



图书馆论坛

Library Tribune

ISSN 1002-1167, CN 44-1306/G2

## 《图书馆论坛》网络首发论文

题目：高价值专利影响因素分析及主题挖掘——以纳米药物领域为例  
作者：奉国和，刘任铎，邓伟伟  
收稿日期：2023-07-04  
网络首发日期：2023-09-18  
引用格式：奉国和，刘任铎，邓伟伟. 高价值专利影响因素分析及主题挖掘——以纳米药物领域为例[J/OL]. 图书馆论坛.  
<https://link.cnki.net/urlid/44.1306.G2.20230915.1247.002>



**网络首发：**在编辑部工作流程中，稿件从录用到出版要经历录用定稿、排版定稿、整期汇编定稿等阶段。录用定稿指内容已经确定，且通过同行评议、主编终审同意刊用的稿件。排版定稿指录用定稿按照期刊特定版式（包括网络呈现版式）排版后的稿件，可暂不确定出版年、卷、期和页码。整期汇编定稿指出版年、卷、期、页码均已确定的印刷或数字出版的整期汇编稿件。录用定稿网络首发稿件内容必须符合《出版管理条例》和《期刊出版管理规定》的有关规定；学术研究成果具有创新性、科学性和先进性，符合编辑部对刊文的录用要求，不存在学术不端行为及其他侵权行为；稿件内容应基本符合国家有关书刊编辑、出版的技术标准，正确使用和统一规范语言文字、符号、数字、外文字母、法定计量单位及地图标注等。为确保录用定稿网络首发的严肃性，录用定稿一经发布，不得修改论文题目、作者、机构名称和学术内容，只可基于编辑规范进行少量文字的修改。

**出版确认：**纸质期刊编辑部通过与《中国学术期刊（光盘版）》电子杂志社有限公司签约，在《中国学术期刊（网络版）》出版传播平台上创办与纸质期刊内容一致的网络版，以单篇或整期出版形式，在印刷出版之前刊发论文的录用定稿、排版定稿、整期汇编定稿。因为《中国学术期刊（网络版）》是国家新闻出版广电总局批准的网络连续型出版物（ISSN 2096-4188，CN 11-6037/Z），所以签约期刊的网络版上网络首发论文视为正式出版。

\*本文系 2020 年广州市科技计划项目“突发词探测理论、方法与应用研究”（项目编号:202002030384）研究成果。

## 高价值专利影响因素分析及主题挖掘 ——以纳米药物领域为例\*

奉国和 刘任铎 邓伟伟

**摘 要** 挖掘高价值专利影响因素及主题，有利于政府和企业进行前瞻性专利战略部署，对抢占技术先机、赢取竞争优势具有积极意义。文章首先依托专利数据，采用因子分析法计算专利综合价值评分，依据 2-8 原则筛选高价值专利；然后对专利指标及其分类标签进行特征分析、差异性分析以及回归分析，探究影响专利价值的因素；最后使用 BERTopic 主题模型挖掘高价值专利主题。选取“纳米药物”专利数据进行实证分析，结果表明该方法能有效探究出专利价值的强影响因素，筛选出的高价值专利与其它方法获得结果高度一致。

**关键词** 高价值专利；影响因素；主题挖掘；BERTopic；纳米药物

**引用本文格式** 奉国和,刘任铎,邓伟伟.高价值专利影响因素分析及主题挖掘——以纳米药物领域为例[J].图书馆论坛, 2024

## A Study on the Factors Influencing of High-value Patent and Topic Mining in the Nanodrugs Area

Feng Guohe& Liu Renhua& Deng Weiwei

**Abstract** Exploring the influencing factors and themes of high-value patents is beneficial for the government and enterprises to carry out forward-looking patent strategy deployment, and has positive significance in seizing technological opportunities and winning competitive advantages. The article first relies on patent data, uses factor analysis to calculate the comprehensive value score of patents, and selects high-value patents based on the 2-8 principle; Then conduct feature analysis, difference analysis, and regression analysis on patent indicators and their classification labels to explore the factors that affect patent value; Finally, use the BERTopic theme model to mine high-value patent themes. Selecting patent data for "Nanodrugs" for empirical analysis, the results show that this method can effectively explore the strong influencing factors of patent value, and the high value patents screened are highly consistent with the results obtained by other methods.

**Keywords** High-value patent; Influencing factors; Topic mining; BERTopic; Nanodrugs

### 0 引言

目前我国每万人高价值发明专利拥有量达 9.4 件，发明专利有效量位居世界第一<sup>[1]</sup>，但是专利质量不高，切实解决产业发展瓶颈问题的核心专利依然不足。提高专利质量已成为关注重点，专利价值则是提升专利质量的关键。“十四五”规划和 2035 年远景目标纲要<sup>[2]</sup>提出将“每万人口高价值发明专利拥有量”纳入经济社会发展主要指标，鼓励各地探讨培育、保护和激励高价值专利。深入研究专利价值影响因素及主题对国家（地区）、企业实施专利战略布局具有重要意义。本文依托纳米药物专利数据，首先选取专利指标，利用因子分析构建高价值专利评价指标体系，依据 2-8 原则筛选高价值专利，并与其它方法对比；然后对高价值专利各项指标及其分类标签进行特征分析、差异性分析以及回归分析，探究影响专利价值的强因素；最后通过 BERTopic 主题模型挖掘高价值专利主题并进行验证。

### 1 研究回顾

Sanders<sup>[3]</sup>首次提出“专利价值”概念，指出仅 55% 专利被实施且各专利间的商业价值差异大。专利价值具有内外两个维度，内在价值主要指专利的技术价值如技术先进性、技术难度或复杂度等<sup>[4]</sup>及授权后所具有的法律价值如权利要求数、同族数、被引证数等<sup>[5]</sup>。外在价值主要指专利潜在市场能力，与专利市场价值、经济价值、新颖性、实用性、专利寿命等相关<sup>[6]</sup>。当前专利价值评估兼顾内外在价值，即综合考量专利的技术价值、法律价值、市场价值、经济价值等<sup>[7]</sup>。下面从评估指标、评估方法两方面对专利价值评估进行梳理。

## 1.1 专利价值评估指标

(1) 单一指标专利价值评估。指仅使用一个指标来评价专利价值的方法,其中被引证数常使用<sup>[8]</sup>,该方法也受到质疑,兰德公司研究报告指出专利的被引用与其价值之间并非线性,二者之间的正相关性不理想<sup>[9]</sup>。专利寿命<sup>[10]</sup>、专利授权率<sup>[11]</sup>、同族数<sup>[12]</sup>、权利要求数<sup>[13]</sup>等也常使用。该方法简单,随着专利技术发展,技术复杂度提高,专利潜在市场能力被发掘,它无法全面反映专利价值。

(2) 多指标综合评估。该方法具有规范性强且模型化程度高的特点,通过构建分层级的价值指标体系全面评估专利价值。CHI Research 与美国国家科学基金会首次提出专利数量、被引证数、技术生命周期等 7 项经典的专利价值评估指标综合评估地区、企业整体的专利价值;Harhoff 等<sup>[14]</sup>选取专利范围、同族数等指标评估专利价值;Park<sup>[15]</sup>选取多个涉及技术内在特征和技术使用情况的相关指标构建专利价值评估指标体系;李清海<sup>[16]</sup>提出多指标专利价值评价的层次结构并基于综合性与专业性总结出被引证数、权利要求数、专利族大小、专利寿命等 9 项指标;李春燕<sup>[17]</sup>首次提出“专利质量指标”的概念,从普遍认可、专利质量、科学指标等六大类筛选出 29 项细分指标作专利价值评价。研究围绕技术、法律、市场、经济、战略、管理等形成三因素、四因素<sup>[18-19]</sup>、五因素专利价值评估指标体系<sup>[20-23]</sup>,多因素指标组合是当前专利价值评估的主流方法,它降低单一指标评估的片面性和主观性,但依然面临指标设置冗余、指标权重设置不合理等问题。

