

面向临床研究的中文概念标准化方法研究

吴 宁^{1#} 李子轩^{1#} 蔡文红² 王怀玉³ 袁 驰¹

(¹ 河海大学计算机与软件学院 南京 211100 ² 江苏大学附属医院 镇江 212003

³ 北京中医药大学国家中医体质与治未病研究院 北京 100029)

[摘要] **目的/意义** 探索医学术语标准化方法, 提高医疗服务质量和效率。**方法/过程** 提出基于观测标识符逻辑命名与编码系统的中文医学术语标准化方法, 包括基于中文大语言模型的关键要素提取与基于 BM25 算法的术语映射两个步骤。**结果/结论** 该方法在国家卫生行业标准已标注数据集上映射匹配率达到 73.04%, 验证了其在医学术语标准化中的可行性, 可为提升医疗数据互认与服务效率提供支持。

[关键词] 临床医学; 术语标准化; 关键要素提取; BM25 算法

[中图分类号] R-058 **[文献标识码]** A **[DOI]** 10.3969/j.issn.1673-6036.2025.09.009

Study on Standardized Methods of Chinese Concepts for Clinical Research

WU Ning¹, LI Zixuan¹, CAI Wenhong², WANG Huaiyu³, YUAN Chi¹

¹School of Computer and Software, Hohai University, Nanjing 211100, China; ²Affiliated Hospital of Jiangsu University, Zhenjiang 212003, China; ³National Institute of Traditional Chinese Medicine Constitution and Preventive Treatment of Diseases, Beijing University of Chinese Medicine, Beijing 100029, China

[Abstract] **Purpose/Significance** To explore the standardization methods of medical terminology, so as to enhance the quality and efficiency of medical services. **Method/Process** A method for standardizing Chinese medical terminology based on the logical observation identifiers names and codes (LOINC) is proposed. The approach consists of two steps; key element extraction based on the Chinese large language model (LLM) and terminology mapping based on BM25 algorithm. **Result/Conclusion** The method achieves a mapping match rate of 73.04% on the annotated dataset of the national health industry standard, verifying its feasibility in medical terminology standardization and providing support for improving medical data mutual recognition and service efficiency.

[Keywords] clinical medicine; standardization of terminology; extraction of key elements; BM25 algorithm

1 引言

医疗数据标准化已成为提升医疗服务质量与效率的关键^[2]。然而, 当前医学检测术语存在标准化难题。一方面, 不同医疗机构之间的设备和标准存在差异, 导致实验室检测结果互认困难; 另一方面,

随着人工智能技术在医疗领域的深入应用^[1],

[修回日期] 2025-07-22

[作者简介] 吴宁, 硕士研究生; 李子轩, 硕士研究生; 通信作者: 袁驰;[#]对本文有相同贡献。

[基金项目] 国家自然科学基金项目 (项目编号: 62302151); 国家中医药管理局高水平中医药重点学科——中医体质学 (项目编号: zyyzdx-2023251); 北京市自然科学基金海淀原始创新联合基金 (项目编号: L232118)。

中文医学术语具有多义性和表述不一致性等特点,对临床研究可靠性^[3]和医疗信息化产生一定影响。医学信息标准化与规范化受到国内外广泛关注。在国际上,世界卫生组织、国际标准化组织等已经建立成熟的术语标准体系^[4-8],如 Forrey A W 等^[4]在 1994 年提出的观测指标标识符逻辑命名与编码系统(logical observation identifiers names and codes, LOINC),用以标识医学实验室观察和临床化验结果^[9],目前已被广泛使用。但国外术语体系并不完全契合国内临床信息化需求^[10]。国内虽已制定《医学主题词表》(Chinese medical subject headings, CMeSH)^[11]、《常用临床医学名词(2023 年版)》^[12]、《临床检验项目分类与代码》^[13]等标准,但仍存在体系分散、覆盖不全等问题。因此,开发适配中文语境、满足临床需求的术语标准化方法具有重要意义。

