

基于 LDA 主题分析模型的突发重大传染病事件 医院抗逆力建设分析——以上海市为例

陈汝婕,王毅欣,刘晶晶,桂莉
(海军军医大学 护理系,上海 200433)

【摘要】 目的 了解医院应对突发重大传染病事件的抗逆力建设,为今后提高其应对能力提供指导。方法 设计聚焦网络爬虫获取国家卫健委、中国新闻网、今日头条和搜狗微信网站内有关上海市医院在 2020 年 1 月 1 日至 2022 年 3 月 1 日间的文本数据,编写 python 程序对纳排后的 179 条数据进行隐含狄利克雷分布(latent Dirichlet allocation,LDA)主题分析。结果 确定 LDA 主题分析的最佳主题个数为 12 个,并根据各主题的前 10 位主题词为主题命名。结论 医院需加强科技、科研以及管理等方面的建设,来提升自身对疫情防控能力,对各类资源的应急准备能力,从而确保医疗服务的稳定提供。

【关键词】 突发重大传染病事件;医院;抗逆力;网络数据;LDA 主题分析

doi:10.3969/j.issn.2097-1826.2023.04.015

【中图分类号】 R473.51 **【文献标识码】** A **【文章编号】** 2097-1826(2023)04-0060-04

Latent Dirichlet Allocation Topic Analysis in Hospital Resilience Construction to Infectious Disease Pandemic: A Case Study of Shanghai

CHEN Rujie, WANG Yixin, LIU Jingjing, GUI Li (School of Nursing, Naval Medical University, Shanghai 200433, China)

Corresponding author: GUI Li, Tel:021-81871481

【Abstract】 Objective To understand the construction of the hospital's resilience to emergencies of major infectious diseases, and to provide guidance for improving its response capacity in the future. **Methods** The web crawler was designed to obtain text data of Shanghai hospitals from the National Health Commission, China News, Toutiao and Sogou wechat from January 1, 2020 to March 1, 2022. The python program was used to analyze 179 items by Latent Dirichlet Allocation(LDA) thematic analysis. **Results** The number of optimal topics for LDA topic analysis is determined to be 12, and the topics are named according to the first 10 topic words of each topic. **Conclusions** Hospitals need to strengthen the construction of science and technology, scientific research and management, so as to improve their own "prevention" of the epidemic, "preparedness" in terms of various resources and capabilities, and the "stability" of medical work.

【Key words】 infectious diseases pandemic; hospital; resilience; web data; Latent Dirichlet Allocation thematic analysis

[Mil Nurs, 2023, 40(04): 60-62, 74]

近年来突发重大传染病事件在全球范围内持续发酵^[1],给国家和人民造成巨大的冲击。医院作为直接承担医疗救治工作的组织,必须在突发重大传染病事件中变得更加游刃有余。目前抗逆力概念被广泛应用于各类系统应对不同类型灾害的研究中,Zhong 等^[2]将医院抗逆力定义为通过提供附加服务(如现场救援、院前护理、紧急治疗、重症监护、去污和隔离)抵御、吸收和应对灾害后果,然后恢复或适

应新情况的能力。医院需要提升其应对突发重大传染病事件的抗逆力,使其能够科学有效地预防或化解突发重大传染病事件造成的冲击或从中恢复过来。隐含狄利克雷分布(latent Dirichlet allocation, LDA)主题分析^[3]是通过机器学习的方式将给定数据中的词语聚类为自定义数量的主题团体的数据分析方法,广泛应用于新闻、文献等各类文本数据的分析和挖掘。在现今多媒体时代,突发重大传染病事件相关的信息通过互联网发布于各类网络平台。因此,本研究基于 LDA 主题分析模型,聚焦上海市医院在疫情爆发前的应对举措,总结突发重大传染病事件医院抗逆力建设的要点。

【收稿日期】 2022-12-04 **【修回日期】** 2023-02-16

【基金项目】 国家自然科学基金项目(72174205)