## 1.2 专利价值评估方法

(1) 定性评估方法。LS 模型为专利价值经典定性分析模型<sup>[21]</sup>,美国知识产权咨询公司专利记分牌方法<sup>[22]</sup>、欧洲 IPScore 系统、我国新三板专利评估法等属于此类。情景分析<sup>[23]</sup>、TRIZ<sup>[24]</sup>等定性评估方法也被用于专利价值评估。定性评估需要投入大量的时间和资源进行专利文件研究和市场调研,随着专利技术不断发展它难以全面评估专利价值。

(2) 定量评估方法。一是综合评价,主要有层次分析法、因子分析法<sup>[25]</sup>、主成分分析法、模糊评价法等。层次分析法最常用<sup>[26]</sup>,传统层次分析法并未考虑人类思维的模糊性而影响判断结果的准确性,模糊理论可弥补这一缺陷<sup>[27]</sup>,出现很多改进方法<sup>[28]</sup>。二是四机器学习,伴随着机器学习理论及技术不断成熟,决策树等特征选择算法、支持向量机等分类算法被用于专利价值研究<sup>[29-32]</sup>。机器学习方法可靠性高、科学性强,但也有缺陷,如它需要一个特征向量作为输入,特征选择困难,不同类型的专利需要不同的特征;或者在计算文本相似度时,由于文本独特性和所在领域单一性,无法建立统一的价值评估体系。三是语义分析,对专利文本进行语义分析评判专利价值。其中狄利克雷分布模型(LDA)应用广泛,它在词袋模型的基础上计算各主题和词分布概率<sup>[30]</sup>。然而该模型并未考虑单词、主题之间的语义关系,主题个数不清晰,主题含义不清晰,识别效果欠佳。Bert 模型<sup>[31]</sup>是一种深度双向的、无监督的语言表示法,它仅使用纯文本语料库进行预训练,能够大幅提升现有语义表示方法的性能,并应用于专利价值分析中<sup>[32-33]</sup>。

## 1.3 述评与设想

单一指标无法全面反映专利价值,多指标组合又存在指标设置冗余、指标权重不合理等问题,LDA 主题模型识别效果欠佳。鉴于此,本文拟首先使用因子分析多指标降维分析并依据结果进行主成分权重分析;其次使用 LightGBM 并结合特征分析和差异分析探究高价值专利的影响因素;最后使用 BERTopic 挖掘高价值专利主题并验证。

## 2 基本理论与研究设计

### 2.1 基本理论与方法

(1) 因子分析综合评价。综合评价在很多领域应用广泛<sup>[34]</sup>, 就专利价值评价而言主要有因子分析、熵权 TOPSIS、层次分析、突变级数等。其中因子分析具有一定的优势, 首先, 它将原始指标降维, 减少数据复杂性和冗余性, 更好反映指标之间的内在关联和共性因素, 揭示潜在指标结构, 使得评价体系更简洁且易于解释。其它方法如熵权 TOPSIS 和突变级数法往往在处理多指标时难于考虑指标之间的内在关系且更易受异常值影响, 导致评价结果片面或不准确。其次, 它在降维时考虑各指标的权重, 而不是简单地对指标进行等权处理, 一定程度上避免因指标权重不同而导致信息损失, 使得降维后的维度更具代表性。再次, 它能够更准确地反映不同指标在专利价值评价中的贡献度。最后, 因子分析在主成分权重分析阶段进一步明确降维后维度的重要性, 为构建综合评价指标体系提供有力支持, 使评价过程更具可解释性, 充分揭示各维度在评价中的相对重要性, 可更准确地理解专利价值构成。基于此, 本研究选择因子分析进行降维并构建高价值专利评价指标体系。

(2) LightGBM 特征权重分析。LightGBM 在传统 GBDT 算法上采用基于直方图的决策树算法、单边梯度抽样算法、互斥特征捆绑算法、带深度限制的 Leaf-wise 叶子生长策略进行优化, 且提供 Split 与 Gain 两个度量标准衡量模型中特征重要性<sup>[35]</sup>。

(3) 主题挖掘。BERTopic<sup>[36]</sup>利用 BERT 嵌入和 c-TF-IDF 创建群集的主题建模技术, 采用统一的平面近似与投射技术, 减少文本的嵌入维数, 很容易对文本进行解析, 允许轻松解释主题, 在主题描述中保留重要词, 该算法包含 3 个阶段: 词嵌入、聚类<sup>[37,38]</sup>与主题表示<sup>[39]</sup>。相较于其他主题抽取模型(如 LDA), BERTopic 模型无需人为设置主题数量等超参数, 免去了复杂的参数尝试步骤<sup>[40]</sup>。

### 2.2 研究设计

#### 2.2.1 专利数据采集

本研究从大为 innojoy 专利数据库(<http://www.innojoy.com/search/home.html>)采集领域专利数据并进行预处理, 具体参见 3.1 节描述。

#### 2.2.2 专利价值评价

(1) 指标选取与预处理。参考前人成果, 结合“十四五”规划和 2035 年远景目标纲要中明确的高价值专利 5 个方面, 依据科学性和易获取性原则遴选出 23 个专利指标: 被引证数<sup>[41]</sup>、被审查员引证数(审查员引用是作为对比文件引用, 说明是审查过程中审查员认为较为接近的技术)、权项数(权利要求数)、独权数(独立权利要求数, 又称主权项)、主权项字数<sup>[42]</sup>、说明书页数、同族数、布局国家/地域数(考察专利海外布局能力)、三方专利(美日欧)(指在美国、日本和欧洲专利局三方都提出专利申请并至少在美国专利商标局获得发明专利权的同一项专利)、引证专利数、引用专利国别/地域数、引证非专利文献数、存活期(专利年龄, 指专利的持续时间)、剩余有效期、申请年、新兴产业分类数<sup>[43]</sup>、IPC 部数、IPC 小类数<sup>[44]</sup>、发明人数、无效次数、许可次数、转让次数、质押次数。

上述指标除三方专利(美日欧)是分类变量外, 其余均为连续变量, 对连续变量进行最大最小归一化处理消除量纲影响, 分类变量则进行虚拟哑变量化处理。

(2) 因子分析及综合价值得分。对预处理后的专利指标进行因子分析, 构建出高价值专利评价指标体系, 计算专利因子得分及综合价值得分。

(3) 高价值专利筛选。任何一组资源中, 最重要的只占约 20%, 其余 80%为多数且次要的, 该原则在各领域得到广泛认同<sup>[45]</sup>。专利价值领域也类似, 大部分高价值往往集中在少数专利上。因此基于此思想, 本研究计算专利综合价值得分并进行降序排序, 依据二



八原则，确定前 20% 专利为高价值专利。这些专利在前述多个指标上综合表现出色，可被认为具有较高价值的专利。

(4) 高价值专利验证。OECD 国际组织、Innography 专利数据库等陆续提出专利强度 (Patent Strength)、专利强度指数 (IP Strength Index)、IPQ 报告、Ocean Tomo 300th 等专利价值评估工具；国内也有专利价值度 (PVD)、合享价值度 (IncoPat)、大为专利指数 (DPI)、P2I (Patent to Intelligence) 等。本文设计方法与已有工具对比，计算高价值专利重叠率，以验证本方法的准确性和有效性。

