

# 基于 PADIS 指南的知识图谱构建研究\*

张山<sup>1</sup> 陆观<sup>2</sup> 刘璐<sup>1</sup> 王慧莹<sup>1</sup> 吴瑛<sup>1</sup>

(<sup>1</sup>首都医科大学护理学院 北京 100069 <sup>2</sup>上海卓越睿新数码科技股份有限公司 上海 200233)

**[摘要]** **目的/意义** 构建基于临床实践谵妄指南的知识图谱,旨在为提升医护人员对指南的依从性提供方法学参考。**方法/过程** 基于知识图谱构建方法,以指南为主要数据来源,基于指南知识架构和专家意见构建本体;采用三元组数据模型进行知识表示;采用“BIOES”标注体系进行人工标注,包含实体、关系和属性等;利用图形数据库 Neo4j 实现知识图谱的可视化。**结果/结论** 所构建的知识图谱包含 112 种实体类型和 16 种实体间关系,形成 124 个三元组,用于描述实体之间的关系、实体的属性及其对应的属性值。未来研究将进一步探索深度学习等模型在知识抽取中的应用,并探究指南自更新机制,以适应当前医学领域指南快速更新的现状。

**[关键词]** 护理管理;临床实践指南;依从性;临床护理决策支持系统;知识图谱

**[中图分类号]** R-058 **[文献标识码]** A **[DOI]** 10.3969/j.issn.1673-6036.2024.12.009

## Study on the Construction of a Knowledge Graph Based on PADIS Guidelines

ZHANG Shan<sup>1</sup>, LU Guan<sup>2</sup>, LIU Lu<sup>1</sup>, WANG Huiying<sup>1</sup>, WU Ying<sup>1</sup>

<sup>1</sup>School of Nursing, Capital Medical University, Beijing 100069, China; <sup>2</sup>Shanghai Zhuoyueruixin Digital Technology Company, Shanghai 200233, China

**[Abstract]** **Purpose/Significance** To construct a knowledge graph based on clinical practice guidelines for pain, agitation/sedation, delirium, immobility, and sleep disruption in adult patients in the ICU (PADIS guidelines), and to provide methodological references for improving the adherence of medical staffs to the guidelines. **Method/Process** Based on the construction method of the knowledge graph, the guidelines are used as the main data source, and the ontology is constructed based on the guideline framework and expert opinions. The triplet data model is used to represent knowledge. The “BIOES” labeling system, including entities, relationships, and attributes, is adopted for manual labeling. The visualization of the knowledge graph is realized by using the Neo4j graph database. **Result/Conclusion** The knowledge graph includes 112 entity types and 16 relationships among entities, and 124 triples are formed to describe the relationships among entities, the attributes of entities and their corresponding attribute values. In the future, it is necessary to explore the application of deep learning and other models in knowledge extraction, and explore the self-updating mechanism of guidelines to adapt to the current situation of rapid updating of guidelines in the medical field.

**[Keywords]** nursing management; clinical practice guideline; adherence; clinical nursing decision-making support system; knowledge graph

**[修回日期]** 2024-09-09

**[作者简介]** 张山,博士,副教授,发表论文 60 余篇;通信作者:吴瑛,教授,博士生导师。

**[基金项目]** 国家自然科学基金项目(项目编号:72304196)。

## 1 引言

近年来医学知识快速更新,临床指南的修订频率也随之加快,这给医护人员精准记忆和应用个性化的处理建议带来挑战,导致全球范围内临床实践指南的依从性普遍较差<sup>[1]</sup>。尤其在重症监护病房(intensive care unit, ICU)环境中,医护人员需要在紧急情况下迅速作出最佳临床实践决策,不可能每次都查阅相关指南后再实施<sup>[2]</sup>。为解决这一问题,本研究团队前期通过优化《ICU成年患者疼痛、躁动/镇静、谵妄、活动受限和睡眠紊乱预防及处理指南》(以下简称PADIS指南)的任务呈现方式,研发了ICU谵妄智能化护理系统<sup>[3]</sup>。该系统通过人工提取PADIS指南的措施,构建临床决策支持系统措施库,并采用“IF...THEN...”规则构建知识库。然而,由于指南更新频繁,基于产生式规则的知识库缺乏自动适应指南变化的能力,因此难以持续发挥有效作用<sup>[4]</sup>。而基于知识图谱表示的知识库能够融合不同图谱,具有知识更新灵活、处理效率高等优势<sup>[5]</sup>。目前,知识图谱已在医学领域得到广泛应用,如乳腺癌<sup>[6]</sup>、老年疾病康复<sup>[7]</sup>、中医养生<sup>[8]</sup>等。鉴于此,本研究在前期研究基础上,采用知识图谱的构建方法,以PADIS指南为主要数据源,构建基于循证指南的知识图谱,以替换ICU谵妄智能化护理系统中现有的“IF-THEN”知识库,旨在为医护人员提供一个更为先进的应用工具,从而提升其对指南的依从性。