【作者简介】 陈汝婕,硕士在读,护士,电话:021-81871481

【通信作者】 桂莉,电话:021-81871481

1 资料与方法

1.1 资料来源 通过八爪鱼采集器设计网络爬虫。因社区或卫生管理部门等组织机构会对医院抗逆力产生影响^[4-5],以检索词“医疗机构”“医院”“发热门诊”“急诊”“诊疗”“社区”“街道”“卫健委”“疾控中心”结合“上海疫情”在国家卫生健康委员会、中国新闻网、今日头条和搜狗微信网络平台采集文本数据。共检索到数据 12 003 条,去重后剩余数据 5525 条。以 2020 年 1 月 1 日至 2022 年 3 月 1 日为限。文献^[6]研究发现,医护人员、医院基础设施、管理以及后勤组织工作是影响医院抗逆力的主要因素。纳入标准:(1)医院运作情况;(2)医院发布的信息;(3)医护应对突发重大传染病事件的工作经历;(4)医院与卫生管理机构或社区等交互的数据。排除标准:(1)医护人员的个人事迹、美好品质等宣传资料;(2)党风廉政建设的资料。最终筛选出 179 条数据进行 LDA 主题分析。

1.2 研究方法

1.2.1 数据预处理 本研究通过计算机语言 python 编写程序完成对纳入的数据的正则清洗、去停用词以及分词处理,并创建自定义词表提高分词的准确性^[7]。程序通过正则表达式“[\u4e00-\u9fa5A-Za-z]+”配数据中的中英文字词;引用哈尔滨工业大学停用词表、百度停用词表和四川大学机器智能实验室停用词库去除数据中如“主张”“举行”等无实意的虚词;导入 jieba 第三方中文分词词库并结合自定义词表(例如定义新词“智慧医院”使其不会被分割为“智慧”和“医院”)将数据切分为词语或短语。

1.2.2 LDA 主题分析 本研究通过在 python 程序中导入 gensim.models 的 LdaModel 函数构建 LDA 主题分析模型(迭代次数为 50),并通过 Coherence-Model 函数计算主题一致性确定最佳主题个数(一

致性越高模型拟合的效果越好)。LDA 主题分析的可视化结果通过在 python 程序中导入 pyLDAvis 模块实现,主题以气泡的形式呈现,各个气泡之间的距离代表主题的相关程度,气泡的直径大小代表在文件集中该主题的比例。

2 结果

2.1 LDA 主题模型一致性测试 本研究测试了数据 LDA 主题模型在 1~100 个主题下的一致性大小,结果显示当主题数为 12 时,模型的拟合效果最好。

2.2 主题可视化 本研究根据 LDA 主题模型,12 个主题下的前 10 位主题词总结医院在突发重大传染病事件尚未大范围爆发时期的应对举措,并以此作为主题名称为各主题命名,可视化结果见图 1,具体信息详见表 1。

