
- [12] 魏文强, 赫捷. 大数据信息化背景下我国肿瘤登记工作的思考 [J]. 中华肿瘤杂志, 2019, 41(1): 15-18.
- [13] 王新正, 元芳梅, 张永贞. 阳城县肿瘤登记资料质量评价 [J]. 中国肿瘤, 2012, 21(1): 39-41.
- [14] 崔亚萍, 周守君, 黄汉明, 等. 以医保人群为基础的肿瘤登记方法研究 [J]. 中国肿瘤, 2018, 27(8): 584-588.
- [15] 杨国庆. 医院为基础肿瘤登记流程分析与设计 [D]. 晋中: 山西医科大学, 2012.
- [16] 魏矿荣, 刘慎超, 魏东霖, 等. 大数据对肿瘤登记发展的影响 [J]. 科学通报, 2015, 60(5): 491-498.
- [17] 徐向东, 胡建平, 周光华, 等. 区域全民健康信息化体制机制建设情况分析 [J]. 中国卫生信息管理杂志, 2018, 15(1): 24-28.