### 2.2.3 影响因素分析

为分析高价值专利影响因素，首先分析专利指标均值、中位数、方差等。其次，以专利综合价值得分作为因变量，各专利指标为影响因素，并依据专利指标原始数据分布标注分类标签，将各指标划分为不同等级水平，构建多因素方差分析模型，探索各专利指标不同水平对专利价值是否存在显著差异。最后，利用 LightGBM 模型计算各特征对专利价值重要性。

### 2.2.4 高价值专利主题挖掘

采用 BERTopic 模型对高价值专利标题摘要进行主题挖掘。首先，提取高价值专利标题和摘要构建语料库。其次，进行预处理，包括分词和去除停用词，分词过程中使用相关领域专利关键词构建自定义词典。最后，基于 BERTopic 获得高价值专利主题。

### 2.2.5 结果验证

利用政策文本、文献、资讯等公开信息验证获得的主题，判定是否为该领域主要核心内容，定性评估模型结果的准确性。

## 3 实证分析

### 3.1 数据采集与预处理

选择纳米药物领域专利开展实证分析，依托大为 innojoy 专利数据库，确定检索式：TA = '纳米药物\*' or TA = 'nanomedicine\*' or TA = 'nanodrug\*' or TA = 'nanomaterial drug\*' or TA = 'nano-drug\*' or TA = 'nano medicine\*' or TA = 'nano drug\*'，检索时间为 2023 年 3 月 23 日，时间跨度为 2015-2022 年；剔除重复和不符合规范专利后，得到 2309 件专利。采集数据中无效次数、质押次数指标值大多为 0，剔除此两项指标，初步选定 21 个指标。对专利数据进行缺失值处理、最大最小归一化及虚拟哑变量化等预处理。

### 3.2 专利价值评价

#### 3.2.1 专利综合价值得分计算

(1) 因子分析可行性及因子提取。得到 KMO 值为 0.649，Bartlett 球形检验  $p$  值显著，因子分析可行。由探索性因子分析公因子方差及验证性因子分析因子载荷系数可知，发明人数、主权项字数、说明书页数、许可次数、转让次数五项指标的公因子方差分别为 0.254、0.323、0.346、0.006、0.069，申请年的标准化载荷系数为 0.169，剔除此六项指标。碎石图特征值曲线前六个公共因子数值均大于 1，因子旋转后提取的 6 个公因子的累计方差贡献率均为 85.152%，据此，选取六个公因子反映 15 个原始指标信息是可行的。

(2) 公共因子命名。表 1 为 15 个原始指标经正交旋转后生成的因子载荷系数矩阵。其中，因子载荷值代表每个原始指标被公共因子解释程度的大小，提取的公因子涵盖原始变量的信息越多，因子载荷的绝对值越大。从表 1 可知，第一个公因子上载荷较大的是同族数、布局国家/地域数、三方专利，主要考察专利的海内外布局能力，故将其命名为布局广泛；第二个公因子上具有高载荷的指标与专利应用范围相关，将其命名为应用范围；第三个公因子与专利引用相关，命名为知识密度；第四个因子与专利被引有关，反映专利的

引领情况，命名为引领水平；第五个公因子上载荷较高的指标与专利的年龄及剩余有效期有关，命名为权利时效；第六个因子与专利权保护相关，反映其权利的保护范围，命名保护范围。具体公因子见表 2。

表 1 因子载荷系数

	旋转后因子载荷系数						共同度(公因子方差)
	因子 1	因子 2	因子 3	因子 4	因子 5	因子 6	
同族数	0.969	-0.037	0.073	-0.013	-0.075	0.063	0.955
布局国家/地域数	0.975	-0.034	0.045	-0.014	-0.073	0.086	0.968
三方专利(美日欧)	0.89	-0.002	0.02	-0.01	-0.056	0.217	0.842
引证专利数	0.048	-0.028	0.913	0.063	0.084	-0.009	0.848
引用专利国别/地域数	-0.003	0.04	0.762	0.051	0.306	-0.012	0.679
引用非专利文献数	0.072	0.012	0.77	0.089	0.07	-0.002	0.611
存活期	-0.089	-0.03	0.177	0.22	0.918	-0.012	0.931
剩余有效期	-0.109	0.039	0.25	0.054	0.925	-0.033	0.936
新兴产业分类数	0.028	0.867	0.019	-0.022	0.005	0.087	0.761
IPC 部数	-0.04	0.93	0.008	-0.031	-0.023	-0.02	0.868
IPC 小类数	-0.054	0.918	-0.006	0.002	0.031	0.043	0.848
被引证数	0.003	-0.027	0.118	0.983	0.117	0.022	0.996
被审查员引证数	-0.035	-0.025	0.085	0.983	0.133	0.017	0.994
权项数	0.324	0.045	-0.024	-0.034	-0.107	0.778	0.725
独权数	0.045	0.06	0	0.062	0.051	0.894	0.811

表 2 公因子载荷指标

因子	因子 1	因子 2	因子 3	因子 4	因子 5	因子 6
荷载成分	(1) 同族数 (2) 布局国家/地域数 (3) 三方专利(美日欧)	(1) 新兴产业分类数 (2) IPC 部数 (3) IPC 小类数	(1) 引证专利数 (2) 引用专利国别/地域数 (3) 引用非专利文献数	(1) 被引证数 (2) 被审查员引证数	(1) 存活期 (2) 剩余有效期	(1) 权项数 (2) 独权数
公共因子命名	布局广泛	应用范围	知识密度	引领水平	权利时效	保护范围

(3) 计算因子得分。表 3 为各成分所含因子得分系数，据此可计算各主成份分值。依据计算公式  $\frac{\text{方差解释差}}{\text{旋转后累积方差解释}}$ ，得到布局广泛、应用范围、知识密度、引领水平、权利时效、保护范围权重分别为 22.092%、19.362%、16.691%、15.708%、14.598%、11.548%，因子分别标识为  $F_1$ 、 $F_2$ 、 $F_3$ 、 $F_4$ 、 $F_5$ 、 $F_6$ 。根据公式 (1) 计算综合得分，并作最大最小归一化处理，使其值落在 [0,100] 区间内。

$$F = 0.22092 * F_1 + 0.19362 * F_2 + 0.16691 * F_3 + 0.15708 * F_4 + 0.14598 * F_5 + 0.11548 * F_6 \quad (1)$$

表 3 成分矩阵表

名称	成分					
	成分 1	成分 2	成分 3	成分 4	成分 5	成分 6
同族数	0.29	-0.012	0.029	-0.008	-0.062	0.06
布局国家/地域数	0.292	-0.011	0.018	-0.008	-0.061	0.082
三方专利(美日欧)	0.267	-0.001	0.008	-0.006	-0.046	0.207
引证专利数	0.014	-0.009	0.364	0.038	0.07	-0.009
引用专利国别/地域数	-0.001	0.013	0.304	0.031	0.254	-0.011
引用非专利文献数	0.022	0.004	0.307	0.053	0.058	-0.001
存活期	-0.027	-0.01	0.07	0.131	0.763	-0.011
剩余有效期	-0.033	0.013	0.1	0.032	0.769	-0.032
新兴产业分类数	0.008	0.29	0.008	-0.013	0.004	0.083
IPC 部数	-0.012	0.31	0.003	-0.019	-0.019	-0.019
IPC 小类数	-0.016	0.306	-0.002	0.001	0.026	0.041
被引证数	0.001	-0.009	0.047	0.585	0.097	0.021
被审查员引证数	-0.011	-0.008	0.034	0.586	0.11	0.016
权项数	0.097	0.015	-0.01	-0.02	-0.089	0.742
独权数	0.013	0.02	0	0.037	0.042	0.853

