

基于慢病防治成果的科研人员研究聚焦度测度

钟 华¹ 范少萍 单连慧 安新颖

中国医学科学院北京协和医学院医学信息研究所 北京 100020

摘 要 目的：构建面向慢病防治成果的科研人员研究聚焦度多层级测度方法，为慢病防治科研机构的人才分类管理和团队优化提供方法学支持。方法：基于某心血管疾病研究机构 76 位研究人员 1900 项科研成果数据，通过集中度指数测度个人研究主题集中程度，构建研究主题一致性分层模型，分析聚焦度特征与团队差异。结果：76 位研究人员中，高度集中型（ $CR_3 \geq 0.8$ ）占比为 59.2%，相对集中型占比为 25.0%，分散型占比为 15.8%。全机构成果中，T1（团队方向层）占 53.6%，T2（领域交叉层）占 42.9%，T3（方向偏离层）占 3.5%。结论：本方法可有效识别慢病防治科研人员的研究方向分布特征，判断个人与团队匹配状况，以及和机构发展定位的偏离程度，为慢病防治领域的差异化人才管理和科研资源配置提供可量化的决策依据。

关键词 医学科研人员；研究聚焦度；慢病防治；主题一致性分层；卫生人才管理

Research on measuring the focus of scientists based on chronic disease prevention and control achievements

ZHONG Hua, FAN Shao-ping, SHAN Lian-hui, AN Xin-ying

Institute of Medical Information, Chinese Academy of Medical Sciences & Peking Union Medical College, Beijing 100020, China

【 Abstract 】 Objective : To construct a multi-level measurement method of researchers ' research focus for the achievements of chronic disease prevention and control, and to provide methodological support for talent classification management and team optimization of chronic disease prevention and control research institutions. Methods : Based on the data of 1900 scientific research achievements of 76 researchers in a cardiovascular disease research institution, the concentration index was used to measure the concentration degree of individual research topics, and a hierarchical model of research topic consistency was constructed to analyze the focus characteristics and team differences. Results : Among the 76 researchers, the proportion of highly concentrated ($CR_3 \geq 0.8$) was 59.2 % , the proportion of relatively concentrated was 25.0 % , and the proportion of dispersed was 15.8 % . Among the results of the whole organization, T1 (team direction layer) accounted for 53.6 % , T2 (field intersection layer) accounted for 42.9 % , and T3 (direction deviation layer) accounted for 3.5 % . Conclusion : This method can effectively identify the distribution characteristics of research directions of researchers in chronic disease

1 基金项目：科技创新 2030-“癌症、心脑血管、呼吸和代谢性疾病防治研究” 重大项目(项目编号：2023ZD0509700；课题编号：2023ZD0509703) 作者简介：钟华（1983 年—），女，副研究员，主要研究方向为医学科技管理与评价研究。E-mail: zhong.hua@imicams.ac.cn
通讯作者：安新颖。E-mail: an.xinying@imicams.ac.cn

10.12201/bmr.202610.00022V1

prevention and control, judge the matching status between individuals and teams, and the degree of deviation from the development orientation of institutions, and provide quantifiable decision-making basis for differentiated talent management and scientific research resource allocation in the field of chronic disease prevention and control.

【 Key words 】 Medical researchers; Research focus; Chronic disease prevention and control; Theme consistency stratification; Health talent management

医学科学研究日益专业化和细分化，是随着医学科技发展和知识体系迭代而形成的基本特征之一。随着学科不断细化，医学科研人员面临一个长期选择，是深耕一个领域以积累深度专长，还是在多个方向保持灵活以捕捉新的研究方向。这一个选择在慢病防治领域表现得尤为突出。癌症、心脑血管疾病、慢性呼吸系统疾病和代谢性疾病等重大慢病已成为威胁我国居民健康的主要公共卫生问题，其防治研究具有病程周期长、致病因素复杂、学科交叉性强等特点。在“健康中国 2030”战略引领下，慢病防治科研机构承担着明确的定向攻关任务，科研人员的方向聚焦不仅关系到个人学术生涯的发展路径，更直接影响机构能否形成可持续的慢病防治研究优势和国家战略目标的实现。《中国心血管健康与疾病报告 2023》显示，我国心血管病现患人数约 3.3 亿，防治形势严峻^[1]。本研究以心血管疾病领域作为慢病防治的切入点。心血管疾病在四大慢病中学科体系成熟、研究力量布局完备，具有一定代表性，此外本研究所构建的聚焦度测度方法也考虑了向肿瘤、代谢病、慢性呼吸系统疾病等其他慢病领域迁移的通用性。

