

# 胃癌数据库建立和管理的质量化控制

吴晓龙 李浙民 陕飞 李子禹

北京大学肿瘤医院暨北京市肿瘤防治研究所胃肠肿瘤中心一病区 恶性肿瘤发病机制及转化研究教育部重点实验室 100142

通信作者:李子禹,Email:ziyu\_li@hsc.pku.edu.cn

**【摘要】** 胃癌高质量数据库的建设有助于提高胃癌规范化诊治的水平和疗效。笔者中心建立了具有特色的单中心胃癌专病数据库,并通过胃肠肿瘤外科联盟推动多中心数据库模式探索。本文通过总结不同层面数据库的建设经验,深入阐述胃癌数据库建设及管理的质量化控制。基于信息化病历系统和多学科会诊模式,我们对数据库进行了模块化、多节点设计。结构化数据的自动调取与手工录入相结合,严格的多学科查房制度与数据录入相辅相成,不同节点实时监督强化了实时质控的效果,并通过实时审核与多级审核制度相结合,强调数据安全性以及管理计划信息化等措施,高质量地维护了数据库建立及管理流程。进一步通过中国胃肠肿瘤外科联盟平台的多中心数据库建设历程总结认为:应通过研究目标导向制定数据计划、充分考虑数据可及性和强调构建中心间数据共识等措施进行质量化控制。多中心数据库未来的建设应立足于稳定的公共数据共享平台,建立数据安全保障机制,并定期进行数据质量评估;同时,加强多中心合作与交流,进一步促进各中心医疗水平同质化。

**【关键词】** 胃肿瘤; 数据库; 数据质量控制; 中国胃肠肿瘤外科联盟

## Quality control in the establishment and management of gastric cancer database

Wu Xiaolong, Li Zhemin, Shan Fei, Li Ziyu

Department of Gastrointestinal Surgery, Key Laboratory of Carcinogenesis and Translational Research (Ministry of Education), Peking University Cancer Hospital & Institute, Beijing 100142, China

Corresponding author: Li Ziyu, Email: ziyu\_li@hsc.pku.edu.cn

**【Abstract】** The establishment of a high-quality gastric cancer database significantly improves the efficiency and standardization of diagnosis and treatment of this disease. Our center has developed a specialized, single-center gastric cancer database and initiated the China Gastrointestinal Cancer Surgery Union, catalyzing the exploration of multi-center databases. This article encapsulates multi-level experience and provides a detailed overview of the quality control methods we implement in both constructing and managing the gastric cancer database. Utilizing an electronic medical record system and a multi-disciplinary treatment (MDT) approach, we have designed the database in a modular and multi-nodal manner. A synthesis of automatic retrieval of structured data and manual entry, coupled with a rigorous MDT system and real-time supervision at various nodes, bolster our real-time quality control efforts. Ensuring data security and digitized management plans alongside real-time review protocol and a multi-level review system, we maintain the highest standards in the initiation and management of the database. Through the establishment of the China Gastrointestinal Cancer Surgery Union platform, we endorse the concept that multi-center database construction should be driven by research objectives, consider data accessibility, while placing an emphasis on building inter-center consensus on data quality control. Moving forward, it is crucial that the development of multi-center databases promotes uniformity in medical standards across centers, cultivates stable public data sharing platforms, ensures robust

DOI: 10.3760/cma.j.cn441530-20231119-00179

收稿日期 2023-11-19 本文编辑 卜建红

引用本文:吴晓龙,李浙民,陕飞,等.胃癌数据库建立和管理的质量化控制[J].中华胃肠外科杂志,2024,27(2):132-136. DOI: 10.3760/cma.j.cn441530-20241119-00179.



data security protocols, routinely conducts data quality assessments, and bolsters multi-center cooperation and exchanges to promote the homogeneity of medical standards.

