

我国慢性病医防融合领域文献主题演化 ——基于Word2vec与LDA模型的可视化分析

李 艳¹, 唐 岚², 黄 豪^{3*}

¹上海市浦东新区金杨社区卫生服务中心科教科, 上海

²上海市浦东新区潍坊社区卫生服务中心全科, 上海

³上海市浦东新区金杨社区卫生服务中心院部, 上海

收稿日期: 2025年2月11日; 录用日期: 2025年3月4日; 发布日期: 2025年3月12日

摘 要

为了深入研究我国慢性病医防融合领域的发展趋势和演化过程, 本文收集了2006~2024年的373篇相关文献, 经过数据清洗和预处理后, 引入Word2vec的LDA模型进行文献的主题挖掘, 确定每个时期的最佳主题数量, 并生成主题演化桑基图。计算不同时间段内各主题强度, 并通过交互式条形图描述热点主题。结果显示, 在第一阶段2006~2020年, 大部分研究主要集中在如何整合医疗服务, 以及如何将慢性病防控与医防结合; 在第二阶段2021~2022年, 除了延续既有的主题, 部分研究焦点转移到如何更好地管理和融合综合医疗服务, 以及如何将公共卫生服务与医疗体系更有效地结合; 在第三阶段2023~2024年, 研究重点在于如何实现健康服务与医防的深度融合, 以及如何在医疗服务中具体落实医防融合的理念, 研究更加注重实际操作和具体应用。通过主题演化分析揭示了不同时期内主题之间的关联和演化过程, 综合医疗服务、慢性病防控与医防结合等主题在不同阶段都有较强的延续性, 而研究重点随着时间的推移逐渐从综合医疗服务向医防融合和健康服务管理方向转移。研究发现, 一些主题在不同时期内保持较高的强度, 从本研究主题强度图可以看出, 在慢性病医防融合领域, 社区基层医疗机构在医防融合中具有重要作用, 此外2021年及以后的阶段中公共卫生体系建设及医防融合成为研究的共识热点。该研究有助于更全面地理解慢性病医防融合领域的研究动态, 为未来的研究方向和政策制定提供有益的参考, 同时也为文本分析方法的应用提供了实践示范。未来的研究可以进一步挖掘基层医疗与医防协同机制以及健康服务管理与慢性病防控方面的潜力, 更好地帮助社区基层医疗机构服务提供者应对来自人口老龄化社会慢性病高发以及多样化健康需求的挑战, 同时也要关注对应的新兴技术如人工智能和大数据分析和对应的数据隐私和伦理挑战, 以及政策实施中的风险。

关键词

文本挖掘, 主题识别, LDA, Word2vec, 慢性病, 医防融合

*通讯作者。

Thematic Evolution of Literature in the Field of Chronic Disease Medical Prevention Integration in China

—Visualization and Analysis Based on Word2vec and LDA Models

Yan Li¹, Lan Tang², Hao Huang^{3*}

¹Department of Science and Education, Shanghai Pudong New District Jin Yang Community Health Service Center, Shanghai

²Department of Family Medicine, Shanghai Pudong New District Weifang Community Health Service Center, Shanghai

³Hospital Department, Shanghai Pudong New District Jin Yang Community Health Service Center, Shanghai

