

精准医学决策支持知识组织研究*

张秀梅

翟运开

(中国科学技术信息研究所 北京 100038)

(郑州大学管理工程学院 郑州 450001)

丁楠 吴晨溪 张兵涛

徐建武

(北京万方数据股份有限公司 北京 100038)

(中国科学技术信息研究所 北京 100038)

〔摘要〕 介绍国内外医学知识组织体系,从数据来源、概念抽取、语义关系构建几方面阐述基于精准医学临床决策支持系统的精准医学知识组织体系设计,分析其在门诊和住院场景的应用。

〔关键词〕 知识组织;临床场景;精准知识库;临床决策支持系统;精准医学临床决策支持系统

〔中图分类号〕 R-056 [文献标识码] A [DOI] 10.3969/j.issn.1673-6036.2020.05.004

Study on Knowledge Organization of Decision Support for Precision Medicine ZHANG Xiumei, Institute of Scientific and Technical Information of China, Beijing 100038, China; ZHAI Yunkai, School of Management Engineering, Zhengzhou University, Zhengzhou 450001, China; DING Nan, WU Chenxi, ZHANG Bingtao, Beijing Wanfang Data Co., Ltd, Beijing 100038, China; XU Jianwu, Institute of Scientific and Technical Information of China, Beijing 100038, China

〔Abstract〕 The paper introduces the domestic and foreign medical knowledge organization systems, expounds on the design of knowledge organization system for precision medicine based on the Clinical Decision Support System for Precision Medicine (PM-CDSS) from aspects of data sources, concept extraction, and semantic relationship building, and analyzes its application in outpatient and inpatient scenarios.

〔Keywords〕 knowledge organization; clinical scenario; precise knowledge base; Clinic Decision Support System (CDSS); Clinical Decision Support System for Precision Medicine (PM-CDSS)

〔收稿日期〕 2020-01-07

〔作者简介〕 张秀梅,博士,发表论文 87 篇;通讯作者:徐建武,发表论文 4 篇。

〔基金项目〕 国家重点研发计划“精准医学研究”重点专项“基于远程/移动医疗网络的精准医疗综合服务示范体系建设与推广”(项目编号:2017YFC0909900)。

1 引言

临床决策支持系统(Clinic Decision Support System, CDSS)是医院信息系统重要组成部分,我国智慧医院及电子病历等相关政策明确对其提出要求^[1]。随着“互联网+”、移动互联、大数据、人工智能、机器学习、区块链等新技术不断发展,医疗及医院信息化领域面临着全新挑战和机遇^[2-3]。CDSS的历史可以追溯到1976年斯坦福大学研究开发的MYCIN,虽然发展历史久远,但现实应用中仍

存在一定问题，如知识与规则流程融合、临床从循证医学走向精准医学、多学科联合决策支持、知识内容严谨性和质控等^[4-8]。本文从精准医学临床决策支持系统（Clinical Decision Support System for Precision Medicine, PM - CDSS）出发，研究构建支撑 PM - CDSS 的精准医学知识组织体系，满足 CDSS 在多角色工作场景中的信息需求。

2 医学知识组织体系

2.1 基本概念

医学知识组织体系或知识组织系统（Medical Knowledge Organization Systems, MedKOS）涵盖医学领域内各种词表、概念体系、分类体系等，具有组织和描述医学知识、概念语义消歧和标准化、概念间层级和非层级结构语义关联展示等功能，属于图书情报学科领域中的一个分支，包括对知识单元或

概念内在结构进行的加工、整理、序化、表示等一系列操作。

2.2 相关研究

国外关于医学知识组织研究较为成熟，其中在临床及学术领域应用较为广泛的有：医学主题词表（MeSH）、一体化医学语言系统（UMLS）、系统化人兽医学学术语表临床术语（SNOMED CT）、观测指标标识符逻辑命名与编码系统（LONIC）、国际疾病分类（ICD）等；精准医学相关知识组织有：人类孟德尔遗传在线（OMIM）、罕见病（ORPHA）、中文人类表型（CHPO）、药物基因组学（PharmGKB_drugs）、基因（HGNC）、变异（DisGeNET）、药物编码查询系统（ATC）等。国内医学知识组织体系有：中国图书馆图书分类法、医学主题词表中文版、UMLS 中文版、中医药主题词表等^[9-17]，见表 1。

