

基于生成式人工智能的公共政策一致性评价研究： 以三明医改为例

蔡媛青¹, 李志军², 何绍斌^{3,4}

(1. 中国社会科学院中国社会科学评价研究院, 北京 100732; 2. 中国海洋大学管理学院, 山东 青岛 266100;
3. 天津大学管理与经济学部, 天津 300072; 4. 北京中关村学院, 北京 100094)

摘要:为何一些地方改革能穿越“执行雷区”而另一些却在部门与层级间支离破碎? 本文以因跨层级协同持续而备受瞩目的中国三明医改为例, 运用生成式人工智能方法深入到政策执行的“黑箱”, 从理想化政策、执行主体、目标群体和政策环境四个核心要素建立公共政策一致性评价框架, 揭示公共政策一致性评价的动态机制。研究发现:①三明医改中医疗、医保、医药三类政策主题词语分布及演变趋势表现出较强的关联程度, 表明其公共政策一致性较高;②在动态主题模型分析中, 三明医改经历了从理想化政策到执行主体, 再到目标群体及政策环境相结合的三阶段动态演变路径;③运用生成式人工智能中的大语言模型对公共政策一致性评价框架的四个维度进行挖掘, 形成了系统性的评价指标, 为推进公共政策一致性评价提供了实践依据。

关键词:公共政策一致性评价; 生成式人工智能; 动态主题模型; 大语言模型挖掘方法; 三明医改

中图分类号:C93-3 **文献标识码:**A **文章编号:**1005-0566(2026)06-0071-12

Public policy consistency evaluation based on generative AI: Evidence from the Sanming healthcare reform

CAI Yuanqing¹, LI Zhijun², HE Shaobin^{3,4}

(1. *Chinese Academy of Social Sciences Evaluation Studies, Beijing 100732, China;*
2. *School of Management, Ocean University of China, Qingdao 266100, China;*
3. *College of Management and Economics, Tianjin University, Tianjin 300072, China;*
4. *Zhongguancun Academy, Beijing 100094, China*)

Abstract: Why can some local reforms cross the “minefield of implementation” while others are fragmented among departments and levels? This paper takes the healthcare reform in Sanming, China, which has attracted much attention due to the continuous cross-level collaboration, as an example. It uses a generative AI approach to delve into the “black box” of policy implementation, and establishes a framework for the consistency evaluation of public policies from four core elements: idealized policy, policy implementers, target groups and policy environment, revealing the dynamic mechanism of the consistency evaluation of public policies. The research findings are as follows: (1) The distribution and evolutionary trends of thematic terms concerning medical care, medical insurance and pharmaceuticals in the

基金项目:北京市社会科学青年基金规划项目“学术评价视域下中国特色公共管理学自主知识体系研究”(24GLC039);中国社会科学院“青启计划”项目“我国卫生健康政策评价的理论与方法研究”(2024QQJH010);中国社会科学院新时代党建研究中心和中国共产党巡视理论研究中心2027年度项目“高质量党建引领教育科技人才一体发展的实践路径研究”。

作者简介:蔡媛青(1993—),女,浙江台州人,中国社会科学院中国社会科学评价研究院助理研究员,研究方向为公共政策评价、卫生事业管理。通信作者:何绍斌。

Sanming healthcare reform shows strong correlations, indicating a high level of public policy coherence. (2) Through DTM analysis, the Sanming healthcare reform has undergone a three-stage dynamic evolution path from idealized policy to policy implementers, then to target groups combined with policy environment. (3) By applying large language models to mine the four dimensions of the public policy consistency evaluation framework, a systematic element deconstruction table was formed, providing a practical basis for promoting the consistency evaluation of public policies.

Key words: public policy consistency evaluation; generative artificial intelligence; dynamic topic model; large language model mining; Sanming healthcare reform

党的十八大以来,以习近平同志为核心的党中央将增强宏观政策取向一致性、提高政策协同作为推动高质量发展的重要举措。2024 年 4 月召开的中央政治局会议和年底召开的中央经济工作会议分别提出“要做好宏观政策取向一致性评估,加强预期管理”“加强预期管理,协同推进政策实施和预期引导”。2025 年政府工作报告提出“健全和用好宏观政策取向一致性评估工作机制,把经济政策和非经济性政策统一纳入一致性评估”,并强调要统筹政策制定和执行全过程,提升政策目标、工具、时机、力度、节奏的匹配度。“十五五”规划建议提出“增强宏观政策取向一致性,强化政策实施效果评价”。公共政策一致性评价逐渐成为学术界和实践领域关注的焦点^[1-2]。公共政策一致性评价是多方协调与协同增效的关键,有助于提高政策的整体效能。然而,政策出台部门化、执行碎片化等导致的“合成谬误”仍然存在^[3]。如何以可验证、可度量、可复制的标准展开系统评价,成为打通宏观治理“最后一公里”的关键瓶颈。本文以福建省三明市“三医联动”改革(以下简称“三明医改”)为切入口展开分析,揭示公共政策一致性评价的动态机制,在“宏观政策取向一致性”的国家命题与“地方差异化改革”的微观场景间架起桥梁^[4]。

本文的研究贡献如下。一是理论层面。已有研究多停留在概念阐释与个案叙事,缺乏可比较的量化指标^[5]。本文使用生成式人工智能方法推动史密斯政策执行模型从静态要素分析走向动态诊断分析,与复杂治理、数智评价等形成理论对话。二是机制层面。已有研究对“政策—执行—预期”的动态耦合缺乏系统分析^[6]。本文深入分析其影响因素和作用机制,有助于提升公共政策制定的连贯性与稳定性。三是情境层面。已有研

究对地方差异化改革场景的关注度不够,难以回应“一致性”与“因地制宜”间的张力^[7]。本文结合“三明医改”的现实案例,探究深化改革背景下公共政策一致性的协同逻辑,为推动公共政策宏观协同与地方因地制宜落实提供参考,为优化公共政策一致性评价提供依据。

一、文献综述与理论框架

公共政策评价作为政策科学的重要组成部分,其理论和方法论经历了长期的演进发展^[8]。从强调客观测量的实证主义,到关注价值多元和参与的建构主义,再到强调复杂系统的动态调适,公共政策评价的理论框架不断丰富^[9-10]。作为衡量政策质量与效能的关键维度,公共政策一致性评价理论模型和分析框架也随之发展。随着国家治理现代化的推进,如何构建系统、动态的理论模型成为关注的焦点。

(一)公共政策一致性评价的内涵

公共政策一致性作为公共政策研究中的核心概念,是确保政策目标实现、提升治理能力的关键。公共政策一致性不仅指政策在时间上的连续性,更涵盖了政策系统内部各要素之间的协调与统一^[11]。在实践中受政策问题的动态性、执行主体的多元性及政策环境的变化影响,政策间难免出现不一致、不协调的情况,削弱了整体效能,还可能导致资源浪费、社会矛盾加剧等负面后果。因此,准确界定其核心内涵与构成,是进行有效政策评价和优化的前提。本文研究认为,理解公共政策一致性评价的本质,须将其置于系统观念下进行研究。Smith^[12]提出的史密斯政策执行模型广泛应用于公共政策一致性评价中,其核心在于分析理想化政策、执行主体、目标群体、政策环境之间的匹配与协调。然而该模型也存在一定适用边界:一方面,史密斯政策执行模型在既有研究中