Received: Feb. 11th, 2025; accepted: Mar. 4th, 2025; published: Mar. 12th, 2025

Abstract

In this paper, in order to deeply study the development trend and evolution process in the field of chronic disease medical preventive integration in China, 373 relevant literatures from 2006~2024 were collected, and after data cleaning and pre-processing, the LDA model of Word2vec was introduced in the theme mining of the literature to determine the optimal number of themes in each period and generate the theme evolution Sankey diagram. The intensity of each topic in different time periods is calculated and hot topics are described by interactive bar charts. The results show that in the first period of 2006~2020, most of the studies focused on how to integrate healthcare services and how to combine chronic disease prevention and control with medical prevention; in the second period of 2021~2022, in addition to the continuation of the existing themes, some of the studies shifted their focus to how to better manage and integrate integrated healthcare services and how to combine public health services with the healthcare system more effectively; in the third stage, 2023~2024, the research focused on how to realize the deep integration of health services and medical preventive, and how to implement the concept of medical prevention integration in health care services, and the research focused more on practical operation and specific application. The analysis of theme evolution reveals the connection and evolution process between themes in different periods. The themes of comprehensive medical service, chronic disease prevention and control and medical prevention integration have strong continuity in different stages, while the focus of research gradually shifts from comprehensive medical service to medical prevention integration and health service management over time. It is found that some themes maintain a high intensity in different periods, and the intensity map of the themes in this study shows that in the field of chronic disease medical prevention integration, community-based primary healthcare organizations have an important role in medical prevention integration, and in addition, public health system construction and medical prevention integration have become consensus hotspots in research in the stage of 2021 and beyond. This study contributes to a more comprehensive understanding of the research dynamics in the field of chronic disease medical prevention integration, provides useful references for future research directions and policy formulation, and also provides a practical demonstration of the application of text analysis methods. Future research can further explore the potential of primary care and medical prevention synergistic mechanisms as well as health service management and chronic disease prevention and control to better help community-based primary care providers to cope with

the challenges from the high prevalence of chronic diseases and diversified health needs of an aging population, as well as to pay attention to the corresponding emerging technologies such as artificial intelligence and big data analytics and the corresponding data privacy and ethical challenges, and the risks in policy implementation.

Keywords

Text Mining, Topic Recognition, LDA, Word2vec, Chronic Disease, Medical Prevention Integration

Copyright © 2025 by author(s) and Hans Publishers Inc.

This work is licensed under the Creative Commons Attribution International License (CC BY 4.0).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

1. 引言

慢性病医防融合是中国国民健康规划的重要组成部分，在政策层面国家出台了《“十四五”国民健康规划》和《中国防治慢性病中长期规划(2017~2025)》等重要政策文件，明确了慢性病综合防控和预防干预的战略目标与具体措施，为医防融合提供了顶层设计和政策保障[1][2]，这些政策强调多部门、多领域的协同合作，推动了多元治理模式的形成，通过整合政府、社会组织、医疗卫生机构等多方资源，形成了全社会共同参与的慢性病防控格局[3][4]。同时，研究还强调了政策分析工具的应用，通过系统评估政策实施效果，为政策优化提供了科学依据[5][6]。此外，彩虹模型等理论工具被用于分析医防融合的治理结构，强调不同层级和部门之间的互动与合作，推动了治理体系的完善[7]。我国医防融合领域的实践模式探索方面不断深入，目前呈现出强化基层医疗服务和各地疾病预防控制机构、医院和基层医疗卫生机构的分工协作、优势互补的慢性病医防融合协作机制[8]-[10]。

尽管我国在慢性病医防融合的政策、理论和实践方面取得了显著进展，但现有的研究多集中于定性分析和政策解读，缺乏对慢性病医防融合研究历程及其演变过程的定量分析，基于数据挖掘技术和机器学习算法的研究方法在慢性病医防融合领域的应用尚未开展，因此难以准确把握不同时间段内的研究重点和发展趋势。LDA (Latent Dirichlet Allocation)主题模型和 Word2vec 等模型可以有效地处理大规模文本数据，挖掘出隐藏在数据中的模式和规律[11]-[13]。因此，本研究首次通过引入 LDA 和 Word2vec 模型，全面展示我国慢性病医防融合领域的研究历程及主题演变过程，为慢性病医防融合研究提供了全新的视角和方法。首先，本研究通过 LDA 主题模型进行定量阶段划分，系统梳理并能够从海量文献中识别出潜在的研究主题，并分析慢性病医防融合研究的演变轨迹，识别出不同时间段内的研究重点和发展趋势，解决了传统定性分析难以准确把握研究脉络的问题，为政策制定和实践创新提供了科学依据。其次，本研究结合 Word2vec 模型，通过词向量表示捕捉词语间的语义关系，进一步提升了主题分析的准确性和深度，这种数据挖掘与机器学习技术的应用，使得我们能够更全面地理解慢性病医防融合研究的复杂性和多样性。总体而言，本文以 2006~2024 年慢性病医防融合发表的文献为研究对象，通过 LDA 和 Word2vec 的主题演化分析构建出慢性病医防融合主题挖掘的研究框架，进行依据模型量化分析结果对研究演化过程进行阶段划分，从主题角度系统梳理研究文本主要主题内容的演化情况，相关研究结论有助于研究者、政策制定者和实践者更好地理解当前的研究热点和发展趋势，还为未来的政策制定和实践创新提供了强有力的决策支持。