表 1 国内外主要医学知识组织体系

名称	介绍
医学主题词表（MeSH）	医学领域标准化叙词表，由美国国立医学图书馆编制，也是动态可扩充的叙词表，汇集 20 000 多个主题词
一体化医学语言系统（UMLS）	由美国国立医学图书馆编制，包括超级叙词表、专家词典和语义网
系统化人兽医学学术语表临床术语（SNOMED CT）	全面的临床医学相关术语集合，包含词库、受控词表、分类系统、代码系统、语义网络，有助于临床信息系统检索和病历信息的语义互操作
观测指标标识符逻辑命名与编码系统（LONIC）	LONIC 术语涉及用于临床医疗护理、结局管理和临床研究等目的的各种临床观测指标，如血红蛋白、血清钾、各种生命体征等
国际医学分类族 WHO - FIC	(1) ICD（International Classification of Diseases），世界卫生组织制定的国际国际疾病/健康分类，是依据疾病某些特征，按照规则将疾病分门别类，并用编码方法来表示的系统。(2) ICF（International Classification of Functioning, Disability and Health），世界卫生组织制定的国际功能 - 障碍 - 健康分类，核心概念用于健康领域的良好状态，健康状况的限定语表现。(3) ICHI（International Classification of Health Interventions），世界卫生组织制定的国际健康干预分类，提供一种用于统计目的并且对健康干预的分布和演变情况加以分析和报告的通用工具
精准医学相关词表	(1) 基因本体（Gene Ontology, GO），在生物信息学领域中广泛使用的本体，涵盖生物学 3 个方面：细胞组分、分子功能、生物过程。(2) 在线人类孟德尔遗传数据库（OMIM），涵盖关于人类遗传病和基因座位等相关信息和文献的中心级数据库。(3) 罕见病（Orphanet），关于罕见病诊断、治疗的专业网站，包含孤儿药和疾病信息。(4) 人类表型术语集（HPO/CHPO），旨在提供人类疾病中用于描述表型异常的标准词汇，中文版为 CHPO。(5) 遗传药理学与药物基因组学数据库（PharmGKB），收集最完整的与药物基因组相关的基因型和表型信息，将基因和药物关系分为两个大的范畴。(6) 人类基因组数据库（HGNC），每个人的基因在 HGNC 数据库中会有一个数字标识的唯一 ID，也会有一个唯一的 symbol。(7) 疾病相关的基因与突变位点数据库（DisGeNET），专门收录人类疾病相关的基因与突变位点信息的数据库。(8) 药物编码查询系统（ATC），即解剖学治疗学及化学分类系统，是世界卫生组织对药品的官方分类系统

续表 1

中国图书馆分类法	简称为中图法，是我国通用的图书分类工具
中文医学主题词表 (CMeSH)	由中国医学科学院医学信息研究所翻译、出版和维护，包括所有主题词、印刷款目词及副主题词的中译名
中文一体化医学语言系统 (CUMLS)	由中国医学科学院医学信息研究所开发研制，由词表构建与维护平台、医学词表、语义网、词表服务平台组成
中医药学主题词表	由中国中医科学院中医药信息研究所研制，是我国第1部中医药专业词表，在中医药学领域应用广泛