**【Key words】** Stomach neoplasms; Database; Data quality control; the China Gastrointestinal Cancer Surgery Union

数据库建设对于恶性肿瘤诊疗质量控制和疗效提高具有重要意义。不同于花费较高、目的单一的多中心随机对照研究,数据库能够反映一个地区诊疗的整体水平,并提供诊疗经验总结,以期进一步提高诊治效率<sup>[1-3]</sup>。在胃癌领域,美国监测、流行病学和结果(Surveillance, Epidemiology, and End Results, SEER)数据库、日本的国家临床数据库(National Clinical Database, NCD)和荷兰的上消化道癌症审计(the Dutch upper gastrointestinal cancer audit, DUCA)数据库是目前国际上较为成熟的数据库,产生了大量循证医学证据,推动了诊疗质量的提高<sup>[4-8]</sup>。

我国胃癌规范化诊疗起步较晚,不同地区胃癌诊疗理念和水平以及医疗记录规范性存在差异,因此,不论单中心或多中心的数据库都还在其建设的初期阶段,建立良好的数据库质量控制体系可以有效提高数据质量<sup>[9-12]</sup>。北京大学肿瘤医院基于信息化病历系统和多学科会诊模式,建立了胃癌专病数据库并迭代更新优化,为临床决策提供大数据层面指导,同时为高质量临床研究课题提供数据支持<sup>[13-16]</sup>。自2014年开始,着手搭建胃肠肿瘤外科联盟(联盟),目前已收集近百家中心30余万例胃癌数据<sup>[17-18]</sup>;基于联盟推动建立外科并发症规范化诊断登记数据库,发表多中心研究论文<sup>[19-20]</sup>。笔者总结团队在不同层面数据库的建设经验,针对胃癌数据库建设及管理的质量化控制进行阐述,以期在分享经验的基础上,提供胃癌数据库质量化控制的思考和启示。

### 一、数据库的模块化和多节点设计

本中心最初基于Excel表格和多种数据库软件的摸索,尝试建立胃癌专病数据库。自2018年起,与医院信息科深度沟通后,依托于医院信息化平台,建立了基于在线网页和服务器架构的胃癌专病数据库平台。该种模式较之传统数据库软件优势如下:在录入和管理上易学易用,多浏览器兼容,不受设备操作系统限制,无需额外安装软件。

数据库采用模块化设计,主要由管理模块和数据录入模块构成。

管理模块由本中心专职数据工程师与本中心

信息部数据库工程师共同维护和优化,涉及词条管理、权限管理和质控管理等内容,例如词条术语的清晰定义、录入必填项的设置、异常值的质控提示以及时限提醒等。

数据模块则根据胃癌综合治疗的趋势,结合临床实际工作中治疗计划的变更等特点,进行多节点设计。目前涵盖的节点包括:(1)多学科查房制定初始治疗计划时;(2)新辅助治疗和转化治疗定期评效时;(3)接受手术治疗前1周内;(4)手术治疗出院后1~2周内;(5)随访期;(6)辅助治疗结束时总结;(7)复发后多学科查房制定治疗计划。不同节点具有对应录入页面,录入词条涉及基线资料、多学科治疗计划信息、综合治疗信息、评效信息、手术治疗信息和随访信息等。

### 二、数据录入的质量化控制

1. 病例纳入与数据来源:所有经门诊就诊、拟继续在本中心住院治疗的胃原发恶性肿瘤的患者,均会分配唯一ID号纳入本数据库。数据库病例自2018年起保持连续性纳入状态。

依托于电子病历、检验检查报告系统等不断改进的结构化数据模式,本数据库中涉及的大量词条可以直接调取获得,实现了病历、报告书写过程即数据库信息录入的过程。住院总医师和门诊医师负责多学科会诊治疗计划和治疗评效信息等需要手工录入的部分;随访信息等需要主管医师和数据工程师通过多渠道获取后录入数据库。

2. 数据的录入流程:数据录入流程遵循多节点录入的规则,除结构化数据的自动调取外,笔者主要阐述手工录入的关键节点:(1)多学科查房制定初始治疗计划时,由门诊医师和总住院医师实时完成基线数据及查房信息的录入;(2)新辅助治疗、转化治疗定期多学科查房评效时,由病房住院医师和总住院医师实时记录治疗不良反应、评效相关词条信息以及录入评效结果和后续治疗计划;(3)手术治疗出院后1~2周内、多学科查房制定后续治疗计划时,由主管医师、住院医师和总住院医师共同录入术中、术后并发症信息以及依据病理结果制定的治疗计划信息;(4)随访信息由专职数据工程师定期随访电话获取,另外,主管医师在门诊获取的患