2 相关工作

2.1 临床术语库

术语库也被称为术语数据库,是一种专门用于组织、存储、管理和检索数据及相关信息的数据库系统。在特定学科领域,术语库扮演着关键角色,其核心内容包含术语以及与术语相关的定义、解释、来源、用法、语境等信息^[14]。在医学领域,术语库可以为专业人员提供规范统一的术语和相关信息,支持其进行更加精确的研究和实践。如美国国立医学图书馆于 1986 年建设的统一医学语言系统(unified medical language system, UMLS),不仅促进了不同医学术语系统之间的融合与转换,还为生物医学研究者提供了一个全面、准确且及时更新的知识资源平台^[15]。

临床术语库作为医学术语库的重要分支,专注于临床实践的专业用语标准化。其核心价值在于为临床诊疗提供统一、规范的术语系统,从而提升医疗质量、优化诊疗流程。临床术语库不仅继承了医学术语库的专业性、规范性、系统性和动态性,更强调实用性、简洁性和国际通用性,以满足临床工作的高效需求。如医学系统命名法-临床术语(systematized nomenclature of medicine-clinical terms, SNOMED CT)^[6],涵盖多方面的临床信息,不仅有极高的灵

活性和可扩展性,也支持多种语言,可满足不同国家、地区的医学信息交流和管理需求。

近年来我国在医学领域的标准化、规范化和信息化建设方面取得显著进展^[16],已连续颁布众多卫生信息标准,但尚未构建起完善且得到广泛应用的临床术语集合,在医学信息的准确性和一致性方面仍待进一步研究。

2.2 基于语料库的概念归一化方法

数据的质量和准确性是影响决策的关键因素。不同特征之间的度量单位和尺度不同,可能使数据分布出现偏差,影响建模和分析结果。数据归一化是一种常见的数据预处理技术,通过对数据进行数学变换、按比例缩放,将其映射到特定的范围内,使不同特征之间具有可比性,是优化数据处理的重要技巧,可用于解决上述问题。

除了最值、均值归一化等线性方法外,还有对数变换、幂函数变换等非线性归一化方法。如在自然语言处理和文本挖掘领域,词频-逆文档频率(term frequency-inverse document frequency, TF-IDF)作为一种广泛应用的统计手段,通过词汇在特定文档中出现的频次(TF)以及在文档集合中的罕见程度(IDF),量化其在单个文档内的关键性^[17]。然而,TF-IDF 主要聚焦于词频和文档频率的考量,在语义相似性评估方面作用有限。最佳匹配(best match, BM)算法是基于词频统计和文档长度归一化的改进型 TF-IDF 检索算法,经过 25 次迭代,得到 BM25。BM25 算法通过引入词项饱和函数和长度因子,能够更准确地评估搜索词与文档的相关性得分。

3 数据集构建

3.1 原始数据来源

在数据科学与机器学习领域,数据来源的规范性与可靠性是支撑研究与应用的核心基础。数据来源体系涵盖公共数据库、私有数据集、互联网公开数据及统计资料等多重维度,其结构特征直接决定了数据的可信度与权威性。鉴于当前缺乏针对临床医学中文概念规范化的标准数据集,本研究通过与医院、检验机

构等建立紧密合作关系，系统性获取高精度实验原始数据，包括来自迪安诊断的 4 268 条精准数据，以及江苏大学附属医院提供的 94 431 条详尽数据。每条数据记录均包括检测项目名称、分析条目、计量单位、样本、参考区间等信息，见表 1—表 2。同时收

集国家卫生行业标准《居民健康档案医学检验项目常用代码》中 204 条已标注数据作为标准数据集，见表 3。其中，两家合作医学实体提供的数据用于检验人工标注结果一致性，标准数据集用于验证本研究方法映射结果的准确率。

表 1 机构提供的数据样例

检测项目名称	分析项目	项目简称	结果类型	计量单位	参考区间
转铁蛋白（TRF）测定	转铁蛋白	TRF	定量型	g/L	2.00 ~ 3.60
血沉（ESR）	血沉	ESR	定量型	mm/h	男：0 ~ 15。女：0 ~ 20
尿肌酐（Ucr）测定	尿肌酐	Ucr	定量型	mmol/L	男：3.90 ~ 27.10。女：2.80 ~ 22.00
粪便常规检查	白细胞	LEU	叙述型	—	阴性或偶见

