



基于人工智能的病历质控系统的应用研究*

温煜, 李雄[△], 曾菲菲, 雷佳雨, 陈绍勇

梅州市人民医院(梅州 514000)

【摘要】目的 通过人工智能技术探索自动化病历质控方法,规范病历书写流程,解决人工质控弊端。**方法** 本文构建了基于人工智能的病历质控系统,该系统首先依据权威标准和专家意见设计并构建质控规则库,通过数据采集引擎自动采集病历数据,然后通过后结构化引擎转换为结构化数据,最后由病历质控引擎结合规则库分析数据,进行质量问题判定,实现自动化智能质控。将该系统应用于病历质控,选取现病史雷同、主诉描述缺陷、初步诊断不全、月经婚育史缺失、主诉现病史不匹配5个质控点,随机抽取2022年1月的2 918份出院病历进行人工智能质控,然后组织病历质控专家进行正确性复核,并对比既往人工质控记录,分析结果。以复核正确的问题数作为金标准,对5个质控点进行受试者工作特征(ROC)曲线分析。**结果** 根据病历质控专家复核,人工智能质控正确率达到89.57%。通过对比抽样病历的人工智能质控和既往人工质控结果,抽样病历既往人工质控检出问题中仅有1个在人工智能质控系统中未检出,人工智能质控正确检出病历质量问题的数量约为人工质控的2.97倍。ROC曲线分析示,人工智能质控组的5个质控点AUC值均有统计学意义($P<0.05$),且AUC值均接近或大于0.9,而人工质控组仅“现病史雷同”质控点AUC值(0.797)有统计学意义($P<0.05$);组间AUC值比较示,人工智能质控组在5个质控点上比人工质控更具有优势。**结论** 通过基于人工智能的病历质控系统的应用,能够实现高效的病历文书全量质控,有效提高质量问题检出率,有助于节约人力,提升病历书写质量。

【关键词】 人工智能 自然语言处理 医学自然语言处理 病历质控

Application of Medical Record Quality Control System Based on Artificial Intelligence WEN Yu, LI Xiong[△], ZENG Feifei, LEI Jiayu, CHEN Shaoyong. Meizhou People's Hospital, Meizhou 514000, China

[△] Corresponding author, E-mail: 443762191@qq.com

【Abstract】 Objective In this study, we used artificial intelligence (AI) technology to explore for automated medical record quality control methods, standardize the process for medical record documentation, and deal with the drawbacks of manually implemented quality control. **Methods** In this study, we constructed a medical record quality control system based on AI. We first designed and built, for the system, a quality control rule base based on authoritative standards and expert opinions. Then, medical records data were automatically collected through a data acquisition engine and were converted into structured data through a post-structured engine. Finally, the medical record quality control engine was combined with the rule base to analyze the data, identify quality problems, and realize automated intelligent quality control. This system was applied to the quality control of medical records and five quality control points were selected, including similarities in the history of the present illness, defects in the description of chief complaints, incomplete initial diagnosis, missing information in the history of menstruation, marriage, and childbirth, and mismatch between the chief complaints and the history of the present illness. We randomly selected 2 918 medical records of patients discharged in January 2022 to conduct AI quality control. Then we organized medical record quality control experts to conduct an accuracy review, made a comparison with previous manual quality control records, and analyzed the results. The number of quality problems that were verified in the accuracy review was taken as the gold standard and receiver operating characteristic (ROC) curves were drawn for the 5 quality control points. **Results** According to the accuracy review performed by medical record quality control experts, the accuracy of AI quality control reached 89.57%. For the sampled medical records, the results of AI quality control were compared with those of previous manually performed quality control and only one problem detected by manual quality control of the sampled medical records was not detected by the AI quality control system. The number of medical record quality problems correctly detected by AI quality control was about 2.97 times that of manual quality control. Analysis of the ROC curves showed that the AUC of the five quality control points of the AI quality control system were statistically significant ($P<0.05$) and all the AUC values approximated or exceeded 0.9. In contrast, results obtained through manually performed quality control found significant AUC (0.797) for only one quality control point—similarities in the history of present illness ($P<0.05$). Comparison of the AUC values of the two quality control methods showed that AI quality control system had an advantage over manually

* 梅州市社会发展科技计划项目(No. 2022B22)资助

[△] 通信作者, E-mail: 443762191@qq.com

出版日期: 2023-11-20

performed quality control for the five quality control points. **Conclusion** Through the application of medical record quality control system based on AI, efficient full quality control of medical record documentation can be achieved and the detection rate of quality problems can be effectively improved. In addition, the system helps save manpower and improve the quality of medical record documentation.