3 精准医学知识组织体系设计

3.1 PM - CDSS 系统

3.1.1 概述 基于场景的精准医学临床决策支持系统 (Scenario based Clinical Decision Support System for Precision Medicine, PM - CDSS) 为精准诊断和治疗提供决策支持，涵盖场景、对象、时间、决策信息的精准等。构建覆盖全学科、多角色 (支持医生、护士、药剂师、麻醉师、检查/验技师等) 主要工作场景，满足诊前、诊中及诊后全流程需求的精准医学决策支持系统，为临床工作提供及时、有效的精准参考建议，进而降低医疗风险，提高总体诊疗水平^[18]。本文基于上述对于精准医学的理解和认知，重点分析 PM - CDSS 的场景需求及知识组织设计。

3.1.2 特点 对知识数据颗粒度要求更细、内容时效和准确度要求更高，需要抽提各类医学概念类；在推理过程中以医学概念类之间的语义关系为基础，建立语义规则库，为输出最佳决策信息提供支撑。能够适应不同终端载体，深入场景与各医院信息系统深度整合，动态、精准地提供决策支持。传统 CDSS 与精准 CDSS 区别，见表 2。

表 2 传统 CDSS 与精准 CDSS 区别

特点	传统 CDSS	精准 CDSS
数据来源	循证证据	循证精准知识
颗粒度	内容片段	碎片化知识
数据关系	无或弱关联	强关联规则
数据展示	静态阅读	动态提醒
数据调用	主动查阅	场景嵌入

3.2 数据来源 (表 3)

3.2.1 概述 一方面遵从循证医学要求，知识数据必须有权威可靠出处，才能够为决策支持所用；另一方面数据广度和深度要求更高，数据广度方面增加精准医学相关术语和知识，深度方面需要对数据进行二次碎片化加工，为临床决策提供支撑。

表 3 精准医学主要概念及数据来源

知识库类型	内容	来源
疾病知识库/ 诊断知识库	临床问题	症状、体征、ICD10/11 诊断术语
	诊断推理规则	临床路径、相关指南
治疗知识库	检查方案	临床路径、临床指南、用药说明
	治疗方案	
药物知识库	药品说明书	药品生产厂家、药典及其他途径
	合理用药	药物相互作用相关书籍及文献
	抗菌药物	国家和地方抗菌药物指南
检查检验及其他	检查检验知识库 正常值及异常值	临床检查检验操作指南等相关工具书
临床指南文献库	公开权威指南	卫健委网站及期刊杂志公开发表的正式指南
临床路径库	公开临床路径	卫健委网站
术语知识库	临床术语词典	SnomedCT, ICD10/11, ICF, ICHI, LOINC, UMLS, MeSH, CMeSH, OMIM, ORP-HA, CHPO, PharmGKB, HGNC, DisGeNET 及非标准术语和值集等
外部数据	临床病例 临床经验	基于临床病例数据挖掘的经验以及临床医生形成共识的经验

南、药物说明书等知识中关于概念与概念之间的描述，梳理标注 22 种精准医学知识语义关系，见表

4。由于医学知识组织体系是不断发展的体系，医学概念及语义关系也会不断更新调整。

表 4 精准医学知识语义关系

关系类型	示例	关系类型	示例
ID	ICD - 10: C50.900	分期	II A 期 分期 乳腺癌
类	乳腺癌 是 肿瘤科 乳腺癌 是 乳腺疾病	并发	放射性肺损伤 并发于 乳腺癌
名称	乳腺癌	参考值	癌抗原 15 - 3 测定 参考值 25 ~ 28U/ml
别名	乳腺癌 别名 breast cancer	危急值	血清肌酐 危急值 > 880μmol/L
致病	Her2 阳性 致病 乳腺癌	标本	癌抗原 15 - 3 测定 标本 血清
多发人群	乳腺癌 多发人群 女性	部位	乳腺 MRI 检查 部位 胸部
发生于	乳腺癌 发生于 胸部	服用方式	枸橼酸他莫昔芬 服用方式 口服
是症状	乳腺肿块 是症状 乳腺癌	服用间隔	枸橼酸他莫昔芬 服用间隔 1 日 1 次
诊断	乳腺细胞学检查 诊断 乳腺癌 乳腺 X 检查 诊断 乳腺癌	不良反应	枸橼酸他莫昔芬 不良反应 月经周期紊乱
鉴别	乳腺良性增生病 鉴别 乳腺癌	禁忌症	枸橼酸他莫昔芬 禁忌症 妊娠期妇女
适应症	枸橼酸他莫昔芬 治疗 乳腺癌	配伍禁忌	阿莫西林 配伍禁忌 糖皮质激素