表 2 医院提供的数据样例

检测项目名称	分析项目	检测样本	简称	计量单位	参考区间
红细胞沉降率测定（ESR）	血沉	全血	XC	mm/h	0 ~ 20
B 型钠尿肽 BNP	B 型钠尿肽	血浆	BNP	pg/mL	< 100
尿常规 + 尿沉渣	尿钙离子	尿	NNCA	mmol/L	2 ~ 9
甲型肝炎抗体检查	甲肝抗体 - IgM	血清	HAVIgM	S/CO	无反应（< 0.8）

表 3 行业标准数据样例

项目类别	项目简称	LOINC 代码	项目全称	结果报告示例
血液一般检查	红细胞沉降率	30341 - 2	全血 · 红细胞沉降率 · 速度（定量型）	n，mm/h
尿液一般检查	尿胆红素	1978 - 6	尿液 · 胆红素 · 质量浓度（定量型）	n，mg/dL
体液与分泌物检查	精液外观	13359 - 5	精液 · 性状 · 外观（名义型）	1 = 正常。2 = 异常
无机物质测定及血气分析	血清钠	2947 - 0	全血 · 钠 · 物质的量浓度（定量型）	n，mmol/L

3.2 数据标注

3.2.1 标注流程设计 为系统评估人工标注的可靠性与一致性，从机构和医院提供的未标注数据集中随机抽取 151 个样本，采用独立标注法，由两名具有医学背景的标注员分别执行标注任务。预先制定严格的标准化操作规程，以最大限度减少系统误差并优化人力资源配置。人工标注流程，见图 1。先根据报告简称（REPORT_ITEM_CODE）在 Athena 平台执行初始检索，通过左侧的筛选栏设置概念标签进行结果过滤，之后结合单位（UNITS）寻找匹配项。如果获得唯一匹配项，则提取对应英文术语完成标注。多数情况下仅凭借报告简称难以直接匹配目标概念，须进一步结合报告项目名称（REPORT_ITEM_NAME）和检测样本（SPECIMEN）

综合选取关键词，重复 Athena 使用流程，见图 2，实现术语映射。作为英文 LOINC 标准的官方检索工具，Athena 平台通过多维度属性关联支持临床术语的标准化查询，其操作逻辑涵盖代码检索、属性过滤等核心功能模块。

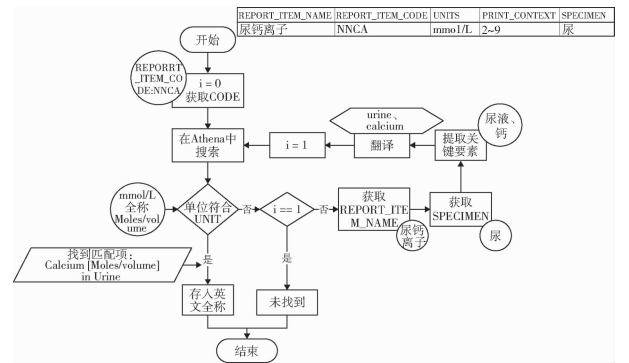


图 1 人工标注流程

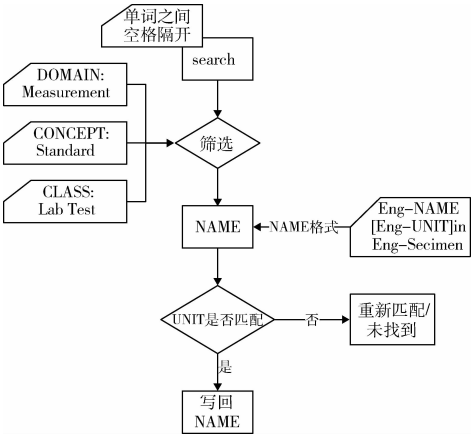


图 2 Athena 使用流程

3.2.2 数据标注结果 为建立高可靠性的标注基准,设计并实施3轮迭代式标注流程。在确保标注员充分理解任务要求后,采用渐进式质量控制策略:

每轮标注完成后均实施双重校验,通过交叉比对发现标注差异,并进行流程优化。首轮标注完成后,经一致性分析发现两名标注员的标注一致性系数仅为 30.46%,显著低于预期水平。经过对标注流程与规则体系的系统性复盘,识别出术语映射规则存在语义覆盖不全、操作指引模糊等问题。为提升后续标注质量,在第 2 轮标注启动前开展预标注测试,通过案例解析与操作示范强化标注员对规范的理解。经多次反馈迭代与标注规则优化,第 3 轮共完成 98 条样本标注,标注一致性系数提升至 73.47%。两名标注员的标注结果,见表 4。结果加粗代表标注结果不同,加粗部分的数据标注正确。由于人工标注过程需要较高的时间与人力成本,且标注一致性尚未达到理想水平,因此该标注结果仅作为后续实验方法对比的参考基准。

表 4 两名标注员标注结果样例

报告项目名称	报告简称	单位	检测样本	标注员 1	标注员 2
尿钙离子	NNCA	mmol/L	尿	Calcium [Moles/volume] in Urine	Calcium [Moles/volume] in Urine
嗜碱性粒细胞绝对值	BASO#	10 ⁹ /L	全血	Basophils [# /volume] in Blood	Eosinophils [# /volume] in Blood
红细胞计数	RBC	10 ¹² /L	末梢血	Erythrocytes [# /volume] in Blood	Erythrocytes [# /volume] in Blood

4 研究方法

4.2 基于中文大语言模型的关键要素提取方法

4.1 总体框架

4.2.1 提示词体系设计 由于 LOINC 标准属性范围具有局限性,设计面向关键要素提取的提示词体系,包括 3 个核心维度:成分、检测样本和精度,分别对应 LONIC 标准的成分、体系和精度。其中成分指医学检测对象,如红细胞、血红蛋白等。检测样本与精度均是可枚举的。检测样本包含全血、动脉全血、静脉全血、血清、血浆、尿液、大便、精液、痰液等。精度即测量标尺类型,包括定量型、名义型、叙述型、序数型和文档型。基于实证数据分析发现,精度与检测项目的计量单位之间存在显著关联性,具体对应关系,见表 5。

构建基于 LOINC 标准的中文医学术语映射方法,见图 3。输入非标准化的医学检验信息,设计提示词,使用大语言模型提取关键要素,采用 BM25 算法计算关键要素与 LOINC 中各项数据之间的相关性得分,最后将各关键要素所对应的得分相加,筛选出最佳匹配项结果,即该条医学检验信息所对应的 LOINC 代码。

表 5 计量单位与精度的对应关系

单位类型	精度	单位类型	精度
质量	定量型	速率	定量型
浓度	定量型	任意型浓度	序数型
比率	定量型	检查诊断意见	叙述型
体积	定量型	外观、类型	名义型

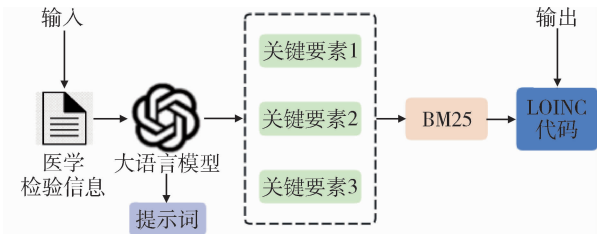


图 3 基于 LOINC 标准的中文医学术语映射框架

4.2.2 基于提示词的关键要素提取和结果归集

随着人工神经网络与深度学习的发展,自然语言处理技术已广泛应用于医疗健康领域^[18]。当前医学大语言模型多基于 GPT、GLM、LLaMA 等通用模型通过微调或二次开发构建^[19]。与仅支持应用程序接口(application programming interface, API)调用的模型(如 OpenAI GPT)相比, GLM 作为开源模型,既支持本地微调定制,也提供在线 API 服务,具备更强的灵活性与可定制性。选用基于 GLM 架构开发的 ChatGLM3-6B 模型^[20]作为核心自然语言理解引擎。该模型于 2023 年 3 月开源,具备 62 亿参数规模,可本地化部署于消费级 GPU,且 INT4 量化级别下最低仅需 6GB 显存即可运行^[21]。模型处理流程,见图 4,基于预定义的提示词模板生成结构化输出文本,为临床术语标准化提供语义解析支持。通过设计合适的提示词精确表达需求,使大语言模型能够有效地协助处理数据,获得理想输出。对国家行业标准中的 204 条已标注数据进行实验,根据数据中的项目名称、单位、样本等信息进行基于提示词的关键要素提取,成功提取出相应关键要素。ChatGLM3 提取的数据以 JSON 格式输出,为了提高数据处理效率、更直观地理解和分析数据、提高数据的可读性和可理解性,对 JSON 数据进行预处理,并将结果归集。