者复发以及生存信息也需及时反馈数据工程师。

3. 质量化控制要点: 高质量数据库建立的实现要点包括: (1) 结构化数据的自动调取提高了数据录入效率以及准确率, 与手工录入相结合具有一定优势。(2) 严格的多学科查房制度与数据录入相辅相成, 保证了关键数据的实时录入; 同时, 在提高录入完成率的基础上, 通过边录入边审核的模式, 充分保证了手工录入的准确率。(3) 不同节点、不同人员参与, 各司其职, 相互监督, 强化了实时质控的效果。

### 三、数据管理的质量化控制

1. 审核与数据清理: 对于涉及诊断准确性或治疗决策的关键性数据词条, 需主管医生和多学科专家协同录入员录入, 如临床分期判定、治疗效果判定、治疗决策制定等。前瞻性地录入避免了回顾偏倚, 做到重要数据的实时审核。

采取多级审核制度: (1) 录入时, 系统后台对异常值或漏缺项以及矛盾数据须实时审核; (2) 数据提交后, 专职数据工程师通过查阅原始记录核查关键词条, 对于需修改数据进行标记, 并返回录入员复核完善; (3) 系统实时逻辑质控, 同时定期随机抽样, 进行整条数据的审核, 收集数据录入时的共性问题, 并持续改进; (4) 定期由数据工程师导出数据进行核查或清理, 进一步处理漏报、缺项、错误及矛盾数据, 形成数据核查报告, 在中心工作平台进行公示督导。

2. 人员培训与管理: 定期对参与数据录入及审核人员进行培训, 提高其对胃癌相关专业知识和患者全流程管理的理解, 有助于提高数据库的准确性和一致性; 定期以数据核查报告为基础, 建立录入人员数据质量评价报告, 督导持续改进。对于所有参与人员进行权限管理, 人员分工包括数据录入、数据核查和数据管理维护等, 需分别设置编辑修改权限、标记权限、检索权限和导出权限等。同时, 在数据临近录入时间窗时, 由平台向相关人员发送提醒并督导, 包括术后并发症、30 d 内返院等信息录入以及定期的随访更新等。

3. 数据安全性与使用规范: 为防止数据泄露及保护患者隐私, 应采用多种措施以保证数据库的安全性。严格的权限设置以保证仅有核心数据管理人员具有检索全部词条的权限。本数据库可用于多种用途, 如临床研究、治疗策略参考及个案报告等。数据使用者根据个人需求向中心工作平台提

出数据申请, 需阐明数据用途, 所需词条项目, 是否涉及临床试验, 是否通过伦理审批等情况, 经由中心核心组审批后, 由数据管理员导出所需数据。在数据管理人员进行数据导出操作时, 患者的身份识别信息会被直接加密, 数据使用者获取的是匿名化的数据, 同时仅包含研究计划所需的特定词条信息。

4. 质量控制要点: 数据库的管理维护, 是实现数据库应用价值的重要保障。其质量控制要点包括: (1) 实时审核与多级审核制度相结合, 夯实数据真实性、一致性和完整性; (2) 高度强调数据安全性, 流程持续优化, 保护患者隐私; (3) 管理计划高度信息化, 录入时限提醒、质控反馈、数据申请及导出等均为线上操作, 降低沟通成本, 提高管理效率。

### 四、多中心数据库建立与管理的探索

近年来, 我们与全国胃肠肿瘤外科同道一起, 搭建了中国胃肠肿瘤外科联盟平台<sup>[17]</sup>。收集了 2014 年起至今的数据进行年度汇报; 另, 借助于该平台经验, 进一步开展了我国首项多中心前瞻性胃结直肠癌术后腹腔感染性并发症现状研究 (PACAGE 研究)<sup>[19]</sup>。在建设多中心数据库的初步探索中发现, 质量化控制的基础涉及以下几个方面: (1) 研究目标导向制定数据计划。应充分考虑数据库建设的目标, 如胃肠肿瘤外科联盟平台建立初始目的是提供我国胃癌外科诊治基本现状的数据, 具有宏观视角, 希望能为医疗政策提供参考, 所以在数据收集计划上偏重于以每个中心为一个条目数据集, 进一步汇总分析汇报。(2) 充分考虑计划内数据参数的可及性。联盟在最初确定数据参数时, 充分考量可及性, 即参数定义的明确性和数据完整性, 首先尝试收集定义更加明确的数据, 例如年度胃癌手术量、手术术式、淋巴结转移数目及病理分期等信息。这些数据具有信息简明、定位明确、完整性高的特点, 保证了最终汇总数据的一致性和准确性。(3) 形成各中心间的数据共识。笔者团队在收集术后并发症信息时发现, 各中心记录差异较大。为了规范我国胃肠肿瘤外科并发症的诊断登记, 联盟和中国抗癌协会胃癌专业委员会于 2018 年发布了《中国胃肠肿瘤外科术后并发症诊断登记规范专家共识(2018 版)》<sup>[21]</sup>。PACAGE 研究也是在此数据共识基础上在 20 家中心开展进行的<sup>[22-23]</sup>。