2. 数据来源与研究方法

2.1. 数据来源

样本数据源自于中国知网(China National Knowledge Infrastructure, CNKI)、万方数据(Wanfangdata)、国家哲学社会科学文献中心数据库和国家卫生健康委员会官方网站, 设定关键词“医防融合”、“医防协同”、“防治结合”、“慢性病管理”、“慢性病”, 检索时间从建库到 2024 年 6 月 30 日, 检索结果获得 373 篇文献, 剔除会议交流等文献, 通过阅读每篇文献的摘要进行内容的进一步确定后纳入 197 篇文献。按照年份将样本数据进行分类, 将总体划分为三个阶段: 第一阶段为 2006~2020 年, 第二阶段为 2021~2022 年, 第三阶段为 2023~2024 年。

2.2. 研究方法

2.2.1. 主题挖掘技术

本文应用 LDA 主题模型和 Word2vec 神经网络模型对我国慢性病医防融合领域文献的主题进行挖掘和演化分析。运用 Python 软件对网络文本数据进行采集、清洗和集成处理, 最终形成本文研究需要的文本资料库, 运用 Python 中的 jieba 工具包对文本语料库进行分析处理, 为了进一步提高分词的精确度, 在分词的过程中增加自定义词表、同义词词表和去停用词表, 采用 LDA 主题模型展开主题分析。LDA 是由文档-主题-词构成的三层贝叶斯概率模型, 通过一篇文章中的每个词以概率的形式抽取某个主题, 再从这个主题中抽取某特征词, 并不断重复这个过程, 对应文档中的每个词可以使用文档-主题分布和主题-词分布来计算器生成概率: P 。通过主题困惑度曲线确定理想主题个数, 提取每个主题的前 30 个高频词汇, 据此进一步归纳主题名称和呈现结果。Word2vec 是一种学习词嵌入的神经网络模型, 通过将词汇映射到高维空间中的向量来捕获词汇之间的语义关系, 将 LDA 和 Word2vec 结合在一起, 可以在文档和单词级别捕获主题和语义信息。

2.2.2. 主题强度分析技术

运用 LDA 模型可得到文档-主题矩阵, 包含每篇文献在每个主题上的概率值, 即该文献在每个主题上的强度, 再结合文献的发表时间, 将每个主题在相同时间窗口的概率值求和并平均即可得到该主题在该时间窗口上的主题强度[14]。

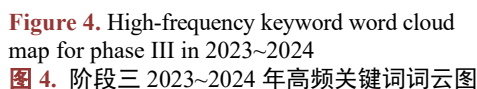
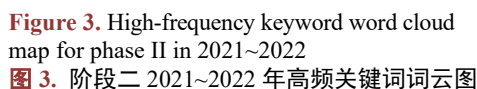
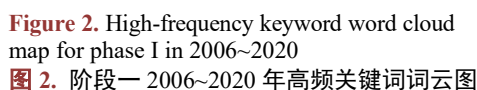
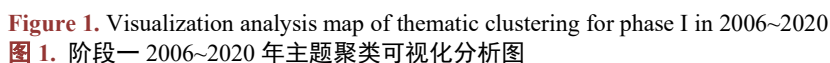
2.2.3. 主题演化分析技术

对不同时间窗口下的文献进行 LDA 主题挖掘, 所获取的主题之间可能存在一定的联系和差异, 为了描述相邻时间窗口下主题之间的相似度及其演化关系, 研究采用余弦值计算相似度来求得主题之间的演化关系, 余弦值的取值范围为 0~1 之间, 数值越接近 1, 表明两个向量越相似, 反之则表明两个向量差异越大[15]。

3. 主题分析

3.1. 主题挖掘分析

应用 LDA 主题建模时需要确定主题数量, 采用困惑度和主题一致性结合的方法来确定最优主题数[16]。一般情况下, 困惑度下降趋势不明显或者处于拐点时取最优主题数, 主题一致性处于最高点时取最优主题数。针对上述三个阶段的文献分别进行 LDA 主题模型挖掘, 展现不同阶段的主题情况及前 30 个关键词进行聚类可视化展示, 并针对每个阶段高频关键词进行词云的展示, 如下图 1~6 所示。