## 2 资料与方法

### 2.1 基于PADIS指南的知识图谱构建框架

通过文献回顾,结合PADIS指南的特点,明确知识图谱的构建框架,该框架主要包含知识体系梳理、知识建模、知识抽取、知识表示和知识存储等环节。

### 2.2 知识体系梳理

基于PADIS指南的知识图谱构建过程中,知识

体系梳理是确保信息准确性和系统性的关键步骤,主要涉及对医学概念、实体及其相互关系的分析,以构建一个结构化的框架。本研究中为了确保基于PADIS指南的知识图谱的准确性、一致性和实用性,研究团队基于指南内容和知识结构,结合专家意见,梳理指南的结构。

### 2.3 知识建模

通过对相关领域知识的深入分析和获取,形成了一系列知识集合。在此基础上,创建一种计算机可解释的模型,以实现知识的保存、改进、共享、替代、聚合和重新应用。由于基于PADIS指南的知识图谱主要为医学领域相关人员提供决策支持,对知识检索效率和准确率的要求较高,因此,基于一体化医学语言系统(unified medical language system, UMLS)、国际医学规范术语全集(systematized nomenclature of medicine, SNOMED)以及专家建议,构建知识图谱的本体模型。

### 2.4 知识抽取

由于知识抽取内容为PADIS指南,属于专业领域知识,且数据量相对较小,采用人工抽取,由具有专业知识和经验的人员进行指南中实体、关系及属性的抽取。这些专业人员能够准确理解特定领域的术语和上下文复杂的语义关系,确保数据的质量和准确性。语料来源方面,以2018年发布的PADIS指南为测试集<sup>[9]</sup>,共计4万余字,主要包含ICU谵妄的危险因素、风险预测、评估方法、药物治疗以及非药物多因素预防和治疗等。语料标注方面,标注体系为“BIOES”,B(begin)代表实体的开头,I(inside)代表实体的中间,O(outside)代表命名实体外部,E(end)代表实体的尾部,S(single)代表该词是一个实体。

### 2.5 知识表示

知识表示是将现实世界的信息转化为计算机能够理解并利用的形式。采用三元组(实体-关系-实体)数据模型进行知识表示,从PADIS指

南中提取医学实体及其关系，并将这些知识点组合成知识网络。此外，使用资源描述框架（resource description framework, RDF）作为描述语言进行知识表示。

### 2.6 知识存储

采用 CSV 文件格式存储数据，便于后续数据实时插入和处理。

## 3 结果

### 3.1 基于 PADIS 指南的知识图谱构建框架

基于 PADIS 指南特点，构建符合临床实际需要和专业认知的知识图谱，构建框架，见图 1。

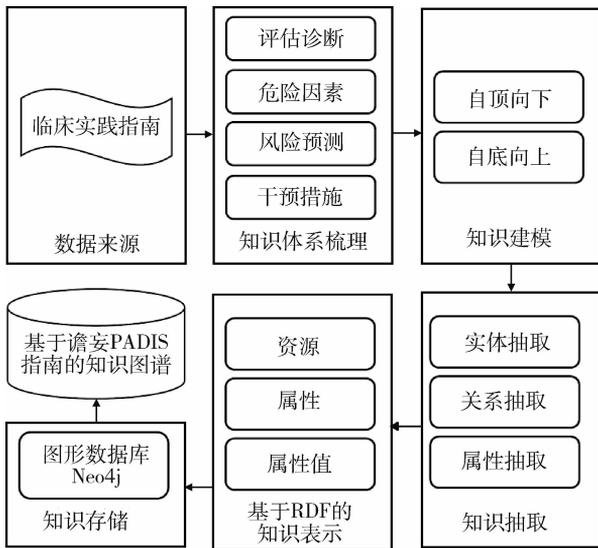


图 1 基于临床实践指南的知识图谱构建框架及流程

### 3.2 基于 PADIS 指南的知识体系

PADIS 指南主要包含 4 个主题：评估诊断、危险因素、风险预测、干预措施，见图 2。该指南在临床实际工作中的应用流程主要是收集患者信息进行疾病评估诊断；收集危险因素进行疾病预测以及提供相应的预防措施；为已发谵妄患者提供相应治疗措施。