4 知识组织应用

4.1 应用范畴

知识组织体系应用范畴较广，在信息检索优化、知识库构建、临床决策支持、命名实体识别、数据挖掘等方面都可以发挥重要作用。本研究仅在场景化的精准医学决策支持方面进行讨论。临床诊疗活动是一个连续的过程，单场景应用包括临床辅助诊断、临床诊疗建议、合理性检查、数值提醒、

数据挖掘等方面，该方式无法完全发挥医学知识组织在临床实践中的作用，如仅有诊断结论而没有后续的治疗建议，或仅有标准化的治疗方案但缺少相应的个性化用药指导等。鉴于此，提出基于临床业务流程的决策场景，在临床业务流程模型中每个业务节点都与知识库进行关联，在不同节点进行相应提醒，以门诊和住院两个场景为例进行分析。

4.2 门诊场景 (图 2)

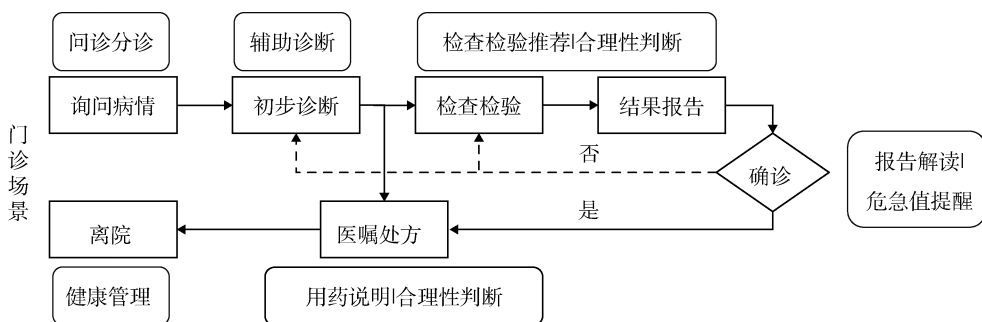


图 2 门诊场景精准医学决策支持

询问病情时能够提供问诊及分诊信息；初步诊断时提供辅助诊断建议；下达检查检验医嘱时推荐

合理的检查检验项目并对其合理性进行判断；阅读结果报告时有相关解读及危急值提醒标注；下达用

药处方时提供药物说明并合理性判断；患者离开医院后能够进行健康管理。

4.3 住院场景 (图 3)

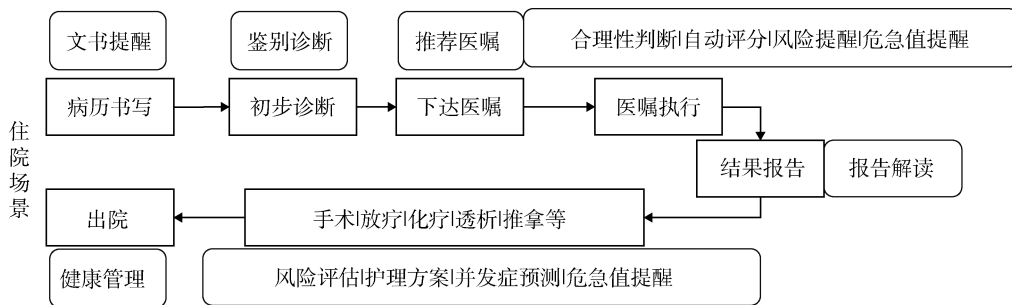


图 3 住院场景精准医学决策支持

撰写电子病历时时有相关提醒及校正；初步诊断时进行鉴别诊断提醒；下达医嘱时有推荐医嘱提醒，并对医嘱内容进行合理性判断及风险提醒；阅读结果报告时有相关解读及危急值提醒标注；制定并执行治疗处理方案时有风险评估、护理方案及相关风险预测提醒；患者出院后能够进行健康管理。