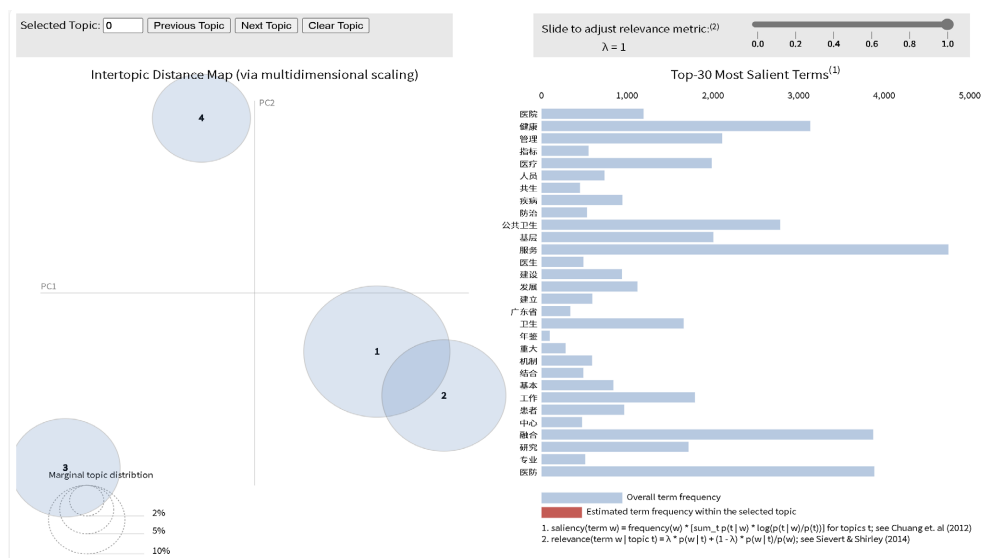


Figure 5. Visualization analysis map of thematic clustering for phase II in 2021~2022

图 5. 阶段二 2021~2022 年主题聚类可视化分析图

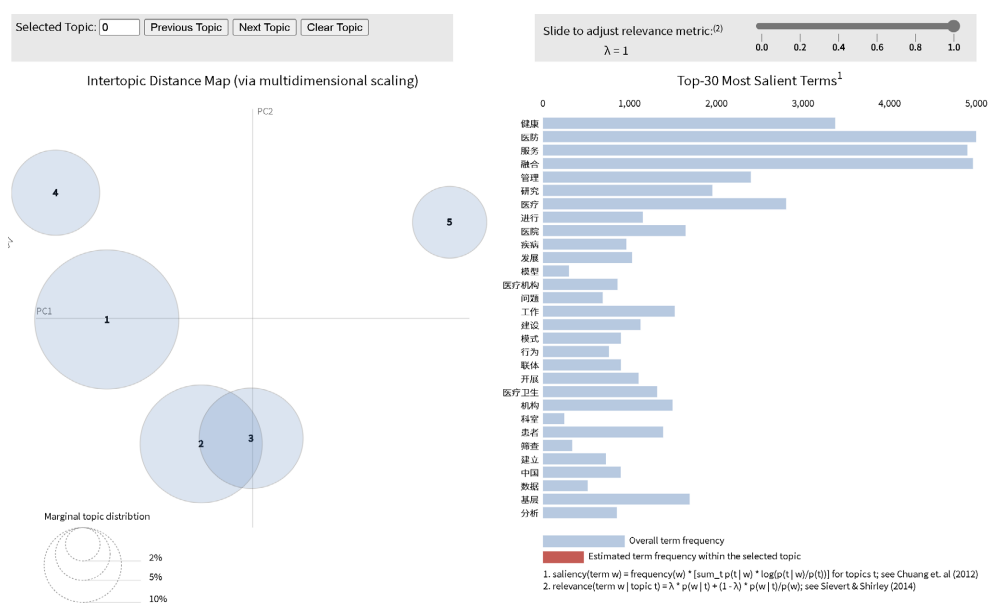


Figure 6. Visualization analysis map of thematic clustering for phase 3 in 2023~2024