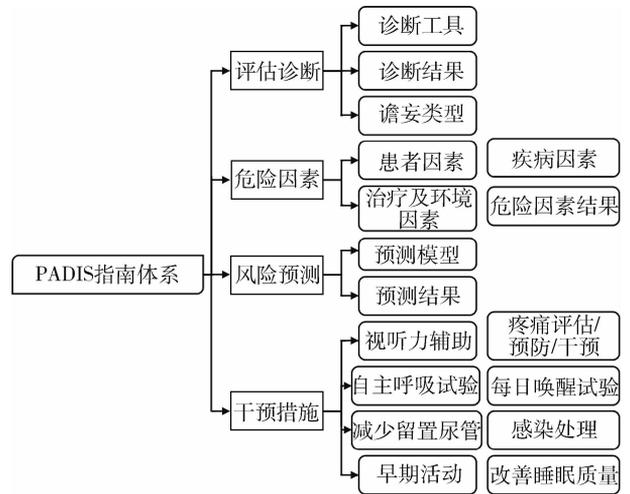


图 2 PADIS 指南知识体系

### 3.3 基于 PADIS 指南的知识建模

应用 Protégé 软件完成知识建模，包含临床表现、危险因素、评估诊断、分型、预防措施和治疗，见图 3。

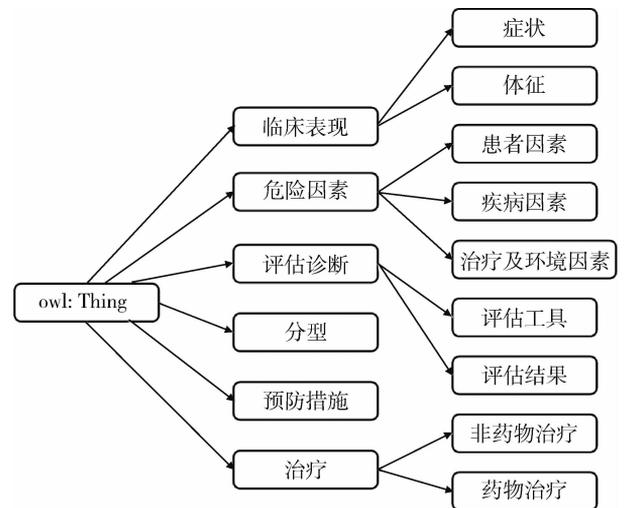


图 3 PAIDS 指南知识建模（部分示例）

### 3.4 基于 PADIS 指南的知识抽取

人工标注完成后，该知识图谱以临床表现、危险因素、评估诊断、分型、预防措施和治疗等为一类实体类型，包含 112 种实体类型及 16 种实体间的关系。知识图谱包含了 124 个三元组，用于描述实

体之间的关系、实体的属性及其对应的属性值。以“危险因素”知识框架为例，部分实体类型框架，见表 1。

表 1 PADIS 指南知识图谱“危险因素”实体类型框架

一级	二级	三级
危险因素	患者因素	视力 (减退)
		听力 (减退)
		疼痛
	疾病因素	疼痛
		感染
		镇静剂 (使用)
	治疗与环境因素	活动水平
		家属陪伴 (无)
		睡眠紊乱
		留置尿管
		机械通气
		机械通气

### 3.5 基于 PADIS 指南的三元组知识表示 (表 2)

表 2 PADIS 指南中三元组知识表示示例

头实体	关系	尾实体
ICU 意识模糊评估法	包含	意识水平
ICU 意识模糊评估法	包含	注意力
活动	包含	被动运动
活动	包含	主动运动
主动运动	包含	站立
主动运动	包含	行走
疾病因素	并列	患者因素
疾病因素	并列	治疗因素
预防措施	涉及	活动
药物治疗	涉及	右美托咪定