(4) 筛选高价值专利。按照专利因子得分、综合得分进行降序排列，筛选前 20% 专利为高价值专利，如表 4。

表 4 2015-2022 年纳米药物专利综合价值分

序号	专利标题	专利价值得分
2173	一种 ROS 响应的纳米药物递送系统及其制备方法与应用	100
2151	卵巢癌特异靶向的生物可降解双亲性聚合物、由其制备的聚合物囊泡及应用	97.32
2183	生物可降解双亲性聚合物、由其制备的聚合物囊泡及在制备肺癌靶向治疗药物中的应用	92.33
1758	VAP 多肽及其在制备靶向诊疗肿瘤药物中的应用	92.12
1334	一种金属有机骨架类纳米药物及制备方法和应用	87.42
2099	一种线粒体靶向的纳米药物递送系统及其制备方法与应用	86.57
1235	用于治疗、成像和治疗应用的高度均匀的小纳米药物组合物	86.43
2292	一种硒化锡纳米材料及其制备方法和应用	85.80
2138	紫杉醇-油酸小分子前药自组装纳米粒的构建	85.55
2028	一种 TDNs-AS1411-核酸药物复合纳米材料载药系统及其制备方法	85.08
.....	.....	.....
1701	GRAPHENE BASE DRUG-ELUTING STENT.	0.13
2263	CARBONATE POLYMER WITH DISULFUR FIVE MEMBERED RING FUNCTIONAL GROUP ON SIDE CHAIN AND APPLICATION THEREOF	0

### 3.2.2 高价值专利对比验证

DPI 为大为专利价值综合评价指数,将价值由高到低依次设为 5 星、4.5 星、4 星.....1 星。以上文确定的检索式及检索条件进行检索,并筛选出 DPI 星级为 4 星及以上的专利共 438 条。将筛选出的高价值专利与 DPI4 星级以上的专利进行对比,重叠数 355 条,重叠率达 81.05%,如表 5,说明本方法是有效及准确的。

表 5 与 DPI 重叠高价值专利

专利标题	DPI 星级
一种两性嵌段共聚物及其制备方法和纳米胶束载药系统	5 星
阳离子纳米药物及其制备方法、载药植入医疗器械	5 星
生物可降解双亲性聚合物、由其制备的聚合物囊泡及在制备肺癌靶向治疗药物中的应用	5 星
纳米药物结合物	5 星
卵巢癌特异靶向的生物可降解双亲性聚合物、由其制备的聚合物囊泡及应用	5 星
.....	.....
Ce6 嵌入型红细胞膜包裹普鲁士蓝纳米颗粒的制备方法	4 星

## 3.3 高价值专利影响因素分析

### 3.3.1 描述性统计分析

对排名前 20% 高价值专利和其它专利各项指标作均值、方差、中位数分析,探索影响专利价值的因素。如表 6 所示,对高价值专利和其它专利进行对比,高价值专利各指标平均值均大于其它专利,高价值专利各指标中位数均大于或等于其它专利,尤其存活期和剩余有效期指标明显,说明专利的年龄及剩余有效期更能影响专利价值。除引用非专利文献数和剩余有效期外,高价值专利各指标方差均大于其它专利,可见高价值专利各指标的数据分布更分散,其它专利各指标方差大多为 0 或较低,数据分布更为集中。整体上,高价值专利各指标值高于其它专利,说明本方法是有效的。

表 6 专利指标特征值

专利指标	前 20% 高价值专利			后 80% 非高价值专利		
	平均值	中位数	方差	平均值	中位数	方差
同族数	0.074	0.043	0.016	0.068	0.043	0.01
布局国家/地域数	0.085	0.053	0.017	0.079	0.053	0.01
三方专利(美日欧)	0.061	0	0.057	0.041	0	0.039
新兴产业分类数	0.383	0.4	0.032	0.296	0.2	0.026
IPC 部数	0.42	0.5	0.026	0.353	0.25	0.023
IPC 小类数	0.456	0.5	0.024	0.378	0.333	0.022
引证专利数	0.146	0.121	0.016	0.069	0.061	0.006
引用专利国别/地域数	0.237	0.25	0.016	0.118	0.125	0.014

引用非专利文献数	0.027	0.023	0.001	0.013	0.006	0.002
被引证数	0.137	0.083	0.02	0.032	0	0.004
被审查员引证数	0.14	0.087	0.021	0.032	0	0.004
存活期	0.534	0.556	0.047	0.102	0	0.042
剩余有效期	0.745	0.789	0.036	0.19	0	0.123
权项数	0.082	0.074	0.011	0.074	0.074	0.004
独权数	0.061	0.04	0.009	0.04	0.04	0.001
样本量	461			1848		

### 3.3.2 多因素方差分析

将各专利指标值由大到小降序排列, 划分不同等级水平, 给出分类标签。以新兴产业分类数为例, 其数值有: 5、4、3、2、1、0, 则分别给出分类标签为 5、4、3、2、1、0, IPC 部数、IPC 小类数、引用专利国别/地域数三项指标与其标注规则一致。三方专利(美日欧)为分类变量, 其原始数据仅为是和否, 分别标注为 1 和 0。其余指标因其数值分布随机, 则划分成四个水平: 前 25%、25%-50%、50%-75%、75%-100%, 并分别标注为 3、2、1、0。将各专利指标其分类标签作为自变量, 各专利的综合价值得分作为因变量, 构建多因素方差分析模型, 得出了多因素方差分析结果。

从表 7 可知, 利用多因素方差分析指标不同水平值影响专利价值的差异, 模型  $R^2$  值为 0.933, 权项数、引用非专利文献数两项指标对专利价值影响不显著。近年专利申请书中的权利要求项数越来越多, 它可以帮助专利持有人获得攻击和先发制人的优势。但 Schettino 等<sup>[46]</sup>发现过多的权利要求会增加审查员负担, 导致其无法完全准确清除所有不合格的权利要求。而且, 过多的权利要求包含不必要的权利声明, 使发明的披露不够清晰, 难以理解, 甚至造成非故意性的侵权, 权项数过多专利价值未必就越高。其余指标  $p$  值均小于 0.05, 对专利价值影响显著。F 是组间均方和组内均方的比值, 值越大组间差异越大。存活期的 F 值明显大于其他指标, 说明不同水平值对专利价值影响更显著。

表 7 多因素方差分析

专利指标分类标签	平方和	df	均方	F	$p$
同族数-分类标签	1074.937	3	358.312	13.255	0.000**
布局国家/地域数-分类标签	1074.937	3	358.312	13.255	0.000**
三方专利(美日欧)-分类标签	6674.056	1	6674.056	246.898	0.000**
新兴产业分类数-分类标签	3229.910	5	645.982	23.897	0.000**
IPC 部数-分类标签	500.245	4	125.061	4.626	0.001**
IPC 小类数-分类标签	2260.422	6	376.737	13.937	0.000**
引证专利数-分类标签	392.070	3	130.690	4.835	0.002**
引用专利国别/地域数-分类标签	3304.062	7	472.009	17.461	0.000**
引用非专利文献数-分类标签	119.160	3	39.720	1.469	0.221
被引证数-分类标签	1074.937	3	358.312	13.255	0.000**
被审查员引证数-分类标签	537.469	3	179.156	6.628	0.001**
存活期-分类标签	62479.292	3	20826.431	770.447	0.000**
剩余有效期-分类标签	9758.553	3	3252.851	120.335	0.000**
权项数-分类标签	192.780	3	64.260	2.377	0.068
独权数-分类标签	2207.326	3	735.775	27.219	0.000**
残差	61091.500	2260	27.032		
$R^2=0.933$					
* $p<0.05$ ** $p<0.01$					