容易被静态化使用,难以充分呈现政策要素随时间演化的过程;另一方面,在多主体共同参与的复杂治理场景中,政策执行往往呈现非线性、情境依赖和涌现性特征,难以完整把握公共政策一致性的动态生成机制^[13]。因此,本文使用生成式人工智能方法,推动史密斯政策执行模型从静态要素分析走向动态诊断分析,以更完整地揭示公共政策一致性的动态生成机制。

(二)公共政策一致性的测量方法

国内外研究围绕公共政策一致性的量化测度,主要形成了三类分析路径。一是运用综合指数的政策文本量化评价。以PMC指数模型为代表的评价方法,运用多投入产出分析表、变量分类及PMC曲面构建等技术,对政策文本的内部一致性、优势特征与潜在缺陷进行系统评价^[14]。后续研究将其拓展应用于分级诊疗政策、商业健康保险等政策领域的文本评估,显著提升了公共政策评价的标准化水平^[15]。然而,此类方法主要聚焦于政策设计端与文本内部逻辑,对政策执行过程中的动态调适与反馈机制关注相对不足。二是采用政策工具与政策组合的内容分析方法。相关研究借助内容编码、文本挖掘等技术手段,从政策工具类型、政策强度等维度展开量化分析,能够较为清晰

地呈现政策工具的结构特征与组合逻辑^[16]。但该类方法对政策主题的历时演化、执行主体的策略响应以及目标群体的行为反馈之间的动态关联缺乏充分解释。三是进行计算文本分析的政策相似性与动态演化测量。例如,通过测算不同政策文本间的语义相似度,为跨地区、跨时期的大规模政策比较提供了方法论支撑,并拓展至政策主题结构识别与演化趋势分析^[17]。综上,现有研究已从PMC指数评价、政策工具编码、文本相似度测算等方面推进了政策一致性的量化研究,但仍存在两方面局限:第一,静态评估取向突出,难以揭示公共政策一致性在政策生命周期不同阶段的形成机理;第二,研究视域多局限于政策文本内部,对政策文本与执行实践、社会反馈之间的映射关系及互动逻辑缺乏系统研究。

(三)研究框架:公共政策一致性评价的动态反馈机制

公共政策一致性评价框架强调了理想化政策、执行主体、目标群体和政策环境的重要性。公共政策一致性评价时需综合考虑四大核心要素及其相互作用(见图1),确保评价在公共政策全局框架内更加精准。该框架不仅是静态分析,还是一个持续、动态的监测诊断过程。

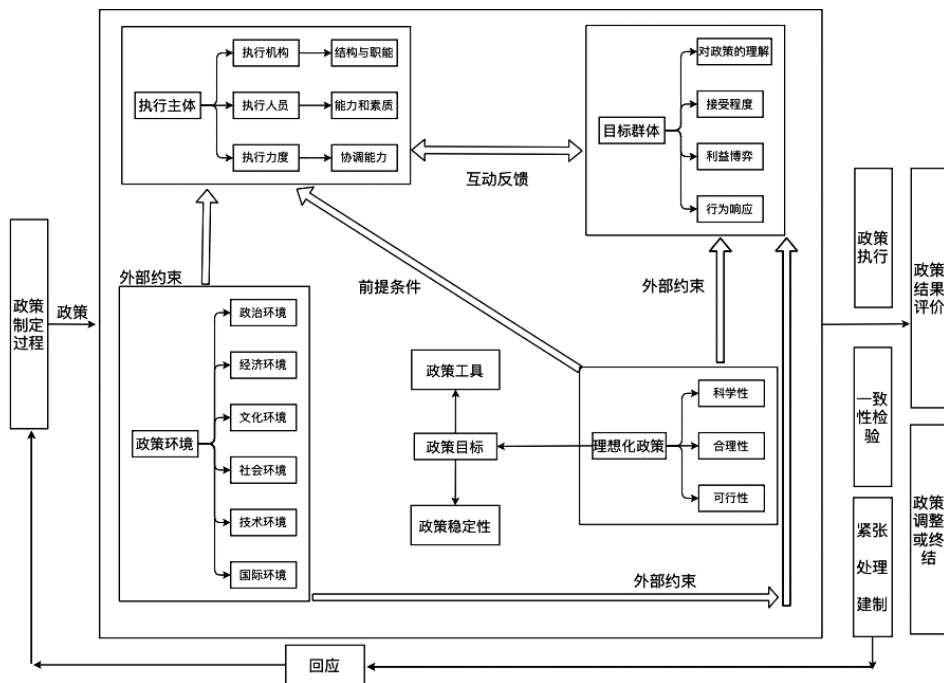


图1 公共政策一致性评价框架

1. 理想化政策与公共政策一致性

理想化政策是公共政策一致性的重要基础,强调公共政策本身应具备科学性、合理性与可行性。一是政策目标一致性。目标应清晰明确,且与其他政策保持方向一致,避免相互掣肘^[18]。二是政策工具一致性。为实现目标所采用的政策工具、管理模式和实施路径应相互补充、形成合力。政策工具的选择应基于对政策问题的准确诊断,确保其能够有效作用于政策目标。三是政策稳定性。政府期望的效果往往通过连贯政策来实现,而非依赖临时性政策。Van Engen 等^[19]将其定义为公共政策随时间推移保持稳定的程度。政策稳定性有助于减少不确定性,增强可信度;反之,政策频繁变动不仅会损害政府公信力,也会增加公共政策执行的不确定性。

2. 执行主体与公共政策一致性

执行主体是公共政策一致性的实施保障,决定了公共政策能否被准确、有力地执行。一是执行机构。执行机构的结构与职能应保持稳定,避免因频繁调整导致政策执行的中断或偏离。涉及多部门协同时,需构建高效协调体系,界定职责,避免推诿。例如,“三医联动”改革中,医疗、医保、医药间的协同治理至关重要,任何环节脱节都可能导致改革受阻^[20]。二是执行人员。关键岗位领导干部应保持任职稳定,以确保其政策持续推动。频繁的人事变动,尤其是“一把手”更替,往往导致“新官不理旧账”现象。执行人员的专业能力和职业素养也是影响能否准确把握政策精神、解决执行问题的重要因素。三是执行力度。执行主体若缺乏协调能力,可能导致政策执行偏离^[21]。

3. 目标群体与公共政策一致性

目标群体是公共政策一致性的效果标尺,关注对政策的理解、接受程度、利益博弈以及行为响应等方面^[22]。一是对政策的理解。政策应充分考虑目标群体利益诉求,避免因设计不当而引发抵触。若目标群体持怀疑或抵触态度,会导致政策执行效果不佳,影响政策预期的实现^[23]。二是接受程度。目标群体的反馈和参与程度是衡量政策一致性的重要指标。信息不对称或误解是导致政