### 3.6 基于 PADIS 指南的知识存储

利用图形数据库 Neo4j 进行 PADIS 指南知识存储并实现可视化展示，见图 4，在 Neo4j Browser 的图形界面查看和浏览 PADIS 指南知识图谱的节点和关系。

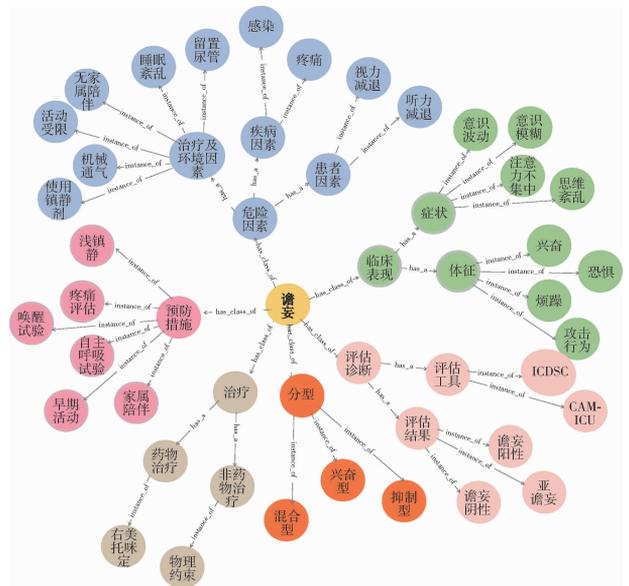


图 4 基于 PADIS 指南构建的知识图谱 (模式图)

## 4 讨论

### 4.1 基于 PADIS 指南的知识图谱构建具有良好的科学性

本研究基于指南知识架构和专家意见构建了 PADIS 指南的本体，包含：临床表现、危险因素、评估诊断、分型、预防措施和治疗等<sup>[10]</sup>。在医学研究中，明确该领域的概念、概念间的关系，有利于消除歧义以及形成丰富的语义网络，进而提高知识表示的准确性和一致性。例如，王永博等<sup>[11]</sup>参照 OMAHA Schema 对非肌层浸润性膀胱癌指南进行了概念层提取。目前，医学领域本体检索的途径有 UMLS、SNOMED、EMBL - EBI Ontology、PubMed 和 Google Scholar 等<sup>[12]</sup>。本研究采用人工知识抽取的方法对 PADIS 指南中相关的实体、关系和属性进行抽取。人工知识抽取在医学专业领域中具有显著优势。一是专业人员在知识抽取过程中能够直接进行质量控制，确保抽取内容的质量和准确性，这对于构建实用且可靠的医学知识图谱至关重要<sup>[13]</sup>。二是考虑到本研究的数据源规模相对较小，数据量不足以支持深度学习模型训练，为避免模型过度拟合导致表现不佳，选择人工知识抽取方式<sup>[14]</sup>。此外，王丽敏等<sup>[15]</sup>也采用人工标注的方式进行知识抽取，