【Key words】 Artificial intelligence Natural language processing Medical Language Processing
Quality control of medical records

医疗质量管理是规范医疗服务行为、保障医疗安全的管理方法,病案质量控制是以医疗质量核心内容为控制对象,以提高医疗及时性、安全性、有效性为目标^[1]。病历质量不但关系到医院整体医疗水平的发展,还会对数据的二次利用产生决定性影响^[2]。目前,大部分医院的病历质控主要还是依靠人工完成,质控效果未能达到预期,存在的问题包括但不限于:①效率低,准确性差;②耗时长,覆盖率低;③偏事后,实时性差;④管理弱,追踪困难。因此,找到更加便捷、高效的质控手段来提高病历质量,是亟待解决的问题。

随着信息技术和医疗科学的快速发展,特别是自然语言处理(natural language processing, NLP)技术实现了对自由文本进行结构化处理^[3],以及由此发展而来的医学自然语言处理(medical language processing, MLP),把NLP应用到医学领域当中,从病历文本当中得到结构化数据,提取有效信息,同时结合机器学习、知识图谱等技术训练医学语义理解模型^[4-5],使人工智能病历质控系统在医院应用有了广阔的应用前景。本研究通过基于人工智能的病历质控系统的应用,探索自动化病历质控方法,以解决传统人工质控费时、费力、滞后、检出率低等弊端,发挥信息系统快速、准确、全覆盖的优势,规范病历书写流程。

1 设备和方法

1.1 硬件

基于人工智能的病历质控系统采用六台服务器,两两互为热备份,分别部署数据采集服务及病历质控数据库、后结构化引擎及高结构化数据库、病历质控引擎及质控管理系统。

1.2 软件

基于人工智能的病历质控系统由数据采集服务及病历质控数据库、后结构化引擎及高结构化数据库、病历质控引擎及质控管理系统三大部分组成。其中,数据采集服务包含数据集成、数据预处理与清洗、数据抽取与融合,将采集处理后的数据保存到病历质控数据库。后结构化引擎主要进行NLP,通过自然语言分词以及上下文语义识别,利用机器学习结合数据模型实现对自由文

本病历、检查报告、检验报告、病理报告等文档进行规范化后结构化处理,得到高结构化数据。病历质控引擎由首页规则库、内涵规则库、基础规则库、预警规则库、自定义规则库组成,规则运算的结果服务于质控管理系统,实现病历内涵质控、环节质控、终末质控等功能。

1.3 系统架构

本系统采用NLP和机器学习等人工智能技术实现自动化病历质控。如图1所示,病历质控系统主要分为数据采集引擎、后结构化引擎、病历质控引擎和质控管理系统四个模块。数据采集引擎负责从医院信息系统(hospital information system, HIS)、实验室信息系统(laboratory information system, LIS)、影像归档和通信系统(picture archiving and communication system, PACS)等各医疗业务系统中采集数据到本系统数据库中;后结构化引擎负责将非结构化的病历文本数据处理成高结构化病历数据,为后续构建病历质控引擎提供数据基础;病历质控引擎从完整性、时限性、逻辑性、合规性四个维度对病历质量进行检查;本系统负责与用户进行交互,提供病历查阅与浏览、质控提醒与报警、质控结果统计与分析、质控规则配置等功能。