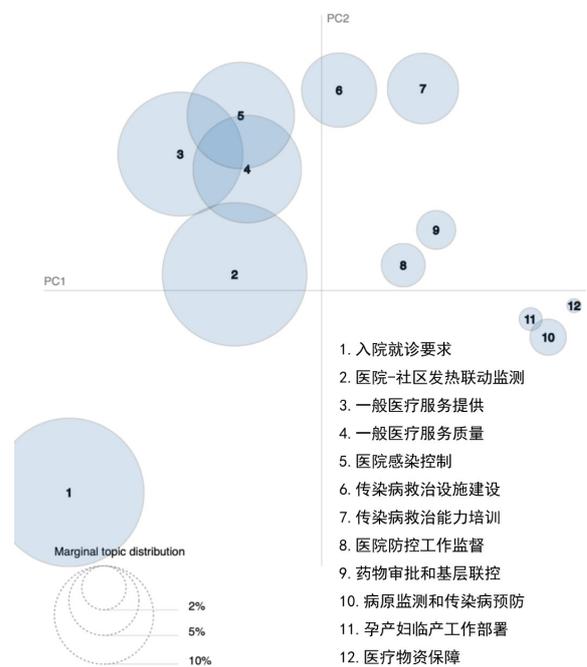


图 1 LDA 主题模型可视化结果

表 1 LDA 主题分析模型主题名称、前 10 位主题词、主题占比情况

主题编号	主题名称	前 10 位主题词	主题占比(%)
1	入院就诊要求	预约、就诊、医院、患者、人员、门诊、陪护、配合、行程码、核酸检测	22.8
2	医院-社区发热联动监测	患者、发热、医院、疫情、哨点、发热门诊、防护、工作、诊室、感染	21.2
3	一般医疗服务提供	发热门诊、医院、患者、上海、发热、疫情、就诊、建设、医疗、新冠肺炎	15.9
4	一般医疗服务质量	医院、医疗、服务、上海市、上海、病人、患者、管理、医疗机构、疫情	12.0
5	医院感染控制	感染、医院、管理、工作、防控、医疗机构、疫情、消毒、疫情防控、相关	11.7
6	传染病救治设施建设	发热门诊、建设、医院、疫情、公共卫生、上海、医护人员、传染病、能力、上海市	5.7
7	传染病救治能力培训	机构、医疗卫生、公共卫生、人员、培训、专业、控制、工作、卫生、应急	5.1
8	医院防控工作监督	医疗机构、防控、新闻、卫生、健康、疫情防控、落实、服务、医疗、发布会	2.0
9	药物审批和基层联控	诊室、医疗机构、发热、设置、疫情防控、制剂、机构、基层、药品、传染病	1.5
10	病原监测和传染病预防	接种、疫苗、老年人、变异株、病毒、人群、新冠病毒、突变、奥密克戎、我国	1.3
11	孕产妇临产工作部署	疫情、分娩、产科、孕产妇、医院、建卡、产房、孕妇、下降	0.5
12	医疗物资保障	上海、医药、药品、物资、生产、累计、小时、供应、保障、采购	0.2

3 讨论

3.1 推动院内外联动疫情预警防线建设 由结果可见,主题1入院就诊要求和主题2医院-社区发热联动监测共同提示医院对建立发热预警防线的重视。重大传染病爆发以来,诸多医院通过流行病学调查、出入口管理、限制陪护等公共卫生应对措施主动筛查病例并减少人员之间的交叉感染^[8]。除医院对传染病病例的排查工作外,主题5指向的医院感染控制工作是保障其他诊疗服务得以持续开展的基础,为防止医院感染事件的发生,医院必须推动感染防控工作的不断改进,并促进医院医务工作者对感染防控工作的重视和落实^[9]。其次,主题10中“新冠病毒”“变异株”和“奥密克戎”等主题词提示医院应当开展对传染病病原的监测工作。医院有必要加强获取传染病病原信息的能力,了解病原的传播速度、致病特点等情况,从而科学地制订院内防控和救治方案并快速转换医院的“平战模式”^[10]。因此,医院有必要不断织密疫情预警防线,并在此基础上探索与社区、药店或其他医疗机构之间的联动防控模式^[11],保障在重大传染病事件突然爆发时能够迅速有序地开展群众的传染病排查和分流工作。疫情的筛查和预警更多依赖于护理人员,护理人员需要具备对传染病的敏感性,通过上报、隔离管控等方式及时遏制感染在社区或者医院内的扩散。

3.2 保证医院医疗救治服务的延续 突发重大传染病事件期间,医疗服务的延续是群体健康的保障。主题3除结果中呈现的主题词外还包括“互联网医院”“智慧医院”等特殊主题词。互联网医院在突发重大传染病事件中为疫情初期发热患者咨询和后期慢性病患者复诊配药等方面提供了重要的支持。医院有必要强化信息化技术的应用,以减少人员交叉感染并保证部分医疗服务的提供^[12]。其次,主题6显示的实体医院建设更是医疗服务延续的关键一环。医院在环境、空间、床位、救治检查设备等基础设施的设计与配备上需要考虑传染病的特殊性以及疫情爆发造成医疗挤兑的可能性^[13]。此外,主题12医疗物资保障提示医院的一般药物以及医疗、防护物资储备是开展一切医疗工作的基础,医院需要保证此类耗材的供给和储备情况能够支持传染病疫情期间的日常消耗和疫情爆发时激增用量^[14]。更重要的是,主题9新冠肺炎临床治疗药物审批指向医院在突发重大传染病事件应对期有必要拓宽例如药物使用权限开放、新型治疗技术引进等传染病救治条件,推动医院在病例救治和治疗方案选择上更具主动性。在突发重大传染病事件中,医疗服务的延续对护理人员的学习能力和工作效率提出更高要求,