### 3.3.3 影响专利价值的特征重要性分析

以各专利指标为影响因素, 综合价值得分作为因变量, 设置 LightGBM 回归模型参数如下: 基学习器数量为 20, 学习率(learning\_rate)为 0.1, 树的最大深度(max\_depth)为 10, 进行影响专利价值的特征重要性分析, 图 1 表明被引证数、引证专利数、新兴产业分类数、存活期的重要性均大于或等于 10%, 即它们是高价值专利的重要影响因素。被引证数、引证专利数与 LS 模型指标相一致; “十四五”规划纲要从 5 个方面明确了高价值专利,



其中第一个方面即“战略性新兴产业的有效发明专利”，再次验证新兴产业分类是专利价值的重要影响因素；存活期则与前述特征分析及多因素方差分析结果相同。此外，引用非专利文献数、权项数的特征重要性均不超过 5%，与之前的理论分析一致。

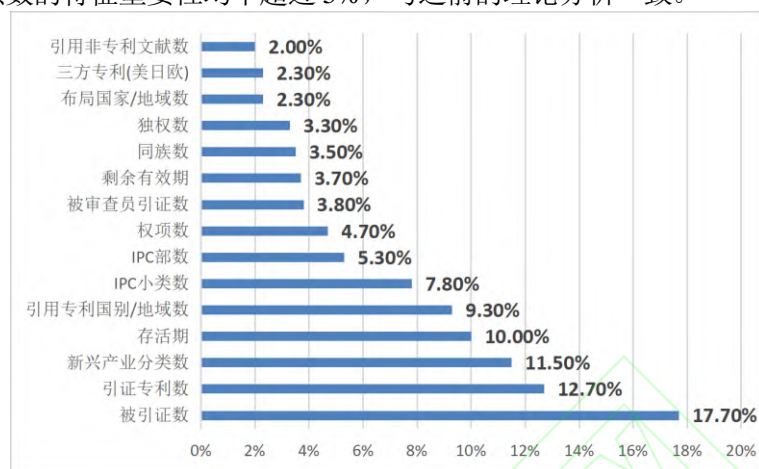


图 1 指标重要性排列

### 3.4 高价值专利主题挖掘

采集高价值专利标题和摘要数据。首先进行分词和停用词处理，使用纳米药物领域论文关键词构建自定义词典，添加到中文分词工具 jieba 中。然后利用 BERTopic 模型挖掘主题。将 nr\_topics 值设置为"auto"，由 BERTopic 自动迭代生成，以确保主题之间的相似度合适。得到 9 个主题，生成交互图谱（图 2），每个圆圈代表一个主题，其大小为该主题在所有文档中出现的频率，相似的主题在坐标系中的距离较近。从图 2 看出，9 个主题较为分散，并以条形图的形式直观地展示了各主题的特征词及其概率大小，如表 8 所示。

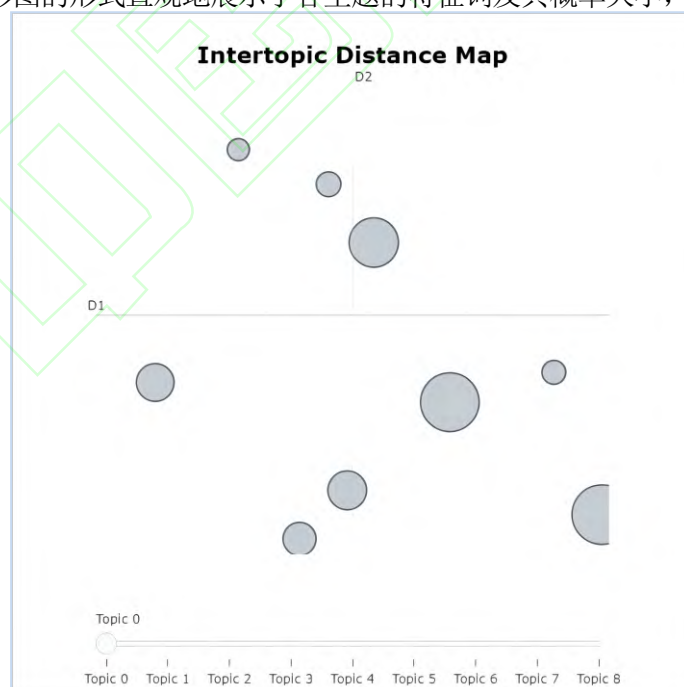


图 2 主题分布图

表 8 主题特征词分布

主题	特征词	概率分布	主题	特征词	概率分布	主题	特征词	概率分布
Topic 0	聚合物	0.0302	Topic 1	水凝胶	0.0372	Topic 2	光敏剂	0.0598
	药物载体	0.0299		美登素	0.0364		荧光	0.0434
	癌细胞	0.0268		聚碳酸酯	0.0276		药物载体	0.0347
	顺铂	0.0265		前药	0.0266		纳米颗粒	0.0329
	脂质体	0.0208		胶束	0.0261		光热治疗	0.0259
Topic 3	脱氧核糖核酸	0.0877	Topic 4	溶液的酸碱度	0.0418	Topic 5	聚乙二醇	0.1592
	序列化共轭聚合物	0.0837		纳米粒子	0.0386		囊泡	0.0511
	一氧化氮	0.0793		层状双氢滑石	0.0368		DNA 石墨烯复合物	0.0456
	疾病诊断	0.0774		果胶	0.0313		纳米探针	0.0416
	纳米抗体	0.0712		纳米胶束	0.0311		纳米钻石	0.0393
Topic 6	聚酰胺胺	0.0555	Topic 7	衍生物	0.0681	Topic 8	石墨烯	0.1643
	粘附肽序列 Arg-Gly-Asp	0.0531		叶面	0.0566		介孔磷灰石	0.1496
	牛血清白蛋白	0.0446		甜菜碱	0.0553		药物载体	0.1373
	T 细胞	0.0439		乙醇胺	0.0547		氧化石墨烯	0.1318
	核定位信号	0.0408		自由基	0.0531		溶液的酸碱度响应性	0.1288

各主题具体内容如下：

(1) Topic 0 (磁性纳米脂质体)。它是一种具有靶向性的药物载体系统，其制备方法包括用二硫键交联单体合成聚合物，并使用磁性纳米颗粒和脂质体来包装药物（如顺铂）。这种药物载体具有优异的靶向性，对癌细胞精确定位并释放药物，最大限度减少毒副作用。

(2) Topic 1 (双功能聚合物纳米胶束)。该技术采用疏水性和两亲性的聚合物为载药系统，并将药物包裹在水凝胶或聚碳酸酯的胶束中，胶束表面被修饰为可使药物在体内释放的前药。此外，这种技术还采用美登素和聚乙二醇来增强纳米粒子的稳定性和药物的生物利用度，这种双功能聚合物纳米胶束通过调整药物在体内的释放速率实现精确治疗。另外，该技术具有高度的灵活性，可根据具体需要调整胶束，被用于输送其他类型的药物和治疗其他类型的疾病。

(3) Topic 2 (光热协同纳米药物治疗)。它将光敏剂和化疗药物结合到纳米载体上，通过纳米颗粒的被动靶向作用，将药物和光敏剂一起输送到肿瘤组织中。在激光照射下，光敏剂产生荧光和热能，使纳米颗粒局部升温，同时激活光敏剂，破坏肿瘤细胞，实现光热治疗，化疗药物也被释放出来杀死癌细胞。该技术可以提高治疗效果，减少药物剂量和副作用，并降低肿瘤的复发率。它将金属有机框架用作药物载体，具有高度可控的结构和孔隙，可容纳药物和光敏剂，具有良好的生物相容性和药物释放能力。