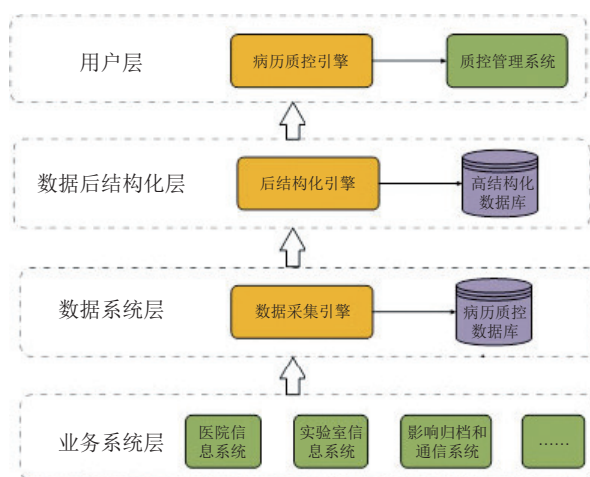


图1 病历质控系统架构图

Fig 1 Architecture diagram of the medical record quality control system

1.4 系统实施

1.4.1 质控规则设计

为确保质控规则的权威、标准,本系统依据《病历书

写基本规范》[卫医政发〔2010〕11号]、《医院住院病历质量评分标准》和《广东省门(急)诊病历质量评定标准》等文件,同时综合三甲医院高年资质控医生专业评审意见,进一步完善、细化病历质控规则。质控规则按类型可分为五类,包括首页规则、基础规则、内涵规则、预警规则和自定义规则。各类规则对应多个质控库,详细质控库和启用情况如表1所示。部分规则基于标准医学术语进行表达,例如首页诊断质控库规则基于ICD-10(医学版),首页手术质控库规则基于ICD-9-CM-3(医学版)。

表 1 质控规则库分类及规则启用情况
Table 1 Classification and rule activation status of the quality control rule base

规则类型	质控库	设置规则数量	启用规则数量
首页规则	首页基础质控库	49	39
	首页一致质控库	14	8
	首页诊断质控库	51	37
	首页手术质控库	18	10
内涵规则	文书一致质控库	41	32
	多源一致质控库	27	23
	多次就诊一致质控库	2	2
	病历诊断质控库	49	38
	专病逻辑质控库	45	36
	真实病情质控库	37	27
	诊治分析质控库	38	31
基础规则	文书缺陷质控库	44	33
	文书超时质控库	26	25
	文书完整质控库	59	52
	数据合理质控库	34	27
	文书重复质控库	13	7
预警规则	书写预警库质控库	1	1
自定义规则	知识库自定义规则	1	1

以文书重复质控库中的“入院记录现病史与首次病程记录病史雷同”规则为例,该规则的设定为“如果入院记录现病史与首次病程记录病史内容相似度超过70%则为雷同”。系统通过TF-IDF结合余弦相似度算法计算入院记录现病史与首次病程记录病史内容相似度,得到的值大于70%时,则判定为雷同。在雷同率判断这一块目前国内常用做法是对比文书数字相同个数,文字相同达到一定比例判定为雷同,没有对段落中文字进行分词、句法分析、赋权等操作,也没有相应的完整语料库供参考。

1.4.2 数据采集引擎

病历是患者在医疗机构诊疗及相关处理过程的资料与信息的汇总,包括:患者历次住院期间的历史就诊信息、检验检查报告及影像资料、患者本次住院期间的病案内容及相关辅助检查、辅助治疗等医疗过程信息。

为了确保病历的医学术语和操作的书写规范,保证病案首页信息资料的准确与统一,减少因数据映射或关联错误所造成的病历质量问题,本系统实现了与其他医疗业务系统的互联互通,通过数据采集引擎从其他医疗业务系统中采集相关数据。

数据采集引擎依照医院信息互联互通标准与质量管理的需要,合理配置需要采集的数据目录,通过应用程序接口与数据库视图的方式相结合进行数据提取、清洗、转换、加载到病历质控数据库中,为后续处理奠定数据基础。清洗过程当中除了去除“脏”数据,同时对患者个人隐私信息如姓名、身份证号码、电话、住址等进行了脱敏处理,以保护患者隐私。经梅州市人民医院医学伦理委员会审议,本研究内容仅使用患者既往住院病历资料,不会对患者造成额外伤害和风险。部分数据目录如表2所示。本研究获得梅州市人民医院医学伦理委员会审批通过,批准号:梅市伦审2022-C-132。