护理人员必须及时适应例如互联网诊疗的模式,快速掌握新的治疗技术并应用,同时能够服务于多于平时数倍的传染病患者。

3.3 提升医院医疗救治工作的质量 医院在保障疫情期间的医疗服务的同时,还需要保证所提供的医疗服务符合疫情防控的标准。主题4突出体现了医院对医疗服务质量的关注。在疫情常态化防控阶段,许多专科诊疗服务的专家逐渐形成共识以指导并规范疫情期间的临床操作^[15]。主题7可见上海市医院及时对医护工作者开展了传染病专科培训以加强医院在传染病救治以及疫情防控方面的软实力。医护人员能力和知识方面的储备将有助于医院应对重大传染病事件的突然爆发^[16]。另外,主题8中的主题词显示医院在对疫情防控、医疗卫生服务等方面执行督导的重要性。突发重大传染病事件期间,通过督导的方式使医院具备院内感染防控意识并科学执行规范,及时发现漏洞、整改问题,降低感染风险^[17]。因此,医院有必要强化在规章制度、专业能力和质量控制方面的建设,提升医疗服务的规范性和科学性,从而在突发重大传染病事件爆发时保证医患的安全和医疗服务的质量。传染病患者的许多救治工作都需要护理人员高频次执行,护理需要在突发重大传染病事件中形成独立的工作规范或指导内容来保证护理工作的质量。

3.4 识别特殊人群医疗需求促进救治 不同的传染病病原有不同的易感人群,主题10提及新冠疫情中老年人接种疫苗的问题。目前研究发现新冠病毒对伴有基础疾病的老年人的致病情况较为严重,医院应当积极开展特定传染病疫情相应易感人群救治方案的探索和改进。其次,医院应当对老年人接种传染病疫苗进行宣教,医护人员的健康教育和疫苗接种行为也能促进群体对传染病疫苗接受^[18],从而降低发展为危重症病例的可能,为可能发生的疫情爆发缓解医院的救治压力。此外,主题11反映疫情期间孕产妇临产或救治等紧急情况需要妥善处置的同时,也折射出医院在救治所有急危重症患者工作上的不足。突发重大传染病事件期间,医院需要具备为急危重症患者进行传染病筛查和救治的能力。例如通过人才队伍的组建、急诊各项设施的建设、流程标准的规范和救治方案的制订等^[19],优化医院对急危重症患者的接收和救治工作。护理应重视对民众的健康教育工作,从而使各类人员明确或避免自身在突发重大传染病事件中可能遇到的需要就医的风险事件,在院前阶段保护他们的健康和安。

(下转第74页)

型的测量不变性[J].现代预防医学,2021,48(9):1600-1603,1608.

[19] NEWTON C R, SHERRARD W, GLAVAC I. The fertility problem inventory: measuring perceived infertility-related stress[J]. *Fertil Steril*, 1999, 72(1): 54-62.

[20] PENG T, COATES R, MERRIMAN G, et al. Testing the psychometric properties of Mandarin version of the fertility problem inventory (M-FPI) in an infertile Chinese sample[J]. *J Psychosom Obstet Gynaecol*, 2011, 32(4): 173-181.

[21] ZURLO M C, CATTANEO DELLA VOLTA M F, VALLONE F. Factor structure and psychometric properties of the fertility problem inventory-short form [J/OL]. [2022-07-10]. <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5779934/>. DOI: 10.1177/2055102917738657.

[22] LI C C, HUANG S M, LAI J C, et al. Development and validation of a fertility intention scale in breast cancer survivors[J]. *J Nurs Res*, 2018, 26(3): 177-184.

[23] HUANG S M, TSENG L M, LAI J C, et al. Impact of symptom and social support on fertility intention in reproductive-age women with breast cancer[J]. *Clin Nurs Res*, 2020, 29(6): 411-

418.

[24] GERSTL B, SIGNORELLI C, WAKEFIELD C E, et al. Feasibility, acceptability and appropriateness of a reproductive patient reported outcome measure for cancer survivors[J/OL]. [2022-07-10]. <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC8396734/>. DOI: 10.1371/journal.pone.0256497.

[25] HOPWOOD P, FLETCHER I, LEE A, et al. A body image scale for use with cancer patients[J]. *Eur J Cancer*, 2001, 37(2): 189-197.

[26] OBERGUGGENBERGER A S, NAGELE E, INWALD E C, et al. Phase 1-3 of the cross-cultural development of an EORTC questionnaire for the assessment of sexual health in cancer patients; the EORTC SHQ-22[J]. *Cancer Med*, 2018, 7(3): 635-645.