策执行偏差的重要原因。有效的沟通和宣传可提高政策认同感,增强配合意愿^[24]。三是利益博弈。政策制定者、执行者与目标群体间的利益博弈会影响公共政策一致性。例如,在制定政策时如果忽视了居民的便利性和习惯,就可能导致政策流于形式。四是行为响应。政策应通过激励或约束机制,引导目标群体朝着政策预期方向调整其行为^[25]。

4. 政策环境与公共政策一致性

政策环境是公共政策一致性的外部约束,包括政治、经济、文化、社会、技术、国际环境等内外部影响因素,强调与公共政策之间的适应性与协调性^[26]。该维度可从宏观和微观两个层面分析:宏观层面,公共政策应同国家发展战略、法律法规相衔接;微观层面,应坚持因地制宜,鼓励地方结合实际探索创新。此外,公共政策还应具备灵活性和调适性。

二、研究设计与方法

(一) 案例选择与概况

2021 年 3 月,习近平总书记视察福建省三明市时指出:“三明医改体现了人民至上、敢为人先,其经验值得各地因地制宜借鉴”。三明医改作为深化医疗体制改革的范例,其经验已被纳入国家顶层设计,成为推进“健康中国”建设和促进卫生健康事业高质量发展的重要实践模式^[27]。本文研究选择三明医改作为案例,其代表性体现在以下方面。其一,具有系统复杂性。其核心“三医联动”是对医疗、医保与医药的协同改革,需考量理想化政策、执行主体、目标群体和政策环境等多维度影响,为公共政策一致性评价研究提供了经验样本。其二,具有政策演进特征。三明医改经历了从“以治病为中心”向“以健康为中心”的持续演进,为政策要素协同演进机制分析提供了素材。其三,具有实践创新性。三明医改作为国家级医改范例,为复杂社会问题领域的公共政策评价提供了重要的实践参照。

(二) 数据来源与处理

本文研究历经 6 年(2019 年 6 月—2025 年 8 月),采用多源数据收集方法以全面审视三明医改

过程。经数据清理后按照来源划分,核心数据包括 245 份源自三明医改领导小组成员和三明医改相关核心部门人员的一手资料和 39 份通过公开检索获取的材料。按照类型划分,其中有 234 份为政

策文本,50 份数据为访谈文本和新闻文本。本文参考 PRISMA 框架^[28],构建了用于分析的数据集,确保了数据的完整性与相关性(见图 2),最终获得数据 284 份,文本总字数为 2 040 038 字。

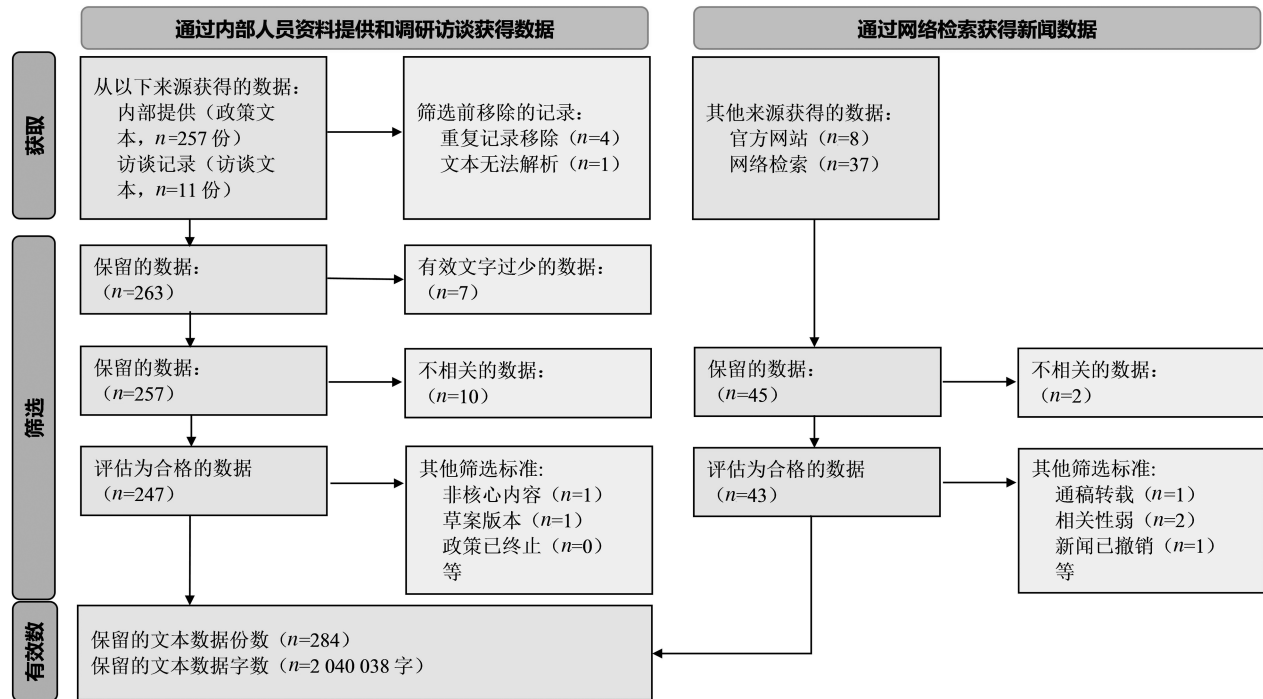


图2 数据处理过程

本文的数据来源包括三类(见表1):第一,政策文本,通过三明市医改领导小组获取 234 份内部文件,涵盖 2012 年启动医改以来三明市涉及“三医联动”的文件、方案、纪要及制度汇编;第二,访谈文本,对三明医改相关核心部门人员进行多轮访谈;第三,新闻文本,通过网络检索获取主流媒体关于三明医改的新闻报道和深度评论共 39 份。

表1 三明医改数据收集概况

数据类型	数据来源	收集方式	数量/份	主要内容
政策文本	三明医改领导小组成员	内部资料提供	83	医保支付、基金管理等相关政策
			123	医院薪酬、医疗服务价格等相关政策
			28	药品采购、配送等相关政策
访谈文本	三明医改相关核心部门人员	深度访谈、参与式观察、小组座谈	11	三明医改政策制定、执行与监管的细节与考量过程
新闻文本	国家、省、市级主流媒体	网络检索	39	关于三明医改的新闻报道、深度评论

注:访谈对象涵盖福建省三明市医改相关核心部门与人员,包括三明市人大常委会、市委党组成员、市医改领导小组、市卫生健康委及公立医院院长。

(三)研究方法 with 实施步骤

研究基于生成式人工智能的方法进行公共政策一致性评价,包含动态主题模型(dynamic topic model, DTM)分析和大语言模型(large language

model, LLM)挖掘分析两个部分。首先,采用 DTM 对政策文本进行纵向的时间序列分析,客观提取并刻画公共政策在不同演进阶段的核心主题,对主题—词语、主题演变等进行分析;其次,将所得政策主题作为参照系,引入大语言模型挖掘方法,对非结构化的访谈、新闻文本进行语义映射与测度。两种方法构成递进关系:前者为后续分析提供结构化的维度框架,而后者则负责测度各维度在实践领域的映射状况,二者协同实现对公共政策一致性的综合评价。