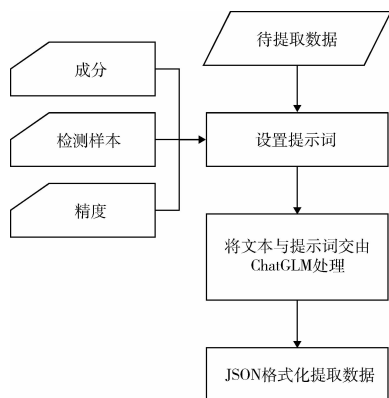


图 4 基于提示词的关键要素提取流程

4.3 基于 BM25 算法的术语映射方法

BM25 算法的核心是得分函数,可以计算文档

与查询条件之间的匹配程度得分,即相关性得分。

$$S(Q, d) = \sum_i^n W_i R(q_i, d) \quad (1)$$

其中 Q 代表查询词, d 代表文档, q_i 表示 Q 中的每个分词, W_i 表示每个分词的权重, R 表示每个分词和文档的相关性。该得分函数的计算主要包括 3 部分。一是 W_i , 即每个分词 q_i 的权重, 其实也是其 IDF 的计算。

$$W_i(q_i) = \log \frac{N - df_i + 0.5}{df_i + 0.5} \quad (2)$$

其中, N 表示所有文档个数, df_i 表示包含 q_i 的文档数, df_i 越高权重越小。二是每个分词 q_i 与文档的相关性计算。

$$S(q_i, d) = \frac{(k_1 + 1)tf_{id}}{K + tf_{id}} \quad (3)$$

$$K = k_1(1 - b + b \times \frac{L_d}{L_{ave}})$$

其中 tf_{id} 单词 t 在文档 d 中的词频, L 表示文档长度, k_1 为正参数, b 控制文档长度对词权重的影响程度。 K 是一个动态归一化因子, 对词频 tf_{id} 进行标准化, 消除文档长度差异带来的不公平影响。三是分词 q_i 与 Q 的相关性计算。

$$S(q_i, Q) = \frac{(k_3 + 1)tf_{iq}}{k_3 + tf_{iq}} \quad (4)$$

其中 tf_{iq} 为单词 t 在 Q 中的词频, k_3 是可调节正参数, 用于控制 tf_{iq} 对最终相关性得分的影响程度和饱和速度。由此得到文档 d 与查询的 BM25 相关性得分的检索状态值 (retrieval status value, RSV)。

$$RSV_d = \sum_{i \in q} \left[\log \frac{N - df_i + 0.5}{df_i + 0.5} \right] \times \frac{(k_1 + 1)tf_{id}}{k_1(1 - b + b \times \frac{L_d}{L_{ave}}) + tf_{id}} \times \frac{(k_3 + 1)tf_{iq}}{k_3 + tf_{iq}} \quad (5)$$

因为直接采用 BM25 算法对医学数据中的项目名称进行 LOINC 映射结果不理想, 所以先从待匹配数据中提取关键词, 再针对各关键词进行筛选和相似度计算, 其算法流程, 见图 5。先依据待匹配项的精度对 LOINC 标准术语集进行初步筛选, 再分别计算成分和检测样本对筛选后信息的相关性得分, 两者相加取最高得分, 输出其所对应的 LOINC 代码即为最终结果。