图 6. 阶段三 2023~2024 年主题聚类可视化分析图

3.2. 主题强度分析

基于模型输出结果, 分别计算 2006~2020 年、2021~2022 年和 2023~2024 年的主题强度, 绘制主题强度的交互式条形图来可视化主题在不同时间段内的强度变化, 强度较高的主题为热点主题。由图 7 可知, 本次研究三个阶段中主题强度超过 0.3 的有 2021~2022 年的 Topic 4 医防融合与公共卫生、2023~2024 年的 Topic 3 医防融合与公共卫生体系建设。主题强度在 0.2~0.3 之间的有 2020 年及以前的 Topic 1 综合医疗服务、2021~2022 年的 Topic 1 综合医疗服务与管理融合, 2020 年及以前的 Topic 3 社区卫生服务、医疗服务与公共卫生融合和 2021~2022 年的 Topic 3 基层医疗和医防协同机制、2023~2024 年 Topic 2 医

防融合与基层医疗服务。由此看出，在慢性病医防融合领域，社区基层医疗机构在医防融合中具有重要作用，此外 2021 年及以后公共卫生体系建设及医防融合成为研究的共识热点。

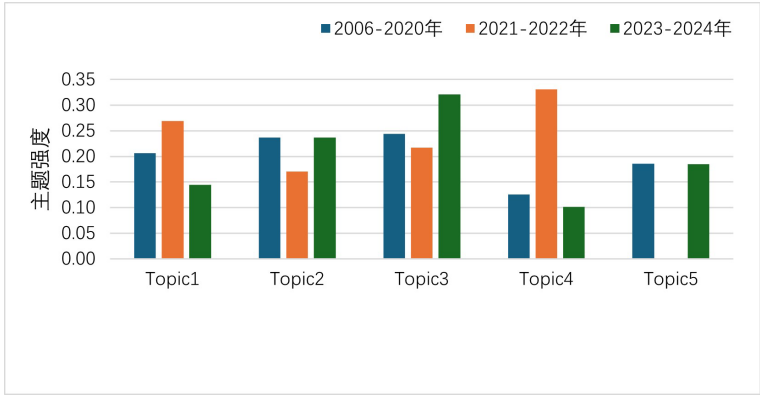


Figure 7. Interactive bar chart of theme intensity
图 7. 主题强度交互式条形图

3.3. 主题演化分析

研究基于 LDA-Word2vec 模型分析了 2006~2020 年、2021~2022 年和 2023~2024 年我国慢性病医防融合领域的主题演化。研究方法上结合三个阶段的主题词词频和主题分类构建了三个时期的演化桑基图 (见图 8)，直观展示主题间的演化路径，每个节点代表一个特定的主题，图中连接线条的粗细展示了主题间的相似度，通过桑基图可以清晰地描述出学术主题随时间的演进，理解三个阶段学者关注的焦点以及研究方向的转变。