(4) Topic 3 (DNA 纳米机器人)。该技术利用 DNA 纳米结构构建出可编程的分子机器人，能检测和治疗疾病。具体地，DNA 纳米机器人可以利用 DNA 序列化共轭聚合物，此共轭聚合物与 DNA 纳米结构上的靶向配体相互作用，实现针对性的药物递送和检测。DNA 纳米机器人还可以利用一氧化氮等分子作为信号，根据一氧化氮的浓度释放药物，实现疾病的诊断和治疗。此外，DNA 纳米机器人也可以利用纳米抗体来实现靶向诊断和治疗。

(5) Topic 4 (pH 响应型/pH 敏感型聚合物胶束)。该技术利用纳米粒子和聚合物胶束的优点，结合了 pH 响应性的特性，具有精准的药物传递和释放特性。其基本原理是，在特定的 pH 值下，聚合物胶束能够自组装成为一定大小的纳米粒子，形成纳米胶束。纳米胶束进入人体内部，当遇到具有不同 pH 值的不同环境时，会发生自组装和解离，释放出药物。药物利用此技术能精确地传递到病变组织，且在达到目标位置时以更精确的方式释放药物。该技术还可以通过改变聚合物胶束的成分和结构，以及调节 pH 响应范围来控制药物释放速度和剂量。该技术中广泛应用的载体是低分子量的果胶和层状双氢滑石，它具有良好的生物相容性和生物降解性及较高的药物包载能力和药物释放性能。

(6) Topic 5 (纳米钻石石墨烯量子点双功能递药系统)。该技术利用纳米钻石和石墨烯量子点的独特性质，将药物包裹在纳米粒子表面上，并通过聚乙二醇 (PEG) 等表面修

饰物质来增强稳定性和生物相容性。这种纳米粒子具有高度可控性和特异性,能通过囊泡或 DNA 石墨烯复合物等将药物精确地输送到病灶部位提高药物的治疗效果。此外,纳米钻石和石墨烯量子点还具有好的光学和电学性质,可作为纳米探针用于生物成像和诊断。

(7) Topic 6 (外泌体纳米颗粒技术)。外泌体是一种直径约 30-150 纳米的细胞分泌小囊泡,其表面含有多种受体和分子,被用于定向传递药物至目标细胞。通过将药物负载至外泌体表面或内部,提高药物的稳定性和生物利用度,同时降低药物的毒性和副作用。外泌体纳米颗粒通过表面修饰来实现细胞核靶向、T 细胞活化、药物释放等功能。此外,外泌体还可与其他纳米材料(如金属纳米粒子、碳纳米管等)结合,形成多功能复合纳米载体,进一步提高药物的传递效率和生物效应。

(8) Topic 7 (亲和型纳米药物缓释系统)。该技术将药物分子与亲和型的分子衍生物结合,使用缓释技术将其封装在纳米颗粒中,增强药物在体内的稳定性和效果。常用的载体材料包括硬脂酸、甜菜碱等,它们具有良好的生物相容性和生物降解性。同时,通过在载体表面修饰不同的叶面分子和自由基等功能分子实现靶向输送和缓释作用。乙醇胺为常用的叶面分子,通过修饰纳米粒子表面,增强与细胞膜的相互作用力,提高药物的靶向性。自由基可与靶向分子中的活性基团进行共价键结合,增强其与药物的结合力。

(9) Topic 8 (氧化石墨烯/介孔磷灰石复合物)。该复合物具有高比表面积、生物相容性好、化学稳定性强等特点,可以作为药物载体实现药物的高效递送。通过将石墨烯氧化得到氧化石墨烯并修饰表面,如引入 pH 响应性分子、稀土元素等增强药物递送性能。此外,通过核壳结构设计提高药物靶向性和控释性能,如将药物包裹在纳米金刚石核内,再将其包裹在氧化石墨烯壳层中实现药物的高效靶向递送和控释释放。

针对药物治疗过程中药物不易靶向、易受生物环境影响等问题,纳米药物技术受到广泛关注。利用纳米材料作为药物载体,构建出具有特定性质的纳米药物复合体,它通过控制药物与载体的比例和配方,实现药物精确负载和靶向释放。总之,纳米药物技术是一个不断发展的领域,其基于纳米材料的药物载体和靶向释放,有望为药物治疗提供更加精确和有效的方法。

### 3.5 高价值专利主题验证

利用政策文本、项目申报指南、科技资讯、相关文献等公开信息,定性评估本研究结果的准确性。

《国家中长期科学和技术发展规划纲要(2006—2020 年)》(以下简称纲要)重点部署纳米研究等四项重大科学研究计划,重点研究纳米材料的可控制备、自组装和功能化,纳米材料的结构、优异特性及其调控机制,纳米生物学和纳米医学,生物分子的光、电、磁学性质及信息传递,纳米材料和纳米技术在信息、医药等领域的应用。纳米胶束的制备(Topic 4)应用了纲要中提到的自组装技术。

2020、2021、2022 年科学技术部分别发布“纳米前沿”重点专项项目申报指南,三份申报指南从纳米尺度出发,围绕纳米物质呈现出的新奇特特性,开展新型纳米材料和器件制备方法、纳米尺度表征新技术等方面的研究,提供多个纳米领域申报主题,其中包括医用磁性微纳器件及技术(Topic 0,磁性纳米脂质体)、纳米结构光学功能设计及其高灵敏增强光谱应用(Topic 2,光热协同纳米药物治疗)、复合热管理材料制备及应用(Topic 2,光热协同纳米药物治疗)、智能型微纳米机器人(Topic 3,DNA 纳米机器人)、仿生纳米药物递送体系、微环境响应纳米生物材料(Topic 4,pH 响应型聚合物胶束)等。

《“十三五”生物产业发展规划》瞄准生物科学发展前沿,围绕生物产业发展技术支撑需求,大力推进生物制造产业创新体系建设,在人工生命体、纳米生物装置、基因组编辑、



仿生制造等方向实现颠覆性技术创新和应用。2022 年 4 月,国家发展改革委批复同意《粤港澳大湾区打造纳米产业创新高地建设方案》,它主要聚焦于纳米材料、纳米器件、纳米医疗、纳米能源与环境四大产业方向。粤港澳大湾区纳米产业创新高地首批发布的成果具有广泛应用场景和广阔市场空间,其中包括由广东粤港澳大湾区国家纳米科技创新研究院(“广纳院”)研发的具有世界领先水平的医用 DNA 纳米机器人。以上文件以及成果引领 DNA 纳米机器人技术(Topic 3)的发展。

《“十三五”医疗器械科技创新专项规划》指出先进医疗器械是健康保障体系建设的重要基础,是推进医学诊疗技术进步的主要动力,应积极推动全球医疗器械领域的创新突破加速演进,碳纳米管、石墨烯、材料组学等新一代材料和技术驱动生物医用材料向高性能、高生物相容等方向发展。《北京市“十三五”时期加强全国科技创新中心建设规划》积极推动以石墨烯、碳纳米管等为代表的纳米材料技术研发、海淀区石墨烯技术研发和产业应用、怀柔区纳米材料产业发展。石墨烯作为纳米钻石石墨烯量子点双功能递药系统(Topic 5)和氧化石墨烯/介孔磷灰石复合物(Topic 8)两项技术的纳米材料,在两项规划推动下得到长效持续发展。

浙江为贯彻落实《国务院办公厅关于促进医药产业健康发展的指导意见》,推进重点领域取得突破,在化学药物领域针对常见病、多发病和重大疾病,鼓励发展脂质体、脂微球、纳米制剂等新型注射给药系统,以及速释、缓释、微囊制剂,纳米制剂和多颗粒系统等口服调释给药系统。在医疗器械领域大力发展医用机器人、电刺激类、新能源类、新材料类等高性能诊疗设备和高价值医用耗材,精准给药系统等智慧医疗技术和产品。《“十三五”中医药科技创新专项规划》重点发展缓控释给药系统、靶向给药系统、基于新型纳米技术的释药系统等中药制剂新技术;开展缓控释、靶向、纳米给药系统等中药高端制剂的应用推广研究,整体提高中药制剂的国际竞争力。两项文件强调的缓释给药系统很好地指引亲和型纳米药物缓释系统(Topic 7)技术的发展。