表 2 数据来源
Table 2 Data sources

数据来源	采集数据
医院信息系统	患者的基本信息,例如住院号、性别、出生日期、血型、过敏史等。
电子病历系统	电子病历文书信息,例如病案首页、入院记录、病程记录、出院记录等。
实验室信息系统	实验室检验指标,例如白细胞、红细胞、血红蛋白等。
影像通讯与归档管理系统	检查报告信息,例如检查所见、诊断意见等。
手术信息系统	手术麻醉信息,例如手术与操作编码、麻醉方式、麻醉分级等。

1.4.3 后结构化引擎

本模块通过NLP技术从文书中提取出结构化信息,为后续构建病历质控引擎提供数据基础。首先根据不同的文书类型(病案首页、入院记录、病程记录等),利用分词、实体识别、句法分析等NLP技术对病历文本内容进行结构化处理,并通过实体链接技术将识别出的疾病知识图谱相互关联。形成包含患者病情信息、诊治过程等在内的结构化数据。

医学知识图谱主要以本体表示法以网络的形式表示知识,即以(实体1、关系、实体2)三元组来表示相关联的两个实体。医学知识主要从非结构化数据中人工或自动地提取实体、关系和属性。人工提取是通过专家依据一定规则收集并整理相关信息提取知识,包括国际疾病分类第九版临床修订本手术与操作ICD-9-CM-3(医学版)、国际疾病分类ICD-10(医学版)、国际疾病诊断编码库ICD-11-MMS、MeSH、人类与兽类医学系统术语(中文SNOMED3.4版)、观测指标标识符逻辑命名与编码系统

(LOINC)等;自动提取则是利用机器学习、人工智能、数据挖掘等信息抽取技术,从本院近五年电子病历抽取数据进行训练,得到的医学知识图谱经专家审核后形成更具有本地特色医学知识库。以“高血压”诊断为例,其医学知识图谱表示为(高血压,症状,胸闷)、(高血压,鉴别诊断疾病,短暂性脑缺血发作)、(高血压,检查,超声心动图)等。

1.4.4 病历质控引擎

病历质控引擎是病历质控系统的核心模块,通过可供灵活配置的规则引擎对结构化后的病历信息进行质量问题判定,更进一步通过机器学习方法挖掘文书中存在的完整性、时限性、逻辑性、合规性等问题。根据贯穿整个疾病诊疗环节的质控规则库构建病历质控引擎,确保病历在形式及内涵上均能得到有效监控。病历质控引擎一共包含2个模块,分别是浅层质量规则引擎模块、深层质量问题挖掘模块,具体如下:

浅层质量规则引擎模块。在浅层质量规则引擎模块中,使用从后结构化处理中得到的结构化数据,进行浅层质量问题的识别。根据国家统一标准规范制定的质控问题分类标准,可以抽象为四个维度,即完整性、时限性、逻辑性、合规性。浅层质量规则引擎模块主要完成对病历书写时限、必写项目、书写频次等初级质控规则的监控,已完成在“完整性、时限性”方面的部分校验,实现了一定程度上的自动质控功能。

深层质量问题挖掘模块。浅层质量规则引擎实现对病历中的完整性、时限性等质量问题的监控和识别,但未能很好地实现对逻辑性、合规性等内涵质量方面的核查。因为其关注点主要在文本内容本身,问题更隐蔽也更复杂,必须经过深度的数据挖掘及逻辑判断才能实现。大数据及人工智能技术的发展,为此提供了思路和契机。首先,本研究对海量的历史病历信息进行数据预处理,规范数据的同时进行数据清洗。然后,利用NLP技术对数据进行处理,形成结构化数据。接着,本研究结合计算机自动匹配映射与人工映射的方式将数据与医学术

语进行映射。最后,通过机器学习构建预测模型,用于对病历一致性和整合型的监控识别,实现“逻辑性、合规性”方面的校验,为病历的全方位自动化智能质控填上了最后两块拼图。