1. 动态主题模型分析

本文采用 Blei 和 Lafferty 提出的动态主题模型^[29]。传统的静态方法,如潜在狄利克雷分布主题模型(latent dirichlet allocation, LDA)或以 TF-IDF 为代表的词汇加权方法,缺乏时间维度的显式建模,无法捕捉主题演化。DTM 是 LDA 在时间维度上的扩展,将文本集合视为有序时间序列,假定主题本身随时间平滑演进。在标准的 LDA 中,主

题—词语分布 (β_k) 从固定狄利克雷先验中抽取,而在 DTM 中,每一时刻 t 的主题—词语分布 $\beta_{t,k}$ 以及文档—主题分布 $\theta_{d,t}$ 都与其前一时刻状态相关联。主题演化被建模为一个状态空间模型,通过高斯分布在 对数空间中关联:

$$\beta_{t,k} | \beta_{t-1,k} \sim N(\beta_{t-1,k}, \sigma^2 I)$$

其中, $\beta_{t,k}$ 是第 t 个时间切片下第 k 个主题的自然参数; N 代表正态分布; σ^2 是过程方差,控制主题随时间演变的速率。该设计使得模型能够捕捉公共政策从萌芽、发展到成熟或消亡的完整生命周期。

具体实施策略如下。首先,引入生成式人工智能技术,利用生成式大语言模型 DeepSeek R1 对

政策的标题进行关键词识别。依据识别结果将其划分为医疗、医保、医药以及不相关四个语料库,只保留前三类数据。为确保分类准确性,对 5% 的样本进行人工抽样验证。研究人员在不知晓 LLM 分类结果的情况下进行独立编码。结果显示,LLM 分类的总体准确率为 91.3%,宏平均 $F1$ 值为 0.91, Cohen's Kappa 系数为 0.87,表明分类结果具有较高信度和效度。其次,依据三明医改的三个阶段,将文本按年份排序,从而构建一个“3 × 3”的分析矩阵(见表 2)。深入挖掘三类政策内部特有的演变路径,可以有效避免领域间信号相互干扰导致的信息混淆。最后,运用 DTM 方法对主题—词语、主题演变等维度进行分析。

表 2 三明医改文本数据分类情况

时间维度	医疗领域	医保领域	医药领域	阶段	特征
2012—2013 年	文本集 A1	文本集 B1	文本集 C1	治混乱、堵浪费	“以赚钱为中心”向“以治病为中心”转变
2014—2015 年	文本集 A2	文本集 B2	文本集 C2	建章程、立制度	“以治病为中心”
2016 年—至今	文本集 A3	文本集 B3	文本集 C3	看好病、大健康	“以健康为中心”

2. 大语言模型挖掘分析

引入生成式人工智能的 LLM 进行深度文本分析。LLM 不仅具备强大的自然语言生成能力,更能够捕捉深层语义、情感色彩与非正式表达等信息,适用于访谈、网络资料等口语化与非结构化文本的信息挖掘^[30]。具体而言,参考金星晔等^[31]学者的做法,研究将 DTM 模型分析得出的政策主题作为参照系,利用 LLM 测度政策主题在访谈、新闻文本中的实际体现程度与反馈评价。LLM 的核心任务是评估政策主题在利益相关方认知中的映射情况。该过程可被表达为式(1)所示的函数:

$$S_{ij} = fLLM(Prompt(T_i), D_j) \quad (1)$$

在此表达式中, $T = \{T_1, T_2, \dots, T_k\}$ 是由 DTM 模型分析得出的 k 个核心政策主题。 $D = \{D_1, D_2, \dots, D_m\}$ 表示 m 份非结构化文本的语料集合。 $Prompt$ 是指将政策主题映射到自然语言提示空间的函数,即 $Prompt(T_i)$ 代表针对特定主题 T_i (如“药品一体化供应管理”)而设计的、结构化的提示词。函数代表 LLM 的语义分析功能,它将 $Prompt(T_i)$ 应用于指定的访谈文本 D_j 。其输出 $S_{ij} = [S_{ij}^1, S_{ij}^2, \dots, S_{ij}^5]^T$ 是多维度的评价向量,其中

S_{ij}^1 是特定政策主题 T_i 在访谈、新闻文本 D_j 中是否涉及的映射判定(正向涉及记为 1,负向涉及记为 -1,未涉及记为 0),而 S_{ij}^2 至 S_{ij}^5 则分别记录了该主题 T_i 在理想化政策、执行主体、目标群体和政策环境中的具体映射内容。基于上述过程,构建“政策实践映射率”这一指标,用来衡量政策主题在访谈、新闻文本中的体现程度^[32],其计算方法定义为:

$$PMR_i = \frac{\sum_{j=1}^m S_{ij}^1}{m} \times 100\% \quad (2)$$

其中, m 为分析的访谈、新闻文本总数(本研究中 $m = 50$)。该指标衡量政策主题在访谈、新闻文本中出现的频率和方向,正值越高表明该主题在实践中得到越多正面关注和执行,负值则表示存在质疑或执行困难。同时 S_{ij} 的其他维度为后续的政策要素分析和内容挖掘提供了基础。为提高 LLM 输出的可复核性,研究采用结构化提示词,统一设定角色、任务、判定标准和输出格式,并要求模型仅依据给定文本作答,不得引入外部知识。

三、基于生成式人工智能的 DTM 分析

主题数量的设定对 DTM 模型效果具有显著影

响。本文采用主题一致性得分确定最优主题数,并辅以主题内高频词语义相似度检验主题可解释性。主题一致性得分越高,表明主题内聚性越强、语义边界越清晰^[33]。经主题一致性得分测算,确定最优主题数为6。

(一)主题—词语分析

1. 第一阶段:治混乱、堵浪费

在改革的启动探索期(2012—2013年),三明医改在医疗费用失控、药品耗材价格虚高和医保基金穿底的多重压力下启动,呈现出需求倒逼下的政策协同演化逻辑。政策文本分析显示,医疗、医保、医药三类政策工具围绕核心治理目标形成高度协同。比如,“医联体”关键词在三类政策文本中均占据核心地位(医疗政策中 $p=0.049$,医保政策中 $p=0.081$,医药政策中 $p=0.124$),不仅作为具体治理手段,更充当了连接三医的关键纽带。

2. 第二阶段:建章程、立制度

在改革的深化攻坚期(2014—2015年),三明医改强调通过多部门协同和层级联动,构建高效政策执行网络。将第一阶段行之有效的改革工具,通过“建章程、立制度”使其成为常规机制。比如,基层医疗机构在医疗和医药政策中的集中出现,特别是“乡镇卫生院”在“医共体与分级诊疗”主题中呈现出较高概率值($p=0.109$),表明基层医疗机构已从边缘化的位置,转变为分级诊疗体系的重要支撑。