2021 年 1 月 *Nature Reviews materials* 发表的 *Extracellular vesicles versus synthetic nanoparticles for drug delivery*<sup>[47]</sup>讨论外泌体纳米颗粒技术(Topic 6)载药和传统合成纳米粒载药的对比分析、面临挑战及未来展望。

2020 年 7 月 *Nano-Micro Letters* 发表的 *A Review on Nano-Based Drug Delivery System for Cancer Chemoimmunotherapy*<sup>[48]</sup>讨论了纳米药物递送系统用于肿瘤化学免疫治疗的最新研究进展。聚合物纳米胶束(Topic 1)被提及具有疏水内核、可同时装载多种药物等特点,并阐述了其作为抗癌药物纳米递送载体的研究。

张婷等<sup>[49]</sup>以德温特专利数据为研究对象,对医用纳米机器人进行分析,识别出磁性纳米粒子(Topic 0)、纳米机器人制备技术(Topic 3)等 5 项技术前沿。

王宇斐等<sup>[50]</sup>同样以纳米药物为主题依托德温特专利数据得出,在材料组成方面纳米粒、脂质体、双亲性纳米载体具有重要地位,聚乙二醇(PEG)依然是使用频率较高的高分子材料,与 Topic0、Topic1、Topic 6 三主题吻合。

上述资料调研有力支撑本研究结果,说明本方法有效。将上述分析整理为表 9。

表 9 资料解读与对应的高价值专利主题

资料	解读	高价值专利主题
“十四五”国家重点研发计划“纳米前沿”重点专项 2021 年度项目申报指南	研制可体内驻留达完整疗程的、由磁性纳米颗粒组装构建的磁性微纳器件;研究磁场调控微纳器件产生磁极化、磁热、磁力等电磁效应及与纳米尺度相关的新现象、新机制,以及对体内特定部位神经系统的调控规律。	Topic 0 (磁性纳米脂质体)
	设计和构筑新型纳米光学材料和结构、建立具有光学功能的纳米材料和结构的理性设计方法。	Topic 2 (光热协同纳米药物治疗)
“纳米科技”重点专项 2020 年度项目申报指南	智能型微纳机器人、创建仿生纳米药物递送体系	Topic 3 (DNA 纳米机器人)
	通过纳米复合材料的精准制备,调控热导、热容、热焓等热管理特性,研发满足隔热、换热、散热等不同应用需求的宏观尺度复合热管理材料。	Topic 2 (光热协同纳米药物治疗)



	针对乳腺癌、结直肠癌和胰腺癌等的转移复发,结合肿瘤转移发生发展的动力学、表现遗传学和代谢组学特征,提出了设计合成具有原位组装微环境响应等功能的纳米生物材料。	Topic 4 (pH 响应型聚合物胶束)
“纳米前沿”重点专项 2022 年度项目申报指南	研究面向精准医疗和环境监测等重大需求的纳米机器人技术,建立纳米机器人功能材料开发和功能集成的交叉科学研究方法。	
“十三五”生物产业发展规划	在人工生命体、纳米生物装置、基因组编辑、仿生制造等方向实现颠覆性技术创新和应用,重点在生物基材料的规模制造等方面实现关键技术与工艺突破。	Topic 3 (DNA 纳米机器人)
粤港澳大湾区打造纳米产业创新高地建设方案	粤港澳大湾区纳米产业创新高地首批发布的成果包括由广东粤港澳大湾区国家纳米科技创新研究院(“广纳院”)研发的医用 DNA 纳米机器人。	
国家中长期科学和技术发展规划纲要(2006—2020 年)	在其重大科学研究计划中,建议重点研究纳米材料的自组装及其调控机制。	Topic 4 (pH 响应型聚合物胶束)
“十三五”医疗器械科技创新专项规划	石墨烯等新一代材料和技术驱动生物医用材料向高性能、高生物相容等方向发展。	Topic 5 (纳米钻石石墨烯量子点双功能递药系统)、Topic 8 (氧化石墨烯/介孔碳复合)
北京市“十三五”时期加强全国科技创新中心建设规划	以石墨烯、碳纳米管为代表的纳米材料的重点研发和产业应用。	
浙江省人民政府办公厅关于加快推进医药产业创新发展的实施意见	希望在速释、缓释等口服调释给药系统中取得突破。	Topic 7 (亲和型纳米药物缓释系统)
“十三五”中医药科技创新专项规划	开展缓控释、靶向、纳米给药系统等中药高端制剂的应用推广研究,并重点发展缓控释给药系统、靶向给药系统、基于新型纳米技术的释药系统等中药制剂新技术。	
Extracellular vesicles versus synthetic nanoparticles for drug delivery (2021)	外泌体纳米颗粒载药和传统合成纳米粒载药的对比研究。	Topic 6 (外泌体纳米颗粒技术)
A Review on Nano-Based Drug Delivery System for Cancer Chemotherapy (2020)	聚合物胶束具有疏水内核、可同时装载多种药物等特点,并阐述了其作为抗癌药物纳米递送载体的研究。	Topic 1 (双功能聚合物纳米胶束)
基于专利引文网络的医用纳米机器人技术前沿识别研究(2023)	识别出磁性纳米粒子、纳米机器人制备技术等 5 项技术前沿。	Topic 0 (磁性纳米脂质体) Topic 3 (DNA 纳米机器人) Topic 6 (外泌体纳米颗粒技术) Topic 0 (磁性纳米脂质体) Topic 1 (双功能聚合物纳米胶束)
纳米药物国际研究态势分析(2020)	在材料组成方面,纳米粒、脂质体、双亲性纳米载体具有重要地位。	

## 4 结语

本文以“纳米药物”领域 2015-2022 年专利数据为分析对象,探测专利价值影响因素及高价值专利主题,通过因子分析、2-8 原则、权重分析、主题挖掘等方法,获得有价值的结果与结论。但本研究存在如下不足:(1)一些指标对评估专利价值很有意义,但其统计数值大部分为 0 或大部分为空,本文仅进行了简单的剔除;(2)指标对专利价值进行评判时,指标具有动态性,其权重并非一成不变;(3)对纳米药物领域缺乏深度了解,对纳米药物领域高价值专利及主题解释不充分。鉴于以上不足,本研究下一步的重点将放在专利指标空值处理和动态演进特征上,并结合相关领域专家意见和权威机构判定,对探测的高价值主题进行更全面的归纳总结。