1.5 系统实施效果评价

以随机抽查的2022年1月的2918份出院病历为例,质控范围是患者此次住院所有归档病历文书,选取5个质控点为基准进行人工质控和人工智能质控的比较:①现病史雷同;②主诉描述缺陷;③初步诊断不全;④月经婚育史缺失;⑤主诉现病史不匹配。

$$\text{正确率} = \frac{\text{复核正确的问题数}}{\text{人工智能检出问题数}} \times 100\% \quad (1)$$

同时,根据专家复核正确的人工智能质控记录,对比抽样病历的既往人工质控记录,计算既往人工质控检出而人工智能质控未检出的问题数(ND)。由于存在假阳性,因此分析以复核正确的问题数作为金标准,对5个质控点进行受试者工作特征(ROC)曲线分析和配对t检验分析。正实际状态是指灵敏性大于特异性的情况,负实际状态是指特异性大于灵敏性的情况,“结”是分类器的灵敏性和特异性相等情况,出现“结”则意味着分类器在该阈值处无法区分正实际状态和负实际状态,其预测能力较弱。

2 结果

如表3所示。人工智能质控系统对于病历质量问题检出的正确率为89.57%,抽样病历既往人工质控检出问题中仅有1个在人工智能质控系统中未检出,但是人工智能质控系统正确检出质量问题数是人工质控检出问题数的2.97倍。对比人工质控,人工智能质控具有快速高效优势:5个质控点的质控,人工质控动用病案科4位高年资技师,持续30 d完成,而人工智能质控仅用时5 s;在文书覆盖方面,人工智能质控能做到100%覆盖,而人工质控只能抽查,存在遗漏问题。

表 3 抽样病历质控结果对比

Table 3 Comparison of quality control results of sampled medical records

质控点	人工智能质控检出问题数	复核正确的问题数	人工智能质控的正确率	人工质控检出问题数	ND
现病史雷同	2048	1720	83.98%	1016	0
主诉描述缺陷	968	902	93.18%	138	0
初步诊断不全	647	612	94.59%	97	0
月经婚育史缺失	624	601	96.31%	57	0
主诉现病史不匹配	365	332	90.96%	96	1
合计	4652	4167	89.57%	1404	1

ND: 既往人工质控检出而人工智能质控未检出的问题数。

ROC曲线分析结果如图2、表4所示。根据AUC值可以看出,人工智能质控组的5个质控点均有统计学意义($P<0.05$),且AUC值均接近或大于0.9,而人工质控组仅“现病史雷同”质控点有统计学意义($P<0.05$)。组间比较

示,5个质控点的两种质控方式均有差别,人工智能质控组在5个质控点上比人工质控更具有优势。本研究ROC曲线没有与对角线交叉,不存在有实际意义的“结”,对结论不会造成偏差。