3. 第三阶段:看好病、大健康

在改革的系统整合期(2016年至今),三明医

改关注从“治病”向“健康”的理念升维,从宏观制度设计转向具体人群需求和环境适应。比如,“城乡居民”在“多层次医保体系构建”主题中的概率值为0.041,以及城镇居民、城镇职工的分别出现,表明政策考虑到不同人群在医疗需求、支付、就医习惯上的差异性。“医务人员”在“公共卫生与医防融合”等多个主题中反复出现,特别是在“药费控制与薪酬联动”主题中的概率值为0.071,表明了其既是改革推动者,也是改革对象。

(二)主题演变分析

本节分析三明医改中医疗、医保和医药政策的DTM主题流行度概率演变过程(见图3)。横轴代表三个改革阶段,纵轴代表主题流行度概率。

一是医疗政策。医疗政策的演变体现了执行主体利益重构与目标群体行为引导的协同机制。医疗政策演变展现了执行主体与目标群体间从博弈到合作性治理的转变。“医院财务与成本管控”始终保持较高水平,反映了医疗机构持续进行运营管理。“公共卫生与医防融合”主题先升后降,反映出初期注重医疗质量与防治,但后期转向大健康策略;“医共体与分级诊疗”则先降后升,反映出政策工具从“堵漏洞”转向“疏堵结合”。通过将医务人员的收入与医疗质量、患者满意度挂钩,而非简单与药品收入脱钩,实现了从零和博弈向正和博弈的转变^[34],使得政策执行不再依赖外部强制,而是通过内生激励进行强化。

二是医保政策。“医保支付改革”等主题逐渐上升,反映出理想化政策的递进式深化,政策重心

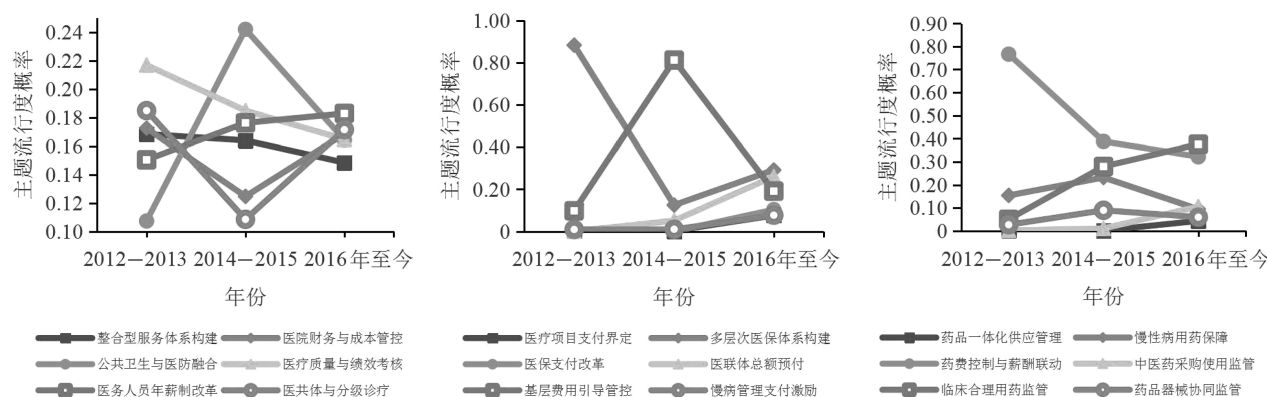


图3 政策DTM主题流行度概率演变过程(从左到右分别为医疗、医保和医药)

适时转向更复杂的治理工具^[35]。“基层费用引导管控”和“医联体总额预付”在第二、三阶段的快速增长,表明医保政策从“保基本”向“促健康”的战略转型。通过调整执行主体(从医保部门主导到医保—医疗协同)和重新定义目标群体(从参保患者到健康人群),实现了政策系统的内在一致性重构。

三是医药政策。医药政策演变趋势体现了跨部门协同与环境适应的耦合效应。“药费控制与薪酬联动”在第一阶段呈主导地位,体现医药改革将药品从利润中心转为成本中心,从根本上重塑了执行主体的行为逻辑。医药政策不再局限于价格管控的单一维度,而是将药品治理置于医保支付、临床路径等多重政策环境的交叉点上,由此形成了深度耦合的政策网络,体现出从“联动治理”到“系统整合”的视角转变。

三明医改主题流行度概率的变化不仅体现为医疗、医保、医药三类政策工具的变化,而且经历了从理想化政策到执行主体,再到目标群体及政策环境相结合的三阶段动态演变路径。第一阶段围绕治混乱、堵浪费等形成理想化政策目标与政策工具设计;第二阶段围绕建章程、立制度进一步转向卫生健康部门、医保部门、医院等执行主体的激励约束重构;第三阶段围绕看好病、大健康延伸至患者等目标群体进行引导,并在医防协同、健共同体建设等政策环境中实现协同整合。

四、基于生成式人工智能的大语言模型分析

前文基于 DTM 提取了三明医改的 18 个核心政策主题。为进一步研究上述政策议题在实践层面的落实与反馈情况,基于上述政策主题引入 LLM 进行挖掘分析,映射与测度不同政策主题在实践层面的表现以开展公共政策一致性评价。

(一) 政策主题映射识别

三明医改政策主题的实践映射率呈现出显著的分层分布特征(见图4)。18个政策主题的 PMR 值均为正值,表明三明医改的各项政策在整体上得到了积极的实践响应,未出现系统性的负面评价或抵制。然而,不同主题之间的映射程度存在显著差异, PMR 值为6%~94%,跨度达88

个百分点,反映出政策执行的选择性和非均衡性特征。映射率超过 80% 的政策主题共有 7 个,包括“医保支付改革”(94%)、“整合型服务体系构建”(94%)等。映射率为 50% ~ 80% 的政策主题共有 5 个,包括“公共卫生与医防融合”(76%)、“基层费用引导管控”(74%)等。以“公共卫生与医防融合”为例,虽然这一理念在政策设计中被反复强调,但在实际执行中面临着部门协调、资源整合等多重挑战,导致其在访谈、新闻中的呈现频率低于预期。值得注意的是,这一组主题多涉及跨部门协作和系统性整合,如医防融合需要医疗和公共卫生系统的深度整合。映射率低于 50% 的政策主题共有 6 个,其中“多层次医保体系构建”仅为 6%,显示出较低的实践关注度。本研究进行了包括替换提示词与更换模型在内的多项稳健性检验,结果均表明大语言模型分析结果具有较高的稳健性。

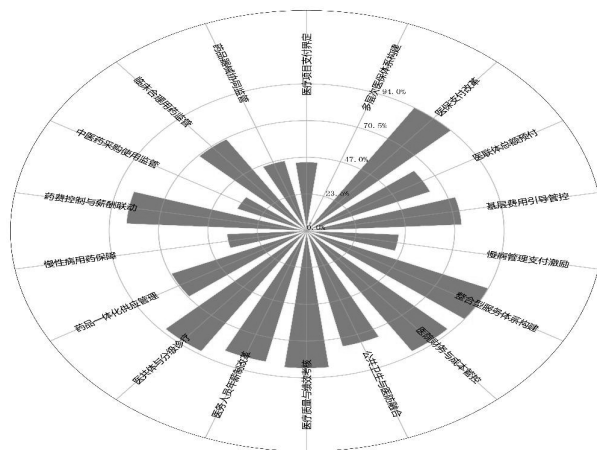


图4 三明医改政策主题实践映射率统计

(二)形成公共政策一致性评价指标

结合前文动态主题建模结果,运用 LLM 方法进行文本挖掘,围绕三明医改的理想化政策、执行主体、目标群体与政策环境四个维度进行分析,形成公共政策一致性评价指标(见表3)。