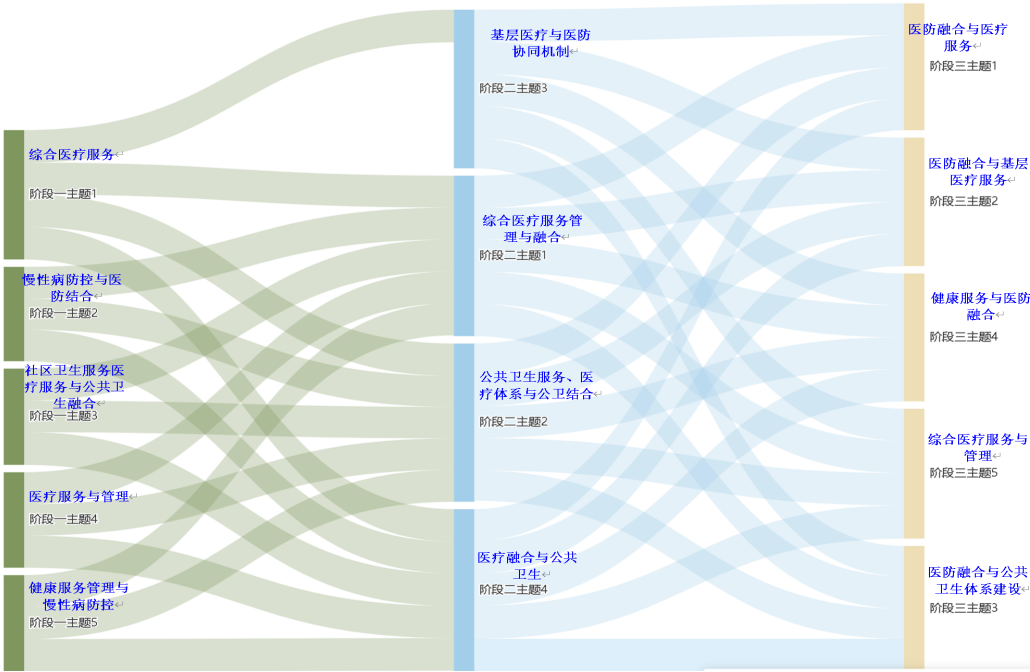


Figure 8. Sankey diagram of LDA theme evolution
图 8. LDA 主题演化桑基图

通过对三个阶段的文献进行 LDA 主题演化桑基图分析,我们可以看出我国慢性病医防融合领域的文献主题演化呈现出延续性、分化、新生的特点。三个阶段中延续性主题包含综合医疗服务的深化与拓展、慢性病防控与医防结合,分化主题包含社区卫生服务医疗与公共卫生融合、医疗服务与管理,而健康管理服务与慢性病防控是新生的主题,这些特点反映了我国在慢性病防控和医疗服务领域逐步深化、系统化的发展方向,同时也展示了不同主题在不同阶段的演变和细化过程。这种趋势有助于我们更好地理解我国慢性病医防融合领域的关键环节和重点领域、发展趋势,为未来的研究和实践提供参考和指导。

两条延续性主题的演变路径中,综合医疗服务(阶段一主题 1)→综合医疗服务管理与融合(阶段二主题 1)→综合医疗服务与管理(阶段三主题 5),这一路径显示了综合医疗服务从初步的整合,到管理与融合的深化,再到服务与管理的进一步发展,表明在慢性病医防融合中,综合医疗服务在不同阶段都有延续性、服务的质量和效率在不断提升。慢性病防控与医防结合(阶段一主题 2)→公共卫生服务、医疗体系与公卫结合(阶段二主题 2)→健康服务与医防融合(阶段三主题 4),该路径展示了慢性病防控从初始的医防结合逐步扩展到公共卫生服务和整体医疗体系的结合,最终实现健康服务与医防融合,这表明慢性病防控越来越注重系统性、整体性的公共卫生策略。

社区卫生服务医疗与公共卫生、医疗服务与管理的主题呈现出分化演变的趋势。其中社区卫生服务医疗与公共卫生融合(阶段一主题 3)→基层医疗与医防协同机制(阶段二主题 3)→医防融合与基层医疗服务(阶段三主题 2),这表明社区卫生服务在不同阶段有分化,初始阶段社区卫生服务医疗与公共卫生融合作为主题出现,这个阶段主要集中在社区卫生服务的初步探索和公共卫生的初步结合,阶段二演化为基层医疗与医防协同机制,这个阶段社区卫生服务开始深入到基层医疗,强调医防协同机制的建立,这意味着社区卫生服务不仅仅是公共卫生的简单结合,而是通过建立协同机制,实现更高效的医疗服务。阶段三进一步演化为医防融合与基层医疗服务,这个阶段社区卫生服务已经深入到基层医疗服务的各个方面,强调医防融合即医疗和预防的深度结合,这反映了社区卫生服务在医防融合中的重要作用,尤其是在基层医疗服务中。医疗服务与管理(阶段一主题 4)→医疗融合与公共卫生(阶段二主题 4)→医防融合与公共卫生体系建设(阶段三主题 3),这个主题初始阶段主要集中在医疗服务的管理层面,探索如何提高服务质量和效率。阶段二演化为医疗融合与公共卫生,强调医疗服务和公共卫生的融合,医疗服务关注治疗,同时还包括预防和健康促进。阶段三进一步演化为医防融合与公共卫生体系建设,此阶段医疗服务与管理已经深入到公共卫生体系的建设,强调医防融合的系统性和整体性,反映了医疗服务与管理在公共卫生体系中的重要性,以及医防融合的系统性建设需求。