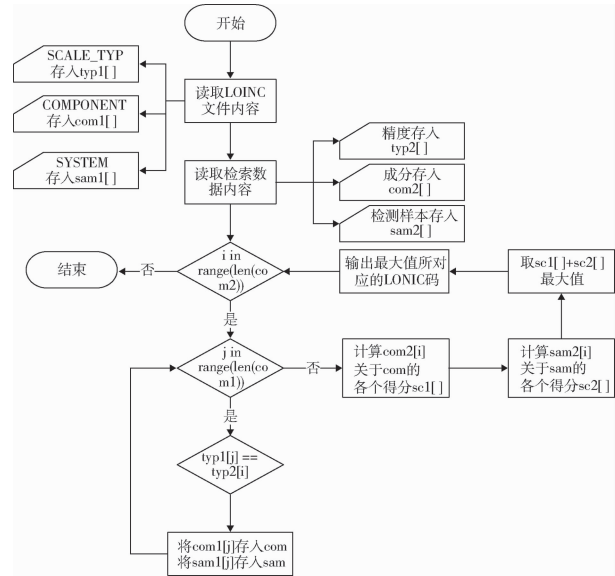


图 5 BM25 算法流程

5 实验与结果分析

5.1 评价指标和实验结果

评价指标能反映模型部分性能，如果选取不合理可能导致错误结论。针对本研究数据，选取准确率分析方法性能。本研究中 BM25 算法对一条数据的处理过程，见图 6。

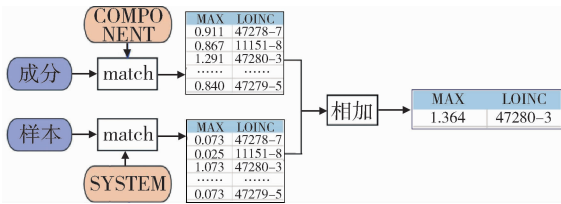


图 6 BM25 算法映射示例

比较每条数据处理结果与标准数据集中同一项目所对应的 LOINC 值，计算映射准确率。经统计，对于《居民健康档案医学检验项目常用代码》中 204 条已标注数据，正确映射 149 条，错误映射 55 条，准确率为 73.04%。

对比本研究方法与其他方法的准确率，见表 6。本研究方法的准确率显著高于其他基准方法。实验过程中发现关键词提取阶段存在关键要素覆盖不全的问题，导致部分映射结果出现偏差，整体准确率尚未达到预期水平，但本研究验证了所提方法的技术可行性。实验结果表明，该模型在中文临床术语标准化领域具有应用价值，为解决医疗信息系统中术语异构性问题提供了技术支撑，对推动我国医疗信息标准化体系建设具有积极意义。

表 6 本研究方法与其他方法的准确率比较

方法	准确率（%）
BM25	12.75
TF-IDF	26.50
Naive Edit Distance ^[22]	72.30
TF-IDF + Jaccard ^[23]	63.56
本研究方法	73.04

5.2 错误分析

分析映射错误的 55 条数据，找出通过术语归一化匹配方法映射出的 LOINC 代码所对应的信息，以及标准数据集中该项数据的 LOINC 代码所对应的信息，进行对比并标注，见表 7。

表 7 映射结果错误标注样例

输入	成分	检测样本	精度	LOINC	标准 LOINC	错误原因
Rh 血型 全血 · Rh · 类型（名义型）	Rh	全血	名义型	10331-7	10331-7	—
Rh 血型 全血 · Rh · 类型（名义型；毒理学确认试验）	Rh	全血	名义型	10331-7	34961-3	方法错误
阴道分泌物滴虫检查 XXX · 阴道毛滴虫 · 任意型浓度（序数型；湿片法）	阴道毛滴虫	阴道	序数型	14367-7	32766-8	检测样本应为 XXX
艾滋病病毒 1+2 抗体定量检测 血清 · HIV 1+2 抗体 · 任意型浓度（定量型）	艾滋病 1+2 抗体	血清	定量型	16796-5	22357-8	标准 LOINC 成分中为 HIV

各种错误原因统计，见表 8。半数以上的错误是由于在提示词设计时未考虑不同检验方法对映射结果的影响。成分问题导致的映射错误原因较复

杂，共占 32.7%。检测样本提取错误导致的映射结果错误占比 16.4%。

表 8 错误原因统计

维度	具体错误原因	示例	数量（个）	占比（%）
成分	成分提取不完全	“皮质醇（上午）”和“皮质醇（下午）”均提取为“皮质醇”	10	18.2
	全称与简称无法匹配	提取：β。标准 LOINC：Beta	3	5.4
	LOINC 成分中未作区分	“淋巴细胞绝对值”和“淋巴细胞百分率”在标准 LOINC 成分中均为“淋巴细胞”	5	9.1
检测样本	样本提取错误	提取：阴道。标准 LOINC：XXX	9	16.4
精度	未识别具体方法	LOINC 中存在成分、体系、精度均相同但方法不同的临床观测指标，无法区分	28	50.9