1. 理想化政策维度:三医联动的工具创新路径

从理想化政策维度来看,三明医改的系列政策要素反映了“三医联动”的工具创新路径,呈现出支付机制创新、管理体制变革与服务模式转型

表 3 基于 LLM 的公共政策一致性评价指标

政策主题要素	理想化政策	执行主体	目标群体	政策环境
医疗项目支付界定	DRG/DIP 动态调整机制、智能审核系统、按人头/病种打包支付、特殊病种单独支付机制	国家医保局推进支付改革、医保部门统筹支付标准、部门协同	医疗机构费用结算、明确患者自付费用标准、县域医疗机构费用控制	适应临床需求、遏制医疗资源浪费、促进医防融合
多层次医保体系构建	构建基本医保+大病保险+医疗救助+普惠医联保	医保局统筹管理、商业保险机构参与	基本医保参保人群	多层次保障需求
医保支付改革	支付方式与控费措施联动、总额预付结余留用、总额包干超支不补	医保部门建立支付杠杆	紧密型医共体资源下沉、健康管护组织效能提升	遏制过度医疗、价格虚高治理、切断药品耗材收入挂钩
医联体总额预付	总额包干结余留用	医保局统筹管理、总医院统筹分配	县域内医疗机构	促进健康绩效导向、分级诊疗推进
基层费用引导管控	工分制分配机制、服务收入计入工资、统一药品目录下沉资源	医共体实施费用减免、市县乡统筹推进	基层诊疗行为规范、乡村医疗机构费用管控	资源下沉需求、双向转诊起付线合并
慢病管理支付激励	零起付线报销、慢性病一体化管理奖励资金、特殊病种报销政策倾斜	总医院统筹、疾控中心参与分配	慢性病患者并发病防控	强化医防协同、从看好病转向健康管理
整合型服务体系构建	组建县域总医院体系、人财物一体化、双向转诊与远程诊疗	总医院统一管理、市县两级医联体参与、各级政府及卫健部门指导	县乡村三级机构协同	分级诊疗推进、优质资源下沉基层需求、健康责任共担、提升县域就诊率
医院财务与成本管控	总会计师制与目标年薪制、药品耗材联合限价采购、工资总额与医疗服务收入挂钩、财政承担基建设备支出、药品零差率销售	总医院统一核算、财政局/人力资源和社会保障局推进	院长及医务人员行为规范/薪酬管理、公立医院运营优化	破除以药养医机制、控制医疗成本
公共卫生与医防融合	建立医防融合项目清单、健康宣传教育、疾控医疗双向挂职	疾控与医疗机构一体化、总医院统筹管理	全民健康管理、公立医院医护人员	强化疾病预防、疫情防控推动改革
医疗质量与绩效考核	院长考核评价体系、突出技术水平和公益性考核、工分制与年薪制考核、职称评审权限下放	医改领导小组统筹、总医院考核	医院管理层、临床医务人员	公益导向治理、过度医疗现象遏制、引导办医方向
医务人员年薪制改革	职称核定基本年薪+绩效年薪结构	医改领导小组主导、成立公立医院薪酬委员会、政府考核发放院长薪酬	公立医院在岗人员	薪酬与药品收入脱钩、破除逐利机制、人才流动促进
医共体与分级诊疗	总院制管办分开机制、专科医生下沉村卫生所、县域医保基金打包支付、建立转诊率考核机制	总医院统筹、县级卫健部门主导实施、县域医共体管理机构	县乡村一体化、基层患者健康改善、向基层首诊患者倾斜	医疗资源结构性调整
药品一体化供应管理	药品集中分类采购机制、统一药品采购平台、两票制限价采购目录	市医改领导小组推进、政府主导采购监管、药采办审定配送	公立医院用药规范、药品供应环节	虚高药价空间压缩
慢性病用药保障	推行一病多方制度、便民门诊免挂号费报销、门诊特殊病种目录动态调整、医保端口直接结算	总医院统筹发放、统一健共体目录、医保部门监管	慢性病患者群体	并发症控制需求、控制长期用药负担
药费控制与薪酬联动	药费结余纳入薪酬总额、药品耗材重点监控	总医院绩效考核体系、三医联动平台、公立医院财务部门执行	医务人员薪酬结构优化、医务人员阳光收入提升、医生处方行为规范	以药补医机制破除、控制医疗成本
中医药采购使用监管	院内制剂研发支持、中医治疗绩效倾斜机制、中西医并重采购政策、中医非药物疗法零起付	市县中医药管理局监管	中医诊疗机构、乡镇卫生院中医科室	中西医同病同价、医保经费用节约需求
临床合理用药监管	建立重点药品监控目录、两票制监管、一品两规限制用药品种	药品监管机构、卫生局动态跟踪管理、医院药事管理部门执行	临床医师群体处方行为	医疗腐败专项整治要求、治理过度用药现象、回扣促销整治
药品器械协同监管	耗材试剂联合限价采购、低价中标质量保障机制、两票制实施	卫健医保市场监管协同、成立三明采购联盟	医疗药品耗材购销环节	全国集采经验推广、价格虚高问题整治

三条相互支撑的改革路径。其一,在支付机制创新路径中,核心指标集中体现为医保支付杠杆的精准运用,包括推行 DRG、DIP 动态调整机制、按人头与病种打包支付,以及落实“总额包干、结余留用、超支不补”的医联体总额预付与慢病特殊病种单独支付机制。其二,在管理体制改革路径中,核心指标聚焦于改变利益分配格局与重塑激励机制,涵盖了实施总会计师制与院长目标年薪制、落实医务人员“职称核定基本年薪与绩效年薪”的薪酬结构改革,以及推行药品耗材联合限价采购、两票制与药品零差率销售等财务与成本管控指标。其三,在服务模式转型路径中,核心指标指向了健

康管护体系的重构,包括建立“人财物一体化”的县域总医院体系、实施双向转诊与专科医生下沉村卫生所的医共体分级诊疗,以及建立医防融合项目清单、中医非药物疗法零起付等整合型公共卫生与慢病防控服务举措。

2. 执行主体维度:协同治理的组织架构

从执行主体维度来看,三明医改展现了高度权责明晰、垂直与扁平相结合的协同治理组织架构(见表 4)。该体系的顶层是由三明市医保局和三明市医改领导小组负责统筹主导、规则制定与资源宏观配置,其要素表现为市级医保部门统筹推进支付方式改革与支付标准界定,市医改领导