多中心数据库具有数据体量大的特点,尽管在数据维度方面不如本中心的专病数据库,但是总结这些数据库的建设历程,加之本中心经验的推广和各中心的经验交流等,对于提升胃癌诊治水平是具有重要意义的<sup>[24-25]</sup>。笔者单中心数据库实质为基于真实世界下的胃癌患者登记研究,是建立在数据录入与临床实际工作深度融合的基础上的;得益于全流程录入模式下的数据完整性,数据库具有很大的灵活性和实用性。在建设多中心数据库时,更加庞大的数据体量带来多方面的挑战,需斟酌评估人力物力等的付出,符合自身的特点和要求。

展望未来笔者认为,建立更加全面、高维度的多中心胃癌数据库,努力方向包括:(1)各中心医疗及信息水平的同质化。这既为始因也是后果,数据录入的全面性、完整性和规范性的要求,意味着各中心在结构化数据建设方面需要同步跟进,同时落实并坚持多学科查房制度等,对手术科室和辅助科室都提出了更高的要求<sup>[9,26]</sup>。(2)实现数据共享和交换。为了促进多中心合作,需要建立稳定的公共数据平台,并持续维护运行;完善数据共享和交换的规范和流程,明确数据使用权限和数据所有权限,确保数据的共享和交换过程符合相关法规和伦理要求。(3)建立数据安全保障机制。多中心数据的建立必然对数据安全提出了更高的要求,因此,需要建立更加完善的数据安全保障机制,包括数据加密、访问控制、备份和恢复等措施,确保数据的安全性和可靠性<sup>[24,26]</sup>。(4)定期进行数据质量评估。跨中心实施数据质量评估,可以通过制定评估指标及开展内部审核等方式,发现和解决数据质量问题,不断提高数据质量水平。(5)加强多中心合作与交流。多中心合作与交流,是促进胃癌数据库多中心推广的关键。可以通过定期举办研讨会或座谈会等方式,增进各中心之间的了解,共同推动胃癌数据库的发展和应用<sup>[27-28]</sup>。

## 五、小结

胃癌数据库的建设需要重视质量化控制和管理,即建立专业的数据库管理团队、制定严格的数据质量标准和安全保障机制,实现数据共享和交换以加强数据分析和应用推广,同时获得持续更新和优化。随着信息化技术的不断发展和应用推广,将会进一步完善和提高胃癌数据库的质量和和应用价值,未来胃癌数据库的建设势必趋向智能化、自动化和高效化,更好地服务于医疗和科研工作。