5.3 中文语境下 LOINC 适配局限性讨论

5.3.1 语义粒度差异导致映射失配 LOINC 以刚性六元组结构构建原子级概念，而中文检验报告常将多项指标合并为单一术语（如“肝功五项”），其语义粒度显著粗于 LOINC。该差异导致中文术语无法直接对应 LOINC 原子概念，须在映射前进行语义拆分，然而拆分规则缺乏统一标准，易造成成分遗漏或属性误配，进而降低映射准确性。

5.3.2 同义异构表达引发的对齐困难 中文语境下存在大量同义异构表述，同一检验项目可呈现全称、简称、俗称等形态（如“红细胞计数”亦称“RBC”或“红血球数”），且缺乏规范化约束。LOINC 英文术语体系未覆盖此类变体，导致基于字面相似度的匹配算法难以捕捉语义等价关系，进而影响映射准确性。

5.3.3 语言结构差异引发的句法解析障碍 中文医学报告中普遍采用“修饰语 + 中心词”的偏正式结构，修饰语可多层嵌套（如“血清游离三碘甲状腺原氨酸测定”）。LOINC 英文术语则以固定语序和介词短语实现语义分解，两者句法之间差异显著。在解析中文长术语时易出现切分歧义，造成成分边界错误，进而导致 LOINC 六元组映射失败。

6 结语

本研究提出一种基于 LOINC 标准的医学术语归一化匹配方法，旨在解决中文医学术语标准化不足的

问题。该方法虽然存在一定局限性，但实验结果表明，其在术语标准化任务中具有可行性，优化标注流程设计或完善提示词策略可能打破局限、弥补该方法缺陷。本研究为中文临床术语标准化提供了新思路，对推动医疗信息标准化建设具有积极意义。

作者贡献：吴宁负责研究实施、数据分析、论文撰写；李子轩负责数据采集、人工标注流程设计；蔡文红负责医院数据预处理；王怀玉负责检验机构数据预处理；袁驰负责提供指导、论文审核与修订。

利益声明：所有作者均声明不存在利益冲突。

参考文献

1 KIM J, LEE D, PARK E. Authors’ reply to: bibliometric studies and the discipline of social media mental health research. Comment on “machine learning for mental health in social media: bibliometric study” [J]. Journal of medical internet research, 2021, 23 (6): e29549.

2 ALONSO I, CONTRERAS D. Evaluation of semantic similarity metrics applied to the automatic retrieval of medical documents: an UMLS approach [J]. Expert systems with applications, 2016, 44 (2): 386 – 399.

3 任慧玲, 李亚子, 李晓瑛, 等. 我国临床医学学术语体系建设路径与实践探索 [J]. 医学信息学杂志, 2022, 43 (8): 2 – 7.

4 FORREY A W, MCDONALD C J, DEMOOR G, et al. Logical observation identifier names and codes (LOINC) database: a public use set of codes and names for electronic reporting of clinical laboratory test results [J]. Clinical chemistry, 1996, 42 (1): 81 – 90.