小组牵头主导薪酬改革、成立市级公立医院薪酬委员会及推进药品一体化供应管理。县域总医院则承担中枢协调功能,负责区域资源整合与绩效的二次分配,其要素表现为县域总医院统一实施内部财务核算与县域医联体医保基金统筹分配,并负责院内绩效考核、慢病管理统筹以及规范药品供应一体化配送。作为落实服务终端的基层医疗机构,主要负责首诊与健康管理的闭环的末端执行,其要素表现为县域医共体管理机构推动市、县、乡三级统筹落实市、县专科医生下沉与基层费用减免等属地化工作。此外,还配套了负责监督与技术支持的辅助执行主体,其要素表现为三明市疾控中心深入参与慢病一体化管理及医防融合指导,三明市疾控中心、三明市市场监督管理局(药监部门)协同实施药品器械采购与动态跟踪管理,以及三明市财政局等部门在医院财务成本管控中的支持。该维度揭示了政策执行的多元主体网络与层级分工机制。多部门协同机制在多个主题中发挥核心作用,形成了以三明市医保局和三明市医改领导小组为统筹主导、县域总医院为枢纽、基层医疗机构为服务终端的三级执行体系。

表 4 三明医改执行主体架构

层级	主体	功能定位	典型表现
决策 统筹层	三明市医保局、三明市医改领导小组	规则制定与资源配置	支付标准设定、薪酬委员会成立
管理 枢纽层	县域总医院	资源整合与绩效分配	总额预付资金统筹、内部考核
服务 提供层	基层医疗机构	基层首诊与健康管理	专科医生下沉、双向转诊执行
专业 监管层	三明市疾控中心、三明市市场监督管理局(药监部门)、三明市财政局	专项监督与技术支持	疾控医疗双向挂职、药品限价采购

3. 目标群体维度:利益相关方的差异化调控

从目标群体维度来看,三明医改反映了政府对各关键利益相关方的差异化调控与精准施策。其调控涵盖了以下三方面群体。一方面,医疗机构作为县域医疗资源运营模式的管理对象,其核心指标聚焦于县域内不同级别公立医院的运营优化、费用控制结算以及药品耗材购销环节的规范性治理。另一方面,患者群体作为健康权益的再分配与保障对象,涉及的主题聚焦于就医负担的实质性减轻与健康获益,核心指标体现为针对慢性病患者群体的并发症防控、基本医疗参保人群

的多层次用药保障以及明确患者自付费用标准的倾斜性设计。此外,医务人员作为政策激励与约束的双重客体,其核心指标既包括医生处方行为规范、过度医疗行为等约束管理,也涵盖了医务人员薪酬结构优化等激励机制设计。

4. 政策环境维度:外部约束与改革驱动

从政策环境维度来看,政策主题要素指标深刻反映了公共政策一致性的外部约束条件与改革驱动力,集中体现为三类要素。第一类是医疗资源配置失衡产生的结构性倒逼。该要素在医联体总额预付、整合型服务体系构建与基层费用引导管控等主题中高频出现,核心指标包括破解优质资源下沉基层的迫切需求、推进分级诊疗与双向转诊、扭转县域就诊率低下以及遏制区域医疗资源浪费。第二类是行为性失范引发的治理危机。该要素普遍存在于药费控制与薪酬联动、临床合理用药监管等相关主题中,核心指标包括破除以药养医机制、遏制过度医疗与价格虚高、切断药品耗材与医生收入的利益挂钩,以及纵深推进医疗腐败、回扣促销专项整治的刚性制度要求。第三类是民众健康需求的转型带来的理念升级。该要素全面贯穿于公共卫生与医防融合、多层次医保体系构建等主题中,核心指标涵盖了全面健康管理的理念转向、人口老龄化背景下慢病并发症长效控制的内在需求。

五、研究结论与政策启示

(一)研究结论

一是三明医改中医疗、医保、医药三类政策主题词语分布及演变趋势表现出较强的关联程度,表明其公共政策一致性较高。三明医改通过宏观层面的高位推动、中观层面的机构整合与信息互通、微观层面的指标量化与闭环管理,实现了医疗、医保、医药三类政策主题的高度关联与协同,形成了“1 + 1 > 2”的协同效应,为公共政策一致性评价与地方因地制宜落实政策提供了情境层面的经验证据。

二是在动态主题模型分析中,三明医改经历了从理想化政策到执行主体,再到目标群体及政策环境相结合的三阶段动态演变路径。运用生成式人工智能方法推动史密斯政策执行模型从静态要素分析

走向动态诊断分析,有助于理解政策执行过程中要素重心的阶段性转换,提升公共政策制定的连贯性与稳定性、提高政策质量与治理能力。

三是运用生成式人工智能中的大语言模型对公共政策一致性评价框架的四个维度进行挖掘,形成了系统性的评价指标,为公共政策一致性评价提供了可操作、可移植、可扩展的分析工具,为推进公共政策一致性评价提供了实践依据。运用生成式人工智能中 LLM 与 DTM 相结合的方法,使政策主题演变轨迹得以追踪,从而将史密斯政策执行模型拓展为面向复杂治理场景的动态诊断理论工具,为其他地区或领域开展公共政策一致性评价提供了可借鉴、可推广、可落地的分析范式与方法论参考,助力我国医改乃至更广泛公共政策领域迈向“数据驱动、模型支撑、精准治理”的新阶段。

(二)政策启示

第一,构建全生命周期的公共政策一致性动态监测体系。一方面,强化顶层设计与政策评价的制度衔接。将包含“理想化政策、执行主体、目标群体、政策环境”四维度的评价框架纳入重大改革项目的标准评价流程。另一方面,建立动态纠偏与即时反馈机制,依据动态监测结果及时识别政策在不同演化阶段的堵点。在政策执行初期重点关注政策工具的协同性,在中期关注政策执行主体的稳定性,在后期关注目标群体的满意度与环境适应性,实现政策系统的动态优化。

第二,加强相关省市之间的政策对接和协调,形成公共政策评价区域发展的整体合力。一方面,通过打破部门壁垒,建立跨区域政策协调委员会,定期召开政策对接会议等,推动区域间政策目标、标准与实施步骤的协同;搭建区域政策信息共享平台,实现各地政策制定过程的透明化与实时互动,减少因信息不对称导致的政策冲突。另一方面,鼓励各地在坚持国家统一政策的前提下,结合自身实际,探索符合地方特点的政策措施,形成各具特色、优势互补的区域发展格局。联合开展区域政策一致性评价,共同制定评价指标和评价标准,确保评价结果的科学性与可比性。

第三,提升政策对目标群体需求与环境变化的敏捷响应水平。一方面,增强公共政策的社会包容

性与回应性。公共政策一致性的核心在于获得目标群体的认同与配合,可以采用生成式人工智能大模型等先进技术持续追踪患者、医务人员及社会公众的舆论反馈,精准画像目标群体的差异化需求。在“以治病为中心”向“以健康为中心”转变的过程中,确保政策红利被人民群众切实感知,提高政策的社会接受度。另一方面,增强公共政策对于复杂环境的适应性和韧性。鼓励地方在保持宏观政策取向一致性的前提下,根据本地资源禀赋与发展阶段进行差异化探索与适应性创新,充分利用比较优势,实现政策系统与外部环境的良性互动。