从已有研究来看，研究方向是否聚焦在学术研究过程中发挥着重要作用。De Rassenfosse 等基于约 30000 名生物医学研究人员的长周期面板数据，以 PubMed MeSH 主题词为分类体系，逐词测算研究者论文中特定主题词占比相对于全球同期同一主题词占比的倍数，经标准化后取各主题词 z-score 分布的第 90 百分位数作为专业化程度的测度值，该值越高表明研究者在投入最多的若干主题上的聚焦强度越高。研究发现，专业化程度每增加一个标准差，论文被引次数提升约 25%，在学术生涯早期可达 75%，为研究方向聚焦的学术价值提供了实证依据^[2]。宋培彦等基于冰山模型将科研人员学术专长分为显性的主要领域与隐性的比较优势领域，提出了自身专注指数和同行相对指数的双维测度框架^[3]。Hladik 和 Renisio 基于布迪厄场域理论，通过主题组合将科研人员定位至三维知识空间，建立了从知识地图到知识生产者分析的理论工具^[4]。上述研究分别从实证验证和测度方法两个层面推进了对研究方向聚焦的认识，但这些研究发现要转化为慢病防治科研管理和科学决策的依据，还需要将聚焦度从抽象概念转化为可用于卫生决策的测度指标。

医学科研人员研究聚焦度测度的意义，一方面可以识别科研人员在慢病防治特定研究方向上的长期积累程度。Kwiek 等基于 32 万名科学家 50 年数据的研究表明，科研生产力具有极强的路径依赖性，早期形成的产出模式倾向于在整个职业生涯中保持稳定^[5]。如果能够在职业生涯早期识别出方向分散的倾向并加以引导，对慢病防治领域的人才梯队建设和长期储备具有重要价值。另一方面，可以评估个人与慢病防治团队的匹配状况。在重大慢病研究项目中，科研人员个人研究方向与团队核心定位之间可能存在偏差，偏差虽然可能带来跨方向的视角和方法创新，但也可能导致个人在团队中缺乏学术认同和协作效率。如果缺少可量化的测度工具，卫生管理和决策者只能依赖主观判断。此外，可以用于建立机构层面的慢病防治学科布局规划。将全机构科研人员的聚焦度进行汇总分析，可从国家慢病防治重大需求出发，分析心血管疾病、肿瘤、呼吸系统疾病、代谢性疾病等不同慢病

方向的人才储备、薄弱环节，以及哪些人员的产出偏离了机构在慢病防治中的核心定位和发展方向。

从方法论角度看，聚焦度的测度已有学者从浓度指数、信息熵到综合多样性指数的角度开展研究^[4]，但已有研究多集中在宏观学科层面或个人学术产出评价，还缺乏面向医学科研管理需求的可落地的测算方案。在“破四唯”“立新标”的政策背景下^[5]，科研成果评价正从重数量转向对研究内容与方向质量的精细化评价。聚焦度测度正是从研究内容维度提供量化依据的工具，对支撑卫生精细化管理和决策分析具有实际需求和重要意义。

针对上述不足，本研究以某心血管疾病研究机构为实证对象，基于 76 位科研人员成果数据，从研究主题集中度、个人与团队方向一致性、个人与机构定位一致性三个维度构建科研人员研究聚焦度的测度体系，并分析不同团队的聚焦模式差异与人才薄弱方向，可为慢病防治科研机构的成果评价从内容维度提供量化依据，并支撑人才分类管理和团队优化。