5 结语

基于场景的精准医学知识体系研发以临床业务流程为中心，满足各种场景临床决策支持需求。本文通过调研国内外医学知识组织体系及临床主要业务场景，对 PM - CDSS 特点、医学知识组织体系、知识数据来源、概念抽取、语义关系梳理及场景应用进行系统性研究，由于精准医学知识组织体系是一个复杂且庞大的工程，现有知识组织体系还存在一定差距，下一步将在 PM - CDSS 实现、知识组织工程化以及知识组织体系完善等方面开展研究。

参考文献

- 1 国家卫生健康委办公厅. 关于印发电子病历系统应用水平分级评价管理办法(试行)及评价标准(试行)的通知 [EB/OL]. [2018 - 12 - 07]. <http://www.nh-c.gov.cn/zyygj/s7659/201812/3cae6834a65d48e9bfd783f3c7d54745.shtml>.
- 2 李国磊, 陈先来, 夏冬. 面向临床决策的电子病历系统概述 [J]. 中国数字医学, 2014, 9 (12): 30 - 32.
- 3 王昱. 基于电子病历的临床决策支持研究 [D]. 杭州: 浙江大学, 2016.

- 4 Sim I, Gorman P, Greenes RA, et al. Clinical Decision Support Systems for the Practice of Evidence - based Medicine [J]. J Am Med Inform Assoc, 2001, 8 (6): 527 - 534.
- 5 周娜, 李爱芹, 刘广伟, 等. 沃森肿瘤人工智能系统在临床中的应用 [J]. 中国数字医学, 2018, 13 (10): 23 - 25.
- 6 洪建, 张洁, 陶敏, 等. 新型临床决策支持系统在智慧医院建设中的应用研究 [J]. 中国数字医学, 2019, 14 (9): 22 - 24.
- 7 王宇, 王心慰, 刘爽, 等. 精准医学的临床部署: 顶层架构设计及关键信息技术 [J]. 转化医学杂志, 2017, 6 (6): 321 - 324.
- 8 杨春华, 王天津, 黄思敏, 等. 支持精准医疗的国外临床决策支持系统 [J]. 中华医学图书情报杂志, 2016, 25 (2): 14 - 19.
- 9 杜也力. 我国关于知识组织的研究述评 [J]. 中国图书馆学报, 2002, 28 (5): 66 - 69.
- 10 蒋永福, 李景正. 论知识组织方法 [J]. 中国图书馆学报, 2001, 27 (1): 3 - 7.
- 11 李洁, 丁颖. 语义网、语义网络和语义网络 [J]. 计算机与现代化, 2007 (7): 38 - 41.
- 12 牟冬梅, 崔艳玲. MeSH、本体论在医学知识组织中的作用 [J]. 情报杂志, 2005, 24 (7): 120 - 122.
- 13 Cimino J J. Auditing the Unified Medical Language System with Semantic Methods [J]. Journal of the American Medical Informatics Association, 1998, 5 (1): 41.
- 14 Stearns M Q, Price C, Spackman K A, et al. SNOMED Clinical Terms: overview of the development process and project status [J]. Proceedings of AMIA Annual Symposium, 2001, 8 (1): 662.