作为构建疾病知识图谱的基础,进一步验证了人工知识抽取的有效性和可行性。

#### 4.2 模型抽取方法在基于 PADIS 指南的知识图谱构建中的展望

人工知识抽取具有一定的局限性,例如,要耗费大量人力资源来进行实体、关系及属性的抽取。与自动化知识抽取方式相比,人工知识抽取的速度较慢,处理的数据量也有限。随着知识图谱规模的扩大,人工知识抽取的方法难以满足大规模数据处理的需求。为克服上述局限性,未来研究将考虑扩大数据来源,并引入模型抽取方法,如基于深度学习的实体识别和关系抽取技术,以提高知识抽取效率<sup>[16]</sup>。BERT - BiLSTM - CRF 模型结合了双向编码器表征 (bidirectional encoder representations from transformers, BERT)、双向长短期记忆 (bidirectional long short - term memory, BiLSTM) 和条件随机场 (conditional random field, CRF) 的优势,在自然语言处理任务中表现出色,特别是在处理上下文信息和标签相互作用的序列标注任务中<sup>[17]</sup>。在医疗领域, BERT - BiLSTM - CRF 模型有助于捕捉医学术语、疾病名称、药物名称等关键信息,对处理具有复杂上下文关系的医疗文本尤为重要,有助于提高知识图谱构建的准确性<sup>[18]</sup>。例如,刘彬等<sup>[19]</sup>应用 BERT - BiLSTM - CRF 模型在中医医案命名实体识别中的  $F1$  值 (84.55%) 高于 BERT 模型 (81.10%) 和 BERT - CRF 模型 (83.28%)。Fang A 等<sup>[20]</sup>发现, BERT - BiLSTM - CRF 模型在中文电子病历中提取垂体腺瘤临床命名实体的综合性能最优,严格  $F1$  值最高为 91.27%, 宽松  $F1$  值最高为 95.57%。然而,医疗领域的专业术语和表达方式可能与通用语料存在差异,可能影响模型在基于指南的知识图谱构建中的性能<sup>[16]</sup>。因此,未来研究需进一步完善电子病历系统,实现数据的标准化、结构化和互通性,以提高 BERT - BiLSTM - CRF 模型学习的数据质量,进而提升辅助决策支持的准确性。此外,为满足临床工作的需求,未来应不断优化算法模型,提高大数据模型的数据处理能力,以提供更加准确的决策支持。

#### 4.3 基于 PADIS 指南的知识图谱应用意义

本研究采用基于知识图谱表示的知识库替代前期研究中基于“IF - THEN”规则的知识库。知识图谱通过编码和表示 PADIS 指南中的知识,实现了知识的有效管理<sup>[21]</sup>。此外,知识图谱能借助算法 (如深度学习、机器学习) 自动识别和分析新的临床实践指南,自动更新内容,以适应不断更新的临床实践指南<sup>[22]</sup>。尹梓名等<sup>[23]</sup>基于非小细胞肺癌与冠心病临床指南分别构建疾病的知识图谱,实证研究发现其能减轻医护人员的记忆负荷,对提升临床决策具有积极作用。李伟等<sup>[24]</sup>采用 Neo4j 图数据库构建知识图谱,作为医疗智能问答系统的知识库,临床试验证实该系统能准确匹配用户问题并提供准确答案。未来,本研究团队将进一步验证基于谵妄 PADIS 指南的知识图谱在提升医护人员指南依从性和改善患者临床结局中的效果。

### 5 结语

PADIS 指南在规范医护人员对疾病的预防、诊疗和护理过程中发挥重要的作用。本研究以医学科学研究的证据增长和知识更新速度导致信息过载为研究问题,基于知识图谱构建的方法,通过人工抽取知识进行图谱构建,并利用 Neo4j 图数据库实现知识图谱的可视化,基于 PADIS 指南的知识图谱能够为临床医护人员提供实用的建议和指导。但由于指南数据量较少,不适合采用深度学习等自动化模型进行知识抽取,未来研究需进一步扩大数据量,并探索指南自更新的机制,以适应当前医学领域指南快速更新的现状。