本研究以某心血管**疾病**研究机构的科研人员及其成果为评价对象。该机构设有 12 个研究团队，研究方向覆盖高血压、心力衰竭、动脉粥样硬化、心脑血管衰老、心血管代谢、心血管流行病学、主动脉与大血管疾病、冠心病与冠状动脉疾病、心律失常、遗传性心血管病、心血管免疫与炎症、睡眠呼吸暂停与肺血管、心血管诊断新技术等领域。

数据来源于科研管理部门提供的资料。一类是研究人员信息，包括 76 位科研人员填报的个人研究方向、所属研究团队、团队研究领域描述、职务职称、学历等字段。另一类是科研成果数据，涵盖学术论文、科研项目、专利、指南与共识、科技奖励、药物、器械、软著、成果转化。成果数据时间跨度为 2006 至 2026 年，共计 1900 项。经过成果标识去重、数据格式统一和人员姓名校验后，形成规范化分析数据集。

基于该机构的**部门设置**、团队的研究领域定位、科研人员填报的研究方向及相关领域科研成果，经与该机构科研管理部门及心血管领域专家进行反复三轮讨论，**归纳形成服务于该机构科研管理实践的分类编码体系（表 1）**。该体系涵盖 16 个研究方向，D01 至 D15 为心血管疾病研究领域，D16 用于标记不属于心血管领域的偏离成果，作为机构定位一致性的判别基准。

表 1 研究方向分类编码体系

编码	研究方向名称	编码	研究方向名称
D01	高血压	D09	冠心病与冠状动脉疾病
D02	心力衰竭与心肌病	D10	心律失常
D03	脂代谢与动脉粥样硬化	D11	遗传性心血管病
D04	心脑血管衰老	D12	心血管免疫与炎症
D05	心血管代谢	D13	睡眠呼吸暂停与肺血管
D06	主动脉与大血管疾病	D14	心血管诊断新技术
D07	心血管流行病学	D15	心血管病其他领域

10.12201/bmr.202610.00022V1

D08	心血管疾病预警与干预	D16	非心血管领域
-----	------------	-----	--------

对每项成果进行研究方向归属标注是后续聚焦度分析的基础。本研究采用方向词典加权自动分类与语义一致性人工复核相结合的双轮收敛标注方法。自动分类方法如基于深度学习的文本分类技术擅长处理大规模文本，但需依赖大量标注数据训练，其输出难以直接映射到科研管理所需的分类体系。自动分类方法如潜在狄利克雷分配（Latent Dirichlet Allocation, LDA）模型擅长处理大规模文本，但输出的主题聚类结果难以直接映射到机构管理所需的分类体系^[9]。Fu 和 Akbar 通过作者、出版物与关键词异质图社区发现方法构建专家画像，其输出同样难以直接服务于机构研究方向管理^[10]。而人工标注虽然精度高但成本过高。本研究的方向词典自动分类以 70%~75% 的基础精度覆盖全量数据，人工复核聚焦于约 23.5% 低置信度样本，从而保证关键样本的标注质量。

为每个研究方向构建三级主题词权重体系。核心词为该方向特有的高度特异性专业术语，如 D02（心力衰竭与心肌病）的核心词包括“射血分数”“心肌纤维化”“利钠肽”“心室重构”等。偏好词是该方向常见但非唯一的专业词汇，如 D02（心力衰竭与心肌病）的偏好词包括“心功能”“心脏超声”“运动耐量”等。通用词为可能涉及但跨方向也通用的词汇，如“心血管”“心梗预后”“危险因素”等。同时为每个方向设置排除词索引，明确列出虽含本方向通用词，但实际属于其他研究方向的术语，防止通用词引起的方向误判。例如 D06（主动脉与大血管疾病）的排除词包含“主动脉瓣狭窄”或“主动脉瓣关闭不全”，以避免心脏瓣膜病方向的泛化。对每项成果的标题、关键词和摘要进行加权打分。得分最高的方向标记为该成果的主要方向 D1，次高方向标记为 D2，第三高方向标记为 D3。