(下转第 48 页)

续表 1

社会合作机制	健康医疗大数据产、学、研、用一体化发展政策, 收费和运营机制, 安全保障机制	社会合作应用数据管理办法、研究成果知识产权办法	社会合作应用数据具体制度、工作流程
研发成果转化机制	成果转化奖励、资金分配、价值转换	研发成果转化制度和 workflows	-

6 结语

区域健康医疗大数据中心是卫生健康信息化发展的方向和必然趋势, 具有广阔的前景。目前处于初始发展阶段, 面临许多困难和挑战, 构建适应区域健康医疗大数据中心运行的体制机制是基础和关键, 对体制机制进行全方位分析和整体构建, 将有助于促进健康医疗大数据健康快速发展。

参考文献

1 国务院办公厅. 国务院办公厅关于促进和规范健康医疗大数据应用发展的指导意见 (国办发 [2016] 47 号) [EB/OL]. [2019 - 06 - 24]. http://www.gov.cn/zhengce/content/2016-06/24/content_5085091.htm.

2 国家卫生健康委员会. 关于印发《国家健康医疗大数据标准、安全和服务管理办法 (试行)》的通知 (国卫规划发 [2018] 23 号) [EB/OL]. [2018 - 09 - 15]. http://www.cac.gov.cn/2018-09/15/c_1123432498.htm.

3 福州市人民政府办公厅. 关于印发《福州市健康医疗大数据资源管理实施细则》的通知 (榕政办 [2017] 299 号) [EB/OL]. [2019 - 11 - 01]. http://www.fuzhou.gov.cn/zfxxgkz/szfbmjxsqxgk/szfbmxxgk/fzsrnzfbgt/zfxxgkml/xzfggzhgfwj_2570/201711/t20171101_1810034.htm.

4 龙虎, 邱航, 吴沧浪, 等. 四川省健康医疗大数据中心构建探讨 [J]. 中国卫生信息管理杂志, 2017, 14 (1): 15 - 18, 23.

5 胡建平. 医院数据治理框架、技术与实现 [M]. 北京: 人民卫生出版社, 2019.

(上接第 29 页)

15 Gene Ontology Consortium. An Introduction to the Gene Ontology. The Gene Ontology [EB/OL]. [2019 - 12 - 01]. <http://www.geneontology.org/GO.doc.shtm>.

16 张晓梅, 李丹亚, 胡铁军. 一体化医学语言系统与本体论研究 [J]. 医学信息学杂志, 2006, 27 (2): 89 - 92.

17 McDonald C J, Huff S M, Suico J G, et al. LOINC, a Universal Standard for Identifying Laboratory Observations: a 5 - year update [J]. Clinical Chemistry, 2003, 49 (4): 624.

18 姜军, 唐鹏. 乳腺癌精准诊疗临床决策支持系统的关键技术 [J]. 中华医学杂志, 2015, 95 (31): 2503 - 2505.

19 张秀梅, 徐建武, 程煜华, 等. 基于知识库的临床决策支持系统构建 [J]. 中华医院管理杂志, 2014, 30 (6): 472 - 475.

20 李艳, 吴梦佳, 张士靖. 语义互操作标准在临床决策支

持系统中的应用 [J]. 医学信息学杂志, 2017, 38 (10): 57 - 61, 83.

21 胡兆芹, 张士靖. 领域本体的构建实践 [J]. 医学信息学杂志, 2008, 29 (1): 12 - 16.

22 张庆, 吕少妮, 轩扬. 本体在生物医学领域中应用研究热点分析 [J]. 医学信息学杂志, 2019, 40 (1): 63 - 67.

23 张娟. 基于本体的单元信息知识组织体系构建 [J]. 图书馆工作与研究, 2017 (12): 62 - 65.

24 苟梦野. 基于主题词的乳腺增生病知识组织体系构建研究 [D]. 重庆: 重庆医科大学, 2018.

25 曹倩, 赵一鸣. 知识图谱的技术实现流程及相关应用 [J]. 情报理论与实践, 2015, 38 (12): 127 - 132.

26 黄恒琪, 于娟, 廖晓, 等. 知识图谱研究综述 [J]. 计算机系统应用, 2019, 28 (6): 1 - 12.