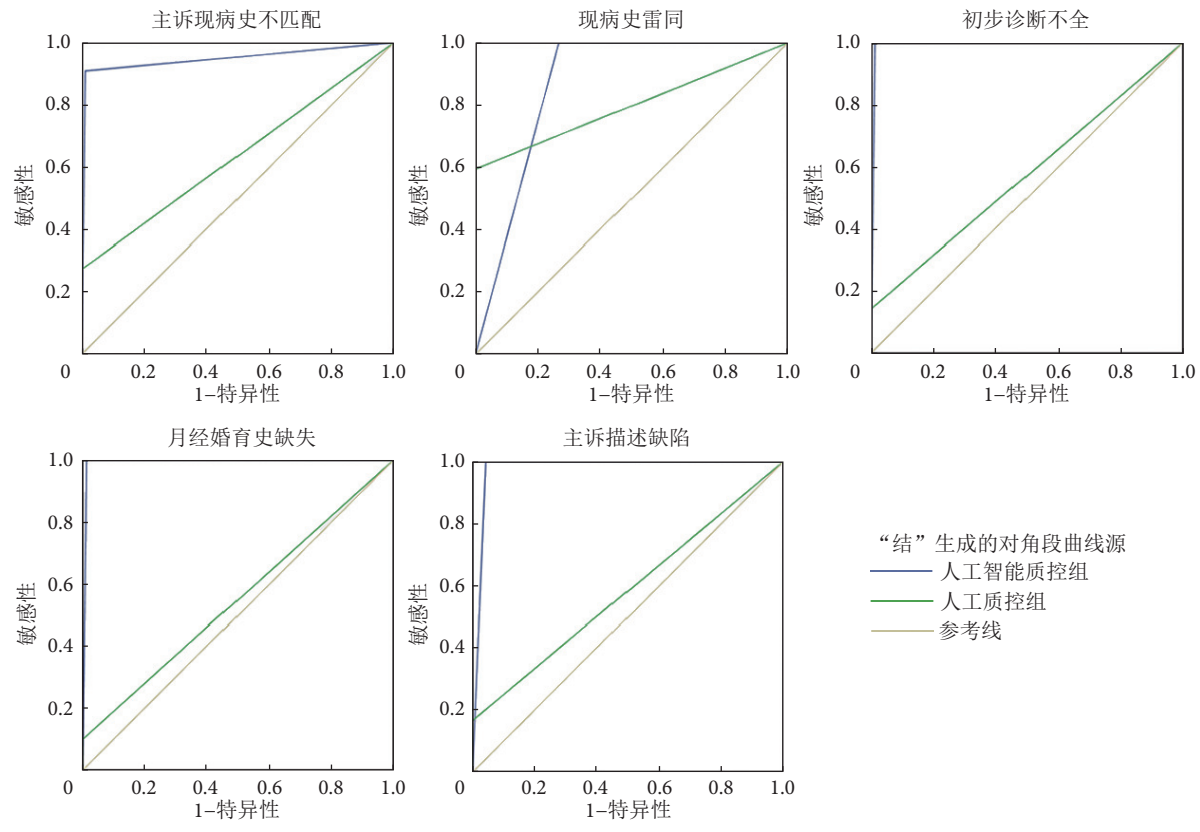


图 2 ROC曲线
Fig 2 ROC curves

表 4 两种质控方式ROC曲线下面积的比较

Table 4 Comparison of the area under the curve of the ROC for the two quality control methods

分组	现病史雷同	主诉描述缺陷	初步诊断不全	月经婚育史缺失	主诉现病史不匹配
人工智能质控	0.866 [*]	0.979 [*]	0.994 [*]	0.994 [*]	0.949 [*]
人工质控	0.797 [*]	0.583	0.571	0.550	0.636
面积差值	0.069	0.396	0.423	0.444	0.313
<i>t</i>	7.301	6.205	4.819	4.975	2.934
<i>P</i>	<0.000 1	<0.000 1	<0.000 1	<0.000 1	0.004

^{*} $P<0.05$, 与ROC曲线下面积=0.5相比。

3 讨论

在采用人工智能技术进行电子病历质控的研究方面,国外在人工智能质控方面相关研究和应用较少,目前国外相关研究主要是针对结构化信息进行分析,而针对非结构化的文本信息分析较少,其主要原因有:①受制于严格的患者隐私保护导致缺乏足够的数

据以开发和测试相关算法;②临床上非结构化文本的细微差使得用标准化的处理技术很难获得可靠的质控结果;③处理电子病历中的自由文本时需要为每个医学数据库甚至是每个临床问题来开发或者改编新的自然语言处理算法,需要耗费大量的人力物力^[6]。

我国相关研究多为质控系统搭建的相关理论介绍以及基于此的质控流程研究,其质控结果能够实现电子病历的形式质控和内涵质控。跟其他医院人工智能质控系统相比,我们所采用技术、算法和模型差异不大,主要区

别在于医学知识库的差异,相比于其他系统采用标准医学知识库形式,本研究在此基础上增加本地跟线上双库运行模式,一方面抽取我院历史数据进行模型训练,另一方面抽取主流杂志和期刊的最新研究成果进行知识点更新,同时在前端推荐治疗方案进行有差异提醒,保证推荐治疗方案既适合院内情况,又能跟上国际最新潮流发展趋势。因此,笔者认为构建适合医院需求知识库是未来人工智能质控发展趋势,同时也是人工智能能否发挥作用的关键。