(三)研究展望

公共政策一致性评价仍面临诸多挑战,未来的研究可以从以下方面进一步深化。一是运用数智赋能等现代信息技术,建立政策评价数据库和模型库,提高评价的科学性和精准性。例如,可以利用大数据技术对政策相关的大量数据进行挖掘和分析,识别政策间的潜在冲突和协同效应;结合知识图谱技术,将分散在各部门的政策数据整合为动态更新的政策知识库,为评价提供全域数据支撑。二是可以利用人工智能技术建立政府政策模拟和预测模型,对不同政策方案的可能影响进行仿真和评价。例如,开发基于机器学习的政策效果预测系统,通过训练历史政策数据,预测新政策在经济、社会、环境等维度的潜在影响;运用多智能体仿真技术,模拟政策实施后各利益相关方的行为响应,以可视化方式对比不同政策组合的长期效应,优化政策设计。

参考文献:

- [1]李志军,李逸飞.建立和完善中国特色公共政策评估制度[J].国家现代化建设研究,2023,2(3):91-99.
- [2]负杰.公共政策评估的制度基础与基本范式[J].管理世界,2023,39(1):128-138.
- [3]孙早,高昕凯.“同频共振”抑或“合成谬误”:人工智能背景下的政策协同如何影响劳动力就业[J].财贸经济,2026,47(2):84-105.
- [4]庄琦,翟绍果,黄春柳等.“三医”联动走向健康治理的实现机制:基于三明医改的纵向案例研究[J].经济与管理研究,2024,45(3):94-108.
- [5]王树涛,熊乐天.省域高校分类评价执行的制约因素与路径选择:基于史密斯政策执行过程模型的分析[J].国家教育行政学院学报,2025(10):25-35.

- [6] BUNDI P, TREIN P. Evaluation use and learning in public policy[J]. *Policy sciences*, 2022, 55(2): 283-309.
- [7] PATASHNIK E M, WEAVER R K. Policy analysis and political sustainability[J]. *Policy studies journal*, 2021, 49(4): 1110-1134.
- [8] 蔡俊, 杨岚, 周亚虹. PSM-DID 在政策评价中的应用现状与改进方法[J]. *管理科学学报*, 2024, 27(2): 30-48.
- [9] 王思琦. 中国政策试点中的随机实验: 一种方法论的探讨[J]. *公共行政评论*, 2022, 15(1): 30-50.
- [10] 黄平平, 刘文云. 中国特色新型智库政策文本量化评价: 基于 PMC 指数模型分析[J]. *情报科学*, 2024, 42(8): 136-143.
- [11] 李志军, 张毅. 公共政策评估理论演进、评析与研究展望[J]. *管理世界*, 2023, 39(3): 158-171.
- [12] SMITH T B. The policy implementation process[J]. *Policy sciences*, 1973, 4(2): 197-209.
- [13] RUIZ ESTRADA M A. Policy modeling: definition, classification and evaluation[J]. *Journal of policy modeling*, 2011, 33(4): 523-536.
- [14] 王文婷, 王佃利. 政策一致性理论的关键议题与研究面向: 政策过程研究新视角[J]. *中国行政管理*, 2025, 41(10): 13-22.
- [15] FU L, WANG R, DONG Y. Policy texts and actions: quantitative evaluation of the hierarchical medical system policy in China based on the PMC-Index model[J]. *BMC public health*, 2025, 25(1): 2358.
- [16] 胡峰, 李加陈, 翟婧. 政策文本计量视角下科技人才政策分析与评价: 基于“工具—效力”的二维框架[J]. *情报科学*, 2024, 42(6): 99-112.
- [17] ZHANG H, ZHU L, ZENG C, et al. Text mining and quantitative research of medical service policy: sichuan province as an example[J]. *Frontiers in public health*, 2021, 8: 509842.
- [18] 王有强, 李海明, 王文娟. 卫生体系和服务能力现代化的实现路径: 基于协同治理视角[J]. *中国行政管理*, 2017(4): 35-39.
- [19] VAN ENGEN N, STEIJN B, TUMMERS L. Do consistent government policies lead to greater meaningfulness and legitimacy on the front line? [J]. *Public administration*, 2019, 97(1): 97-115.
- [20] YIP W, FU H, CHEN A T, et al. 10 years of health-care reform in China: progress and gaps in Universal Health Coverage[J]. *The lancet*, 2019, 394(10204): 1192-1204.
- [21] 顾昕. “健康中国”战略中基本卫生保健的治理创新[J]. *中国社会科学*, 2019(12): 121-138.
- [22] 张兮, 石焱文, 杜万里, 等. 基于 PMC 指数模型的数据要素政策量化评价及优化路径[J]. *管理评论*, 2024, 36(12): 3-14.
- [23] 杨斌. 地方政府的政策整合如何促进政策有效执行: 云南省 S 县政策“组合拳”的案例研究[J]. *公共管理学报*, 2024, 21(2): 128-140.
- [24] 任弢, 宋王杰. 村庄发展政策何以连续? 全过程人民民主视阈下的一种解释[J]. *西北大学学报(哲学社会科学版)*, 2025, 55(1): 117-129.
- [25] 李砚忠. 理解中国经济奇迹: 基于政策连续性和制度弹性的视角[J]. *经济社会体制比较*, 2016(2): 174-183.
- [26] 姚东旻, 董鸿儒, 徐傲晗. 财政补贴面向家庭还是面向企业: 基于存量—流量一致模型的系统评估[J]. *中国社会科学*, 2025(5): 72-91.
- [27] 李玲. 探索中国式医疗保障制度: 三明医改实践[J]. *社会保障评论*, 2024, 8(2): 12-23.
- [28] PAGE M J, MCKENZIE J E, BOSSUYT P M, et al. The PRISMA 2020 statement: an updated guideline for reporting systematic reviews[J]. *BMJ*, 2021, 372: 71.
- [29] BLEI D M, LAFFERTY J D. Dynamic topic models[C]// *Proceedings of the 23rd international conference on Machine learning*. New York, NY: Association for Computing Machinery, 2006: 113-120.
- [30] ZIEMS C, HELD W, SHAIKH O, et al. Can large language models transform computational social science? [J]. *Computational linguistics*, 2024, 50(1): 237-291.
- [31] 金星晔, 左从江, 方明月, 等. 企业数字化转型的测度难题: 基于大语言模型的新方法与新发现[J]. *经济研究*, 2024, 59(3): 34-53.
- [32] LAI Y, ZHAO H. Comparative analysis of smart city scientific research trends in the USA and China[J]. *Nature cities*, 2025, 2(9): 875-883.
- [33] RIJCKEN E, ZERVANOU K, MOSTEIRO P, et al. Topic specificity: a descriptive metric for algorithm selection and finding the right number of topics [J]. *Natural language processing journal*, 2024, 8: 100082.
- [34] MILSTEIN R, SCHREYOEGG J. Pay for performance in the inpatient sector: a review of 34 P4P programs in 14 OECD countries[J]. *Health policy*, 2016, 120(10): 1125-1140.
- [35] TA Y, ZHU Y, FU H. Trends in access to health services, financial protection and satisfaction between 2010 and 2016: has China achieved the goals of its health system reform? [J]. *Social science & medicine*, 2020, 245: 112715.