## 参考文献

- [1]中国知识产权局.中国打击侵权假冒工作年度报告(2022)[R].北京:知识产权出版社,2022.
- [2]国务院,国家发展和改革委员会.“十四五”规划和 2035 年远景目标纲要[D].北京:中国发展出版社,2021.
- [3]SANDERS B S, ROSSMAN J, HARRIS L J. The economic impact of patents[J]. Patent, Trademark and Copyright Journal, 1958, 2 (2): 340-362.
- [4]BEKKERS R, BONGARD R, NUVOLARI A. An empirical study on the determinants of essential patent claims in compatibility standards[J]. Research Policy, 2011, 40(7): 1001-1015.
- [5]THOMA G. Composite value index of patent indicators: Factor analysis combining bibliographic and survey datasets[J]. World patent information, 2014, 38: 19-26.
- [6]GRIMALDI M, CRICELLI L. Indexes of patent value: a systematic literature review and classification[J]. Knowledge Management Research & Practice, 2020, 18(2): 214-233.
- [7]刘妍.专利价值评估研究综述与趋势展望[J].图书情报工作, 2022, 66(15): 127-139.
- [8]TRAJTENBERG M. A penny for your quotes: patent citations and the value of innovations[J]. The Rand journal of economics, 1990, 21(1): 172-187.
- [9]王功,吴新年.新兴技术识别方法研究综述[J].图书情报工作, 2020, 64(04): 125-135.
- [10]SCHANKERMAN M, PAKES A. Estimates of the value of patent rights in European countries during the post-1950 period[J]. The economic journal, 1986, 96(384): 1052-1076.
- [11]GRILICHES Z. Patent statistics as economic indicators: a survey[M]//R&D and productivity: the econometric evidence. University of Chicago Press, 1998: 287-343.
- [12]NEUHÄUSLER P, FRIETSCH R. Patent families as macro level patent value indicators: applying weights to

- account for market differences[J]. *Scientometrics*,2013,96(1):27-49.
- [13]LLANES G, TRENTO S. Patent policy, patent pools, and the accumulation of claims in sequential innovation[J]. *Economic Theory*,2012,50(3):703-725.
- [14]HARHOFF D, SCHERER F M, VOPEL K. Citations, family size, opposition and the value of patent rights[J]. *Research policy*,2003,32(8):1343-1363.
- [15]PARK Y, PARK G. A new method for technology valuation in monetary value: procedure and application[J]. *Technovation*,2004,24(5):387-394.
- [16]李清海,刘洋,吴泗宗,等.专利价值评价指标概述及层次分析[J].*科学学研究*,2007(2):281-286.
- [17]李春燕,石荣.专利质量指标评价探索[J].*现代情报*,2008(2):146-149.
- [18]HOU J L, LIN H Y. A multiple regression model for patent appraisal[J]. *Industrial Management & Data Systems*,2006,106(9):1304-1332.
- [19]刘勤,杨士淦,刘友华.高价值专利评估方法、存在问题及对策[J].*科技管理研究*,2022,42(4):147-152.
- [20]GRIMALDI M, CRICELLI L, ROGO F. Valuating and analyzing the patent portfolio: the patent portfolio value index[J]. *European Journal of Innovation Management*,2018,21(2):174-205.
- [21]胡元佳,卞鹰,王一涛.Lanjouw-Schankerman 专利价值评估模型在制药企业品种选择中的应用[J].*中国医药工业杂志*,2007(2):148-150.
- [22]曹晨,胡元佳.综合专利价值指数与药物经济价值的相关性研究[J].*中国医药工业杂志*,2011,42(7):560-562.
- [23]文豪,胡晓阳.基于质量评价模型的专利质押价值评估体系修正研究[J].*科技管理研究*,2022,42(21):168-175.
- [24]林静,连晓振,侯亮.面向主控式创新的高价值专利挖掘研究[J].*情报杂志*,2022,41(6):164-172,163.
- [25]刘勤,王少康,胡良龙,等.农机装备专利价值评估研究[J].*中国农业科技导报*,2017,19(5):86-91.
- [26]万小丽,朱雪忠.专利价值的评估指标体系及模糊综合评价[J].*科研管理*,2008,154(2):185-191.
- [27]张黎,李倩.基于直觉模糊层次分析法的专利质量模糊综合评价[J].*科技管理研究*,2019,39(7):85-92.
- [28]伊惠芳,吴红,马永新,等.改进的柔性专利价值评价方法——基于时域-领域双重视角[J].*情报杂志*,2019,38(3):53-60.
- [29]赵蕴华,张静,李岩,等.基于机器学习的专利价值评估方法研究[J].*情报科学*,2013,31(12):15-18.
- [30]BLEI D M, NG A Y, JORDAN M I. Latent dirichlet allocation[J]. *Journal of machine Learning research*,2003,3(4-5):993-1022.
- [31]DEVLIN J, CHANG M W, LEE K, et al. Bert: Pre-training of deep bidirectional transformers for language understanding[EB/OL].[2021-05-31].<https://arxiv.org/pdf/1810.04805.pdf>.
- [32]HUANG Y, FENG X, ZHU M. Research on Patent Technology Value Evaluation Method Based on Knowledge Graph[C]//2022 5th International Conference on Data Science and Information Technology (DSIT). IEEE,2022:1-7.
- [33]THOMPSON L, MIMNO D. Topic modeling with contextualized word representation clusters[J]. *arXiv preprint arXiv:2010.12626*,2020.
- [34]刘照德,詹秋泉,田国梁.因子分析综合评价研究综述[J].*统计与决策*,2019,35(19):68-73.
- [35]KE G, MENG Q, FINLEY T, et al. Lightgbm: A highly efficient gradient boosting decision tree[J]. *Advances in neural information processing systems*,2017,30.
- [36]GROOTENDORST M. BERTopic: Neural topic modeling with a class-based TF-IDF procedure[J]. *arXiv preprint arXiv:2203.05794*,2022.
- [37]MCLNNES L, HEALY J, MELVILE J. Umap: uniform manifold approximation and projection for dimension reduction [EB/OL].[2021-12-21].<https://arxiv.53yu.com/abs/1802.03426>.
- [38]MCINNEN L, HEALY J, ASTELS S. Hdbscan: hierarchical density based clustering[J]. *The Journal of Open Source Software*,2017,2(11): 205.
- [39]QAISER S, ALI R. Text mining: use of TF-IDF to examine the relevance of words to documents[J]. *International Journal of Computer Applications*,2018,181(1): 25-29.
- [40]张清慧,陈谊,武彩霞.基于词表示模型的领域文献数据可视分析方法[J].*图学报*,2022,43(4):685-694.
- [41]TRAJTENBERG M. A penny for your quotes:Patent citations and the value of inventions [J]. *Rand Journal of Economics*,1990(21):172 -187.
- [42]REITZIG M. Improving patent valuations for management purposes—validating new indicators by analyzing application rationales[J]. *Research policy*,2004,33(6-7):939-957.
- [43]国家统计局.战略性新兴产业分类(2018) [S].北京:中国统计出版社,2018.
- [44]LERNER J. The importance of patent scope:An empirical analysis [J]. *Rand Journal of*

Economics,1994,25(2):319-333.

[45]徐剑,黄秋月.“二八定律”在图书馆管理中的应用[J].中国图书馆学报,2007(5):106-108.

[46]SCHETTINO F, STERLACCHINI A, VENTURINI F. Inventive productivity and patent quality: Evidence from Italian inventors[J]. Journal of policy modeling,2013,35(6):1043-1056.

[47]WITWER K W, WOLFRAM J. Extracellular vesicles versus synthetic nanoparticles for drug delivery[J]. Nature Reviews Materials,2021,6(2):103-106.

[48]MU W, CHU Q, LIU Y, et al. A review on nano-based drug delivery system for cancer chemoimmunotherapy[J]. Nano-Micro Letters,2020(12):1-24.

[49]张婷,陈娟,杨满逸,等.基于专利引文网络的医用纳米机器人技术前沿识别研究[J].中国医疗设备,2023,38(1):122-126.

[50]王宇斐,马嘉虹,梁兴杰.纳米药物国际研究态势分析[J].中国科学:生命科学,2020,50(7):698-714.

**作者简介** 奉国和（通信作者，ghfeng@163.com），博士，华南师范大学经济与管理学院教授、硕士生导师；刘任铎，华南师范大学经济与管理学院硕士研究生；邓伟伟，博士，华南师范大学经济与管理学院特聘副研究员。

**收稿日期** 2023-07-04

（责任编辑：吴卫娟）