**作者贡献:** 张山负责研究设计、资料收集与分析、论文撰写;陆观负责数据分析、技术指导;刘璐负责文献分析、数据整理;王慧莹负责资料收集、技术指导;吴瑛负责提供指导、论文审核。

**利益声明:** 所有作者均声明不存在利益冲突。

#### 参考文献

- 1 黄超,杜亮,陈耀龙,等.中国医务人员对临床实践指

- 南的使用和需求调查 [J]. 中国循证医学杂志, 2019, 19 (6): 635 - 638.
- 2 张山, 崔薇, 吴瑛. 基于临床实践指南的临床决策支持系统使用依从性的影响因素 [J]. 中国实用护理杂志, 2023, 39 (13): 98 - 103.
- 3 张山, 崔薇, 吴瑛. 以智能化谵妄护理临床决策辅助支持系统为例探讨规则驱动的临床护理决策支持系统的基本组成 [J]. 护理研究, 2022, 36 (24): 4464 - 4468.
- 4 BERNER E S. Clinical decision support systems: theory and practice [M]. 3rd edition. New York: Springer, 2016.
- 5 HOGAN A, BLOMQUIST E, COCHEZ M, et al. Knowledge graphs [J]. ACM computing surveys, 2021, 54 (4): 71.
- 6 靳淑雁, 王爽, 黄琼, 等. 基于乳腺癌专病库的知识图谱构建研究 [J]. 医学信息学杂志, 2023, 44 (12): 65 - 70.
- 7 方攀, 曹宇汀, 丁子啸, 等. 老年疾病康复知识图谱构建与应用探索 [J]. 医学信息学杂志, 2023, 44 (8): 42 - 48.
- 8 李琳菊, 李亚春, 徐宏宁, 等. 基于知识图谱的中医养生知识智能问答 App 构建研究与实践 [J]. 医学信息学杂志, 2022, 43 (7): 50 - 54.
- 9 DEVLIN J W, SKROBIK Y, GELINAS C, et al. Clinical practice guidelines for the prevention and management of pain, agitation/sedation, delirium, immobility, and sleep disruption in adult patients in the ICU [J]. Critical care medicine, 2018, 46 (9): e825 - e873.
- 10 BRIGNARDELLO - PETERSEN R, CARRASCO - LABRA A, GUYATT G H. How to interpret and use a clinical practice guideline or recommendation: users' guides to the medical literature [J]. JAMA, 2021, 326 (15): 1516 - 1523.
- 11 王永博, 高旷, 李绪辉, 等. 临床实践指南实施性促进研究之二: 基于非肌层浸润性膀胱癌指南的知识图谱框架设计 [J]. 医学新知, 2021, 31 (6): 419 - 432.
- 12 周祎灵, 石清阳, 陈向阳, 等. 本体在糖尿病临床决策支持系统中的应用 [J]. 四川大学学报 (医学版), 2023, 54 (1): 208 - 216.
- 13 黄贺瑄, 王晓燕, 顾正位, 等. 医学知识图谱构建技术及发展现状研究 [J]. 计算机工程与应用, 2023, 59 (13): 33 - 48.
- 14 FRAILE N D, IJAZ K, REZAZADEGAN D, et al. Clinical named entity recognition and relation extraction using natural language processing of medical free text: a systematic review [J]. International journal of medical informatics, 2023, 177 (9): 105122.
- 15 王丽敏, 陈泓伯, 王琦, 等. 以公众健康教育与非药物干预为主的膝关节骨性关节炎疾病知识图谱的构建 [J]. 中华护理杂志, 2022, 57 (10): 1172 - 1177.
- 16 杜晋华, 尹浩, 冯嵩. 中文电子病历命名实体识别的研究与进展 [J]. 电子学报, 2022, 50 (12): 3030 - 3053.
- 17 LIU F, LIU M, LI M, et al. Automatic knowledge extraction from chinese electronic medical records and rheumatoid arthritis knowledge graph construction [J]. Quantitative imaging in medicine and surgery, 2023, 13 (6): 3873 - 3890.
- 18 LIU H, ZHANG Z, XU Y, et al. Use of BERT (bidirectional encoder representations from transformers) - based deep learning method for extracting evidences in Chinese radiology reports: development of a computer - aided liver cancer diagnosis framework [J]. Journal of medical internet research, 2021, 23 (1): e19689.
- 19 刘彬, 肖晓霞, 邹北骥, 等. 融合汉字部首的 BERT - BiLSTM - CRF 中医医案命名实体识别模型 [J]. 医学信息学杂志, 2023, 44 (6): 48 - 53.
- 20 FANG A, HU J, ZHAO W, et al. Extracting clinical named entity for pituitary adenomas from Chinese electronic medical records [J]. BMC medical informatics and decision making, 2022, 22 (1): 72.
- 21 许思特, 孙木. 基于命名实体识别与 Neo4j 的中文电子病历知识图谱构建和应用 [J]. 医学信息学杂志, 2022, 43 (12): 50 - 56.
- 22 LIU Q, TIAN Y, ZHOU T, et al. An explainable and personalized cognitive reasoning model based on knowledge graph: toward decision making for general practice [J]. IEEE journal of biomedical and health informatics, 2024, 28 (2): 707 - 718.
- 23 尹梓名, 杜方芮, 赵紫彤, 等. 基于临床指南的知识图谱构建技术研究 [J]. 软件, 2020, 41 (9): 178 - 184, 197.
- 24 李伟, 王竣生, 秦鹏. 基于知识图谱的医疗问答系统研究 [J]. 长江信息通信, 2023, 36 (6): 107 - 109.