- 5 U. S. National Library of Medicine. Medical subject headings [EB/OL]. [2025 - 04 - 30]. <https://www.nlm.nih.gov/mesh/meshhome.html>.
- 6 SNOMED International. SNOMEDCT [EB/OL]. [2025 - 04 - 30]. <https://www.snomed.org/snomed-ct>.
- 7 DOAN V T T. Establishing a meddra safety surveillance unit [J]. Therapeutic innovation regulatory science, 2000, 34 (1): 245 - 250.
- 8 World Health Organization. International classification of diseases [EB/OL]. [2025 - 04 - 30]. <http://www.who.int/classifications/icd/en/>.
- 9 张林, 赵英杰, 陈兴. 观测指标标识符逻辑命名与编码系统 (LOINC) 数据库介绍 [J]. 河北省科学院学报, 2004 (4): 66 - 68, 71.
- 10 任慧玲, 郭进京, 孙海霞, 等. 医学术语标准化研究的思考 [J]. 医学信息学杂志, 2018, 39 (5): 2 - 7.
- 11 中国医学科学院医学信息研究所. 中文医学主题词表 [EB/OL]. [2025 - 04 - 30]. <http://cmesh.imicams.ac.cn/index.action?action=index&noMsg=1>.
- 12 国家卫生健康委印发常用临床医学名词 (2023 年版) [J]. 中国乡村医药, 2024, 31 (7): 25.
- 13 原国家卫生和计划生育委员会. 临床检验项目分类与代码 [EB/OL]. [2025 - 04 - 30]. <https://www.nhc.gov.cn/zwgkzt/s9492/201212/33995.shtml>.
- 14 冯志伟. 术语数据库 [J]. 语文建设, 1989 (2): 53 - 58.
- 15 BODENREIDER O. The unified medical language system (UMLS): integrating biomedical terminology [J]. Nucleic acids research, 2004, 32 (S1): 267 - 270.
- 16 杨威, 乔政, 晁燕丽. 国外医学术语数据库浅析及我国医学术语库发展建议 [J]. 中国继续医学教育, 2021, 13 (17): 85 - 88.
- 17 鲍治国, 王海安, 胡士伟, 等. 基于内容相似度的相关性评分算法对比分析研究 [J]. 电子测试, 2022 (19): 52 - 55.
- 18 沈红, 崔子樟, 曾淑君, 等. 基于自然语言处理 (NLP) 的医学知识挖掘探索与实践 [J]. 健康教育与健康促进, 2024, 19 (2): 155 - 157, 217.
- 19 王泰一, 靳擎, 范梦月, 等. GPT 还是 GLM? 大模型应用于中医药现代化的机遇与挑战 [J]. 中国药理学与毒理学杂志, 2023, 37 (S1): 5 - 9.
- 20 ZENG A, LIU X, DU Z, et al. Glm - 130b: an open bilingual pre - trained model [EB/OL]. [2024 - 12 - 20]. <https://keg.cs.tsinghua.edu.cn/glm-130b/posts/glm-130b>.
- 21 顾鹏辉, 李涛, 高阳. MedKGGPT: 基于知识图谱的医疗大型语言模型设计方法 [J]. 计算机技术与发展, 2024, 34 (6): 178 - 184.
- 22 DOGAN R, LU Z Y. An inference method for disease name normalization [EB/OL]. [2024 - 12 - 20]. <https://aaai.org/papers/05604-5604-an-inference-method-for-disease-name-normalization/>.
- 23 林楠铠, 林晓钿, 吴凯莹, 等. 基于多策略的临床术语标准化 [EB/OL]. [2025 - 04 - 30]. <https://di4s.chinaxiv.org/abs/202307.00104v1>.

(上接第 29 页)

- 18 PRINS F J, VEENMAN M V J, ELSHOUT J J. The impact of intellectual ability and metacognition on learning: new support for the threshold of problematicity theory [J]. Learning and instruction, 2006, 16 (4): 374 - 387.
- 19 国家卫生健康委员会. 2023 中国卫生健康统计年鉴 [M]. 北京: 中国协和医科大学出版社. 2024.
- 20 李韬, 张珺, 李亚哲, 等. 人工智能在医疗健康领域的创新应用、风险挑战与治理对策 [J]. 医学信息学杂志, 2025, 46 (1): 2 - 8, 16.
- 21 马勇. 数智赋能 助推基层服务能力提升 [J]. 中国卫生, 2025 (5): 22.
- 22 向茜, 张跃祥, 栾艳, 等. 紧密型医共体基层全科医师知识技能培训需求调查及对策研究 [J]. 医学理论与实践, 2025, 38 (13): 2334 - 2338.
- 23 董怡, 冉晔, 余中光. 我国医疗人工智能风险研究现状及风险识别 [J]. 医学信息学杂志, 2025, 46 (3): 24 - 28.
- 24 于良芝. 从社会实践到个人信息世界——基于“实践—经验—心灵”交互性的信息分化理论框架 [J]. 图书馆理论与实践, 2024 (1): 1 - 8, 35.
- 25 邢春国, 夏迎秋, 廖云霞, 等. 江苏省基层医疗卫生机构公共卫生信息化建设现状研究 [J]. 中华临床医师杂志 (电子版), 2022, 16 (12): 1201 - 1206.
- 26 林恺, 姚弥, 林桂丽, 等. 青年全科医师科研素养及培训需求调查 [J]. 中华全科医师杂志, 2022, 21 (11): 1056 - 1062.