健康管理服务与慢性病防控是新生的主题。健康管理服务与慢性病防控(阶段一主题 5)→医疗服务与公共卫生(阶段二主题 5)→医防融合与医疗服务(阶段三主题 1),这个主题展示了健康管理服务与慢性病防控从初始的管理与防控逐步扩展到医疗服务与公共卫生,最终实现医防融合与医疗服务的结合,这表明健康管理服务与慢性病防控在不同阶段有新生的内容和方向,逐步深化和完善。

从桑基图中可以看出,综合医疗服务(阶段一主题 1)、慢性病防控与医防结合(阶段一主题 2)、社区卫生服务医疗与公共卫生融合(阶段一主题 3)、医疗服务与管理(阶段一主题 4)、健康管理服务与慢性病防控(阶段一主题 5)这几个主题在不同阶段都有不同程度的延续性和分布,表明这些内容始终是研究的核心内容。

综合医疗服务(阶段一主题 1)与综合医疗服务管理与融合(阶段二主题 1)、综合医疗服务与管理(阶段三主题 5)是紧密相连的主题。此外,慢性病防控与医防结合(阶段一主题 2)与公共卫生服务、医疗体系与公卫结合(阶段二主题 2)、健康服务与医防融合(阶段三主题 4)之间也有较强的主题紧密联系。

结合主题强化表和主题演化桑基图可知各阶段的研究重点的变化,2020 年及以前的阶段研究主要集中在综合医疗服务(阶段一主题 1)和慢性病防控与医防结合(阶段一主题 2)上,这两个主题的主题强度较

高。2021~2022 年阶段研究重点逐渐减少了对综合医疗服务(阶段一主题 1)的关注,转而增加了对公共卫生服务、医疗体系与公卫结合(阶段二主题 2)和综合医疗服务管理与融合(阶段二主题 1)的关注。2023~2024 年阶段研究主题健康服务与医防融合(阶段三主题 4)和医防融合与医疗服务(阶段三主题 1)成为新的关注主题,表明研究的重心进一步向医防融合和健康服务管理方向转移。

综上所述,综合医疗服务、慢性病防控与医防结合等主题在不同阶段都有较强的延续性,而研究重点随着时间的推移逐渐从综合医疗服务向医防融合和健康服务管理方向转移。这些变化反映了政策推动下,研究领域逐步深化和系统化的趋势。

4. 结语与展望

本文在中国知网、万方、国家哲学社会科学文献中心数据库中收集建库到 2024 年“慢性病医防融合”主题下的中文文献标题和摘要列表,并将其划分为 2006~2020 年、2021~2022 年和 2023~2024 年三个阶段。经过数据预处理、Word2vec 模型训练、LDA 主题模型训练,得到两阶段主题的相似性矩阵及文献主题演化桑基图。

在 2006~2020 年,许多学者主要关注的是综合医疗服务(阶段一主题 1)和慢性病防控与医防结合(阶段一主题 2)。这一阶段的研究主要集中在如何整合医疗服务,以及如何将慢性病防控与医防结合,研究重点在于建立基础的服务体系和初步的医防结合模式。

进入 2021~2022 年,学者们的研究更加深入和具体。他们不仅关注综合医疗服务管理与融合(阶段二主题 1),也对公共卫生服务、医疗体系与公卫结合(阶段二主题 2)进行了探讨。这一阶段的研究重点在于如何更好地管理和融合综合医疗服务,以及如何将公共卫生服务与医疗体系更有效地结合,研究开始关注系统性的管理和更深层次的医防结合机制。

到了 2023~2024 年,学者们的研究继续关注健康服务与医防融合(阶段三主题 4),同时细化关注了医防融合与医疗服务(阶段三主题 1)。这一阶段的研究重点在于如何实现健康服务与医防的深度融合,以及如何在医疗服务中具体落实医防融合的理念。研究更加注重实际操作和具体应用,力求在实践中实现医防融合的目标。