**利益冲突** 所有作者均声明不存在利益冲突

## 参 考 文 献

- [1] Pascucci S, Langella F, Franzò M, et al. National spine surgery registries' characteristics and aims: globally accepted standards have yet to be met. Results of a scoping review and a complementary survey[J]. *J Orthop Traumatol*, 2023, 24(1): 49. DOI: 10.1186/s10195-023-00732-4.
- [2] Niederländer CS, Kriza C, Kolominsky-Rabas P. Quality criteria for medical device registries: best practice approaches for improving patient safety—a systematic review of international experiences[J]. *Expert Rev Med Devices*, 2017, 14(1): 49-64. DOI: 10.1080/17434440.2017.1268911.
- [3] Larsson S, Lawyer P, Garellick G, et al. Use of 13 disease registries in 5 countries demonstrates the potential to use outcome data to improve health care's value[J]. *Heal Aff*, 2012, 31(1): 220-227. DOI: 10.1377/hlthaff.2011.0762.
- [4] Merrill RM, Capocaccia R, Feuer EJ, et al. Cancer prevalence estimates based on tumour registry data in the Surveillance, Epidemiology, and End Results (SEER) Program[J]. *Int J Epidemiol*, 2000, 29(2): 197-207. DOI: 10.1093/ije/29.2.197.
- [5] Kurita N, Miyata H, Gotoh M, et al. Risk model for distal gastrectomy when treating gastric cancer on the basis of data from 33,917 Japanese patients collected using a nationwide web-based data entry system[J]. *Ann Surg*, 2015, 262(2): 295-303. DOI: 10.1097/sla.0000000000001127.
- [6] Beck N, Bommel AC van, Eddes EH, et al. The Dutch institute for clinical auditing: achieving codman's Dream on a nationwide basis[J]. *Ann Surg*, 2020, 271(4): 627-631. DOI: 10.1097/sla.0000000000003665.
- [7] Watanabe M, Miyata H, Gotoh M, et al. Total gastrectomy risk model: data from 20,011 Japanese patients in a nationwide internet-based database[J]. *Ann Surg*, 2014, 260(6): 1034-1039. DOI: 10.1097/sla.0000000000000781.
- [8] Voeten DM, Busweiler LAD, Werf LR van der, et al. Outcomes of esophagogastric cancer surgery during eight years of surgical auditing by the Dutch Upper Gastrointestinal Cancer Audit (DUCA) [J]. *Ann Surg*, 2021, 274(5): 866-873. DOI: 10.1097/sla.00000000000005116.
- [9] Mandavia R, Knight A, Phillips J, et al. What are the essential features of a successful surgical registry? A systematic review[J]. *BMJ Open*, 2017, 7(9): e017373. DOI: 10.1136/bmjopen-2017-017373.
- [10] 吴舟桥,石晋瑶,陕飞,等.从北京大学肿瘤医院11年胃癌术后并发症登记数据的学习曲线看并发症的规范化登记[J]. *中华胃肠外科杂志*, 2017, 20(2): 177-183. DOI: 10.3760/cma.j.issn.1671-0274.2017.02.013.
- [11] 张军豪,刘睿清,高源,等.厚植临床优势推动创新发展:中国结直肠癌手术病例登记数据库与国外数据库的比较[J]. *中华外科杂志*, 2023, 61(5): 362-367. DOI: 10.3760/cma.j.cn112139-20221113-00480.
- [12] 国家癌症中心,国家肿瘤质控中心胃癌质控专家委员会.中国胃癌规范诊疗质量控制指标(2022版)[J]. *中华肿瘤杂志*, 2022, 44(10): 997-1002. DOI: 10.3760/cma.j.cn112152-20220803-00532.
- [13] Li Z, Wang Y, Shan F, et al. ypTNM staging after neoadjuvant chemotherapy in the Chinese gastric cancer population: an evaluation on the prognostic value of the AJCC eighth edition cancer staging system[J]. *Gastric*