应用本系统之前,我院依靠质控人员抽查的质控形式,质控覆盖范围小,检出率低。通过本系统的应用,实现了病历文书全量质控,有效提升质控效率,提高检出率,能够发现以往依靠人工不易觉察的问题,有效提高病历文书质量,但本系统仍然存在一些影响质控效果的因素,例如:医学术语词汇量大,术语复杂,存在一词多义的不同表达形式,人工智能难以做出准确判断,容易出现误判问题;异构系统数据结构复杂,形式多样,数据采集和存储各异,导致异构数据结构化处理过程中容易出现误差,影响病历质控准确性;新开展检查、检验项目没有相应质控规则,需人工发现问题后再添加规则算法并训练模型。因此,人工智能质控也需要在人工校准基础上进行优化和学习,针对逻辑性错误、信息不匹配等严重错误,人工校准环节可以进行提醒和告警,促使系统自身进行机器学习优化或者反馈给开发人员进行修改。

随着医学领域新项目、新技术应用,人工智能也要进行算法学习,提倡医疗专家、管理人员、计算机开发和运维人员相结合形式,实时更新质控规则和指控点,同时结合模型训练,才能保证本系统处于最优状态;提倡全员参与质控,发现问题及时提出并弥补缺陷,相比只有终末质控,通过环节质控和终末质控相结合的模式,在医生工作站收到提醒,及时发现并解决缺陷问题,可以减轻终末质控的压力,为提升医院病历文书质量提供可靠的保障,实现最大价值。

* * *

作者贡献声明 温煜负责论文构思、数据编审、调查研究、研究方法、研究项目管理、初稿写作和审读与编辑写作,李雄负责数据编审、正式分析、调查研究、验证、初稿写作和审读与编辑写作,曾菲菲负责正式分析、调查研究、验证、初稿写作和审读与编辑写作,雷佳雨负责数据编审、可视化、初稿写作和审读与编辑写作,陈绍勇负责可视化和初稿写作。所有作者已经同意将文章提交给本刊,且对将要发表的版本进行最终定稿,并同意对工作的所有方面负责。

利益冲突 所有作者均声明不存在利益冲突

参 考 文 献

- [1] 左海霞. 智能化病案质量控制系统的构建与应用. *中国病案*, 2021, 22(8): 40-43. doi: 10.3969/j.issn.1672-2566.2021.08.015.
- [2] 胡敏, 茅建华, 李永昌, 等. 依托信息化手段实现有效的病历三级质控. *中国数字医学*, 2017, 12(12): 108-109. doi: 10.3969/j.issn.1673-7571.2017.12.036.
- [3] 阮彤, 高炬, 冯东雷, 等. 基于电子病历的临床医疗大数据挖掘流程与方法. *大数据*, 2017, 3: 83-98. doi: 10.11959/j.issn.2096-0271.2017054.
- [4] 胡海洋, 赵从朴, 马琰, 等. 基于膨胀卷积神经网络的中文医疗命名实体识别研究. *医学信息学杂志*, 2021, 42(9): 6. doi: 10.3969/j.issn.1673-6036.2021.09.008.
- [5] MAJID A, DMITRIY D, BRIHAT S, *et al.* Development and application of a high throughput natural language processing architecture to convert all clinical documents in a clinical data warehouse into standardized medical vocabularies. *J Am Med Inform Assoc*, 2019, 26(11): 1364-1369. doi: 10.1093/jamia/ocz068.
- [6] 侯李萍. 基于人工智能的电子病历质量控制研究. 武汉: 华中科技大学, 2021.

(2023-06-15收稿, 2023-09-01修回)

编辑 吕 熙



开放获取 本文遵循知识共享署名—非商业性使用

4.0国际许可协议(CC BY-NC 4.0), 允许第三方对本刊发表

的论文自由共享(即在任何媒介以任何形式复制、发行原文)、演绎(即修改、转换或以原文为基础进行创作), 必须给出适当的署名, 提供指向本文许可协议的链接, 同时标明是否对原文作了修改; 不得将本文用于商业目的。

CC BY-NC 4.0许可协议访问<https://creativecommons.org/licenses/by-nc/4.0/>。

© 2023《四川大学学报(医学版)》编辑部 版权所有