综上所述,研究主题随着时间的推移逐渐从基础的医防结合和综合医疗服务,过渡到更深层次的管理融合和系统性结合,最终细化到具体的健康服务和医疗服务中的医防融合。这一演化过程反映了学术界和政策制定者对医防融合和综合医疗服务管理的关注点逐渐深入和具体化。未来可以期待更多关于基层医疗与医防协同机制和医防融合与基层医疗服务的研究,以适用不同地区和人群的具体环境需求。这些研究将有助于推动社区卫生服务和公共卫生体系的进一步发展,为居民提供更加个性化和高效的服务,达到提升全民健康水平的效果。

基金项目

2023 年度上海市浦东新区科技发展基金事业单位民生科研专项医疗卫生项目, PKJ2023-Y76。

参考文献

- [1] 国务院办公厅印发《中国防治慢性病中长期规划(2017-2025 年)》[EB/OL]. https://www.gov.cn/xinwen/2017-02/14/content_5167942.htm, 2017-02-14.
- [2] 国务院办公厅关于印发“十四五”国民健康规划的通知[EB/OL]. https://www.gov.cn/gongbao/content/2022/content_5695039.htm, 2022-04-22.
- [3] 刘劲松, 王华东, 查震球, 等. 政策工具视角下我国慢性病医防融合政策分析[J]. 安徽预防医学杂志, 2021, 27(6): 501-506.
- [4] 马文雯, 李超凡, 刘聪慧, 孙强. 我国医防协同政策文本量化分析——基于政策工具、协同层次和协同机制的三

- 维框架[J]. 中国卫生政策研究, 2022, 15(7): 24-29.
- [5] 宋大平, 张玉鹏, 崔雅茹, 等. 慢性病基层治理网络、治理工具和治理路径[J]. 卫生经济研究, 2024, 41(3): 5-9.
- [6] 龙春晓, 李承璐, 范阳东, 等. 我国慢性病防治政策文本的量化分析: 基于政策工具、政策演进和政策主体的三维框架[J/OL]. 中国全科医学: 1-8. <https://www.chinagp.net/CN/10.12114/j.issn.1007-9572.2024.0480>, 2025-02-15.
- [7] 王汉文, 汪卓赞, 顾维波, 等. 基于彩虹模型的医防协同实施路径探索[J]. 南京医科大学学报(社会科学版), 2023, 23(1): 21-25.
- [8] 李明, 张伟, 王芳. 我国慢性病医防融合的现状与对策研究[J]. 中国医院管理, 2021, 41(4): 15-19.
- [9] 王敏, 李华, 赵强. 慢性病医防融合的实践与探索[J]. 中国医院管理, 2021, 41(4): 20-24.
- [10] 李辉, 陈洁平, 崔军, 等. 宁波市慢性病“1+X”医防整合模式实践与思考[J]. 中国健康教育, 2015, 31(9): 909-911.
- [11] Blei, D.M., Ng, A.Y. and Jordan, M.I. (2003) Latent Dirichlet Allocation. *Journal of Machine Learning Research*, **3**, 993-1022.
- [12] Nguyen, D.Q., Billingsley, R., Du, L. and Johnson, M. (2015) Topic Modeling with Word Embeddings. *Proceedings of the 53rd Annual Meeting of the Association for Computational Linguistics and the 7th International Joint Conference on Natural Language Processing*, Beijing, 26-31 July 2015, 1455-1464.
- [13] 席笑文, 郭颖, 宋欣娜, 等. 基于 Word2vec 与 LDA 主题模型的技术相似性可视化研究[J]. 情报学报, 2021, 40(9): 974-983.
- [14] Blei, D.M. (2012) Probabilistic Topic Models. *Communications of the ACM*, **55**, 77-84. <https://doi.org/10.1145/2133806.2133826>
- [15] Tang, J., Yao, L., Zhang, D. and Li, J. (2012) Topic Evolution in Scientific Literature. *Proceedings of the 21st ACM International Conference on Information and Knowledge Management*, Maui, 29 October-2 November 2012, 2172-2176.
- [16] 钟健. 基于 LDA 模型的京津冀协同发展政策文本主题分析[J]. 长春师范大学学报, 2024, 43(5): 119-124.