- Cancer, 2018,21(6):977-987. DOI: 10.1007/s10120-018-0830-1.
- [14] Xing J, Wang Y, Shan F, et al. Comparison of totally laparoscopic and laparoscopic assisted gastrectomy after neoadjuvant chemotherapy in locally advanced gastric cancer[J]. Eur J Surg Oncol, 2021,47(8):2023-2030. DOI: 10.1016/j.ejso.2021.02.002.
- [15] Li Z, Wu X, Gao X, et al. Development and validation of a novel staging system integrating the number and location of lymph nodes for gastric adenocarcinoma[J]. Br J Cancer, 2021, 124(5): 942-950. DOI: 10.1038/s41416-020-01190-z.
- [16] Tang X, Wu X, Gan X, et al. Using normalized carcinoembryonic antigen and carbohydrate antigen 19 to predict and monitor the efficacy of neoadjuvant chemotherapy in locally advanced gastric cancer[J]. Int J Mol Sci, 2023, 24(15): 12192. DOI: 10.3390/ijms241512192.
- [17] 中国胃肠肿瘤外科联盟. 中国胃肠肿瘤外科联盟数据报告(2014-2016)[J]. 中国实用外科杂志, 2018,38(1):90-93. DOI: 10.19538/j.cjps.issn1005-2208.2018.01.20.
- [18] 吴舟桥,李子禹,季加孚. 中国胃肠肿瘤外科联盟数据库建设及对胃肠外科诊疗的推动作用[J]. 中华胃肠外科杂志, 2020, 23(1): 26-32. DOI: 10.3760/cma.j.issn.1671-0274.2020.01.005
- [19] 王琦,吴舟桥,侯士阳,等. 胃结直肠癌术后腹腔感染性并发症的现状研究(PACAGE研究)[J]. 中华消化外科杂志, 2019, 18(3):229-234. DOI:10.3760/cma.j.issn.1673-9752.2019.03.007.
- [20] Wu Z, Yan S, Liu Z, et al. Postoperative abdominal complications of gastric and colorectal cancer surgeries in China: a multicentered prospective registry-based cohort study[J]. Sci Bull, 2022, 67(24): 2517-2521. DOI: 10.1016/j.scib.2022.12.008.
- [21] 中国胃肠肿瘤外科联盟,中国抗癌协会胃癌专业委员会. 中国胃肠肿瘤外科术后并发症诊断登记规范专家共识(2018版)[J]. 中国实用外科杂志, 2018,38(6):589-595. DOI: 10.19538/j.cjps.issn1005-2208.2018.06.01.
- [22] 王琦,吴舟桥,刘子宁,等. 胃结直肠癌术后腹腔感染性并发症多中心、前瞻性数据库的建立及数据质量控制[J]. 中华胃肠外科杂志,2023,26(2):154-159. DOI:10.3760/cma.j.cn441530-20221214-00524.
- [23] 吴舟桥,李子禹. 基于临床实践的胃肠道术后感染性并发症数据库的设计及质量控制[J]. 中华胃肠外科杂志, 2023, 26(9): 818-823. DOI: 10.3760/cma.j.cn441530-20230830-00075.
- [24] Lazem M, Sheikhtaheri A. Barriers and facilitators for the implementation of health condition and outcome registry systems: a systematic literature review[J]. J Am Med Inform Assoc, 2022,29(4):723-734. DOI: 10.1093/jamia/ocab293.
- [25] Bulusu VR, Fullarton J, Leahy M, et al. Rationale and design of a UK database for a rare cancer type: the GEM Registry for gastrointestinal stromal tumours[J]. Br J Cancer, 2013,109(6):1403-1407. DOI: 10.1038/bjc.2013.406.
- [26] McCulloch P, Cook JA, Altman DG, et al. IDEAL framework for surgical innovation 1: the idea and development stages [J]. BMJ, 2013,346:f3012. DOI: 10.1136/bmj.f3012.
- [27] 季加孚,何琦非,王晓云,等. 胃肠肿瘤医疗大数据的机遇与挑战[J]. 中华消化外科杂志, 2019,18(3):199-202. DOI:10.3760/cma.j.issn.1673-9752.2019.03.001.
- [28] 王敬普, de Groot Eline M, Visser Maurits R,等. 科学严谨的患者信息登记是高质量胃癌外科研究的基石—荷兰上消化道癌症审计数据库 11 年经验分享[J]. 中华胃肠外科杂志, 2023,26(2):148-153. DOI:10.3760/cma.j.cn441530-20221026-00432.

· 读者 · 作者 · 编者 ·

## 本刊“胃肠新视野”栏目征稿启事

“胃肠新视野”栏目为本刊特设的视频栏目。视频内容通过“e-Surgery 伊索云®/医路有伴®平台”为我刊设置的“专区”呈现,大家可通过手机进行观看。同时,视频内容的相关文字内容(包括手术方式的介绍、新技术的创新背景、病例介绍、手术相关并发症的处理要点等)会在相应的杂志上刊登并附二维码。诚挚欢迎各位同道积极投稿,具体投稿要求如下。

1. 内容:主要为手术视频,侧重展示胃肠新技术、新术式以及术中并发症的处理等;并附相应的文字介绍(1000字左右)。

2. 视频:视频时长不超过9 min,视频附带解说,大小<1 GB,格式:MPEG、MOV、MP4、AVI或WMV。请注明解剖部位;无背景音乐,避免“花俏”转场。已发行的具有著作权的视频资料DVD不宜。

3. 本栏目的视频及文字内容请以“胃肠新视野栏目投稿”为主题,发至我刊 Email: china\_gisj@vip.163.com。