[27] MITCHELL A J, BAKER-GLENN E A, GRANGER L, et al. Can the distress thermometer be improved by additional mood domains? Part I. Initial validation of the emotion thermometers tool[J]. *Psychooncology*, 2010, 19(2): 125-133.

(本文编辑:王园园)

(上接第 62 页)

【参考文献】

[1] 韩辉,伍波,吴海磊,等.2022年8月全球传染病疫情概要[J].*疾病监测*,2022,37(9):1140-1142.

[2] ZHONG S, HOU X Y, CLARK M, et al. Disaster resilience in tertiary hospitals: a cross-sectional survey in Shandong province, China[J/OL]. [2022-10-28]. <https://bmchealthservres.biomedcentral.com/articles/10.1186/1472-6963-14-135>. DOI: 10.1186/1472-6963-14-135.

[3] 叶艳,吴鹏.循证决策视角下的患者健康咨询主题分析[J].*情报理论与实践*,2022,45(2):198-203.

[4] BARKER K M, LING E J, FALLAH M, et al. Community engagement for health system resilience: evidence from Liberia's Ebola epidemic[J]. *Health Policy Plan*, 2020, 35(4): 416-423.

[5] SMAGGUS A, LONG J C, ELLIS L A, et al. Government actions and their relation to resilience in healthcare during the covid-19 pandemic in New South Wales, Australia and Ontario, Canada [J]. *Int J Health Policy Manag*, 2021, 11(9): 1682-1694.

[6] KHADEMI J A, AHMADI K R, HEYRANI A. Factors influencing hospital resilience[J]. *Disaster Med Public Health Prep*, 2021, 15(5): 661-668.

[7] 黄仁东,韩静,崔乃雪.公众对“网约护士”的态度和关注点:基于微博评论的文本挖掘研究[J].*护理研究*,2021,35(17):3021-3025.

[8] 付陈超,吴安华,黄勋.新冠肺炎疫情医院内流行病学调查及思考[J].*中国感染控制杂志*,2022,21(8):723-728.

[9] 侯冷晨,王鹏,钱香玲,等.上海市公立医院医院感染管理现状调查[J].*中国感染控制杂志*,2021,20(11):1016-1021.

[10] 周建琿,孟蕾,汪丽娟,等.甘肃省 2009—2019 年发热伴出疹症

候群病原分布特征多重对应分析[J].*中国公共卫生*,2022,38(3):356-359.

[11] 石磊,朱锦生,杨冬林,等.新冠疫情下高风险地区集中救治点防控机制探讨[J].*现代医院*,2022,22(1):137-140.

[12] 杨玲,宋宝香.互联网医院患者线上线下转化的制约因素及对策探析[J].*卫生经济研究*,2022,39(11):65-68.

[13] 钟正东,廖芃芃,吴文琪,等.突发公共卫生事件下医疗卫生机构基础设施建设调查:基于武汉市数据[J].*中华医院管理杂志*,2020,36(11):886-890.

[14] 史艳茹,闫沛,胡雪慧,等.新冠肺炎疫情防控期间医院防护物资管理实践探讨[J].*中国卫生质量管理*,2020,27(6):59-61.

[15] 国家消化内镜专业质控中心,中国医师协会内镜医师分会,中华医学会消化内镜学分会.新型冠状病毒肺炎疫情防控期间 ERCP 诊疗专家共识[J].*中华消化内镜杂志*,2020,37(9):609-613.

[16] 孙梦圆,杨艳,赵勋,等.护理人员突发公共卫生事件应对能力的培训管理现状与思考[J].*解放军护理杂志*,2021,38(8):78-80,84.

[17] 黄青菁,张瑾,秦赛君.医院纪检监察在疫情防控工作中的监督保障作用探讨[J].*中医药管理杂志*,2022,30(8):219-221.

[18] BETSCH C, SCHMID P, HEINEMEIER D, et al. Beyond confidence: development of a measure assessing the 5C psychological antecedents of vaccination [J/OL]. [2022-10-28]. <https://journals.plos.org/plosone/article?id=10.1371/journal.pone.0208601>. DOI: 10.1371/journal.pone.0208601.

[19] 何昊,张海南,陶澄,等.新型冠状病毒肺炎疫情防控期间急危重症患者的筛查与救治[J].*中国急救复苏与灾害医学杂志*,2021,16(6):589-591,594.

(本文编辑:王园园)