主题词词典经历六轮迭代进行修正。首轮以专家知识构建初始词典，基于分类结果的人工抽样审核发现系统性问题，调整特定方向的主题词和排除词，重新分类，再次审核。经过六轮迭代后，自动分类的总体精度从首轮的约 55%~60% 提升至约 70%~75%。引入团队先验规则作为辅助依据，即当成果的主题词与作者所属团队核心方向的专有词匹配时，该方向得分予以加权调整，以降低领域内交叉词汇的干扰。

自动分类的局限在于词语层面的匹配无法充分捕捉语义层面的研究方向归属。对于自动分类置信度较低的样本，启动人工复核。复核由两名具有心血管研究背景的评审人独立阅读成果的标题、关键词和摘要原文，结合研究人员所属团队方向信息，给出主要方向、次要方向、补充方向及置信度的判断。评审人间一致性系数 Kappa=0.86。分歧样本经讨论达成共识后标注。

采用自动分类和复核的合并策略，通过人工复核的成果以复核判断为准，未进入复核的成果以自动分类结果为准。复核共覆盖 330 项成果，占全部成果的 23.5%，最终形成每项成果的研究方向归属。

10.12201/bmr.202610.00022V1

在方向分类完成的基础上，构建研究主题一致性分层模型（Thematic Consistency Stratification, TCS），将每项成果按其研究主题与团队方向及机构定位的一致性程度划分为三个层级（表 2）。TCS 模型的设计逻辑是个人研究与团队方向一致性，并非绝对的二分变量。T2（领域交叉层）成果占比本身就是对心血管研究机构内部交叉合作程度的一项度量。T3（方向偏离层）承担偏离预警功能，当某个研究人员成果的 T3（方向偏离层）占比超过一定阈值时，提示其研究方向与机构核心定位之间可能出现了需要关注的偏差。

表 2 研究主题一致性分层模型（TCS）

层级	判定规则	分析维度
T1（团队方向层）	成果 D1 方向等于所属团队核心研究方向	个人与团队一致性
T2（领域交叉层）	成果 D1 方向不等于团队核心方向，但在心血管领域内（D01~D15）	个人与团队一致性
T3（方向偏离层）	成果 D1 方向为非心血管领域（D16）	个人与机构定位一致性

集中度指数（Concentration Ratio, CR_n ）是产业组织经济学中衡量市场集中程度的常用指标^[8-9]。它通过计算行业中前 n 家最大机构的市场份额总和，来反映集中程度，已广泛应用于经济研究与市场分析。公式如下：

$$CR_n = \sum_{i=1}^n S_i \tag{1}$$

其中， S_i 为第 i 个研究方向的占比， $S_1 \geq S_2 \geq \dots \geq S_n$ ，按占比从大到小排序。 CR_n 值越高，说明研究方向越聚焦，反之， CR_n 值越低，则说明研究方向趋于分散。常用形式有 CR_1 、 CR_3 、 CR_5 等。本研究将集中度指数应用于慢病防治领域科研人员的研究方向集中度的测度。其中 CR_1 （第一大方向占比）为最大研究方向成果数除以个人成果总数，直接反映科研人员在最大研究方向上的投入强度。 CR_3 （前三大方向占比）为前三大研究方向成果数之和除以个人成果总数，比 CR_1 （第一大方向占比）更全面地反映了研究产出的集中程度。根据 CR_3 将研究人员划分为三种类型，高度集中型（ $CR_3 \geq 0.8$ ），研究产出高度集中于少数方向；相对集中型（ $0.6 \leq CR_3 < 0.8$ ），有一定主攻方向但保持适度的方向多样性；分散型（ $CR_3 < 0.6$ ），研究产出广泛分布于多个方向。

个人与团队一致率（T1%）为 T1（团队方向层）成果数除以个人成果总数，衡量科研人员研究产出与其所属团队核心方向的相符程度。领域交叉率（T2%）为 T2（领域交叉层）成果数除以个人成果总数，反映科研人员在保持团队归属前提下参与跨方向学术合作的程度。T2%（领域交叉率）本身不含有优劣判断，适度的 T2%（领域交叉率）可能是学术创新的积极因素，但 T2%（领域交叉率）持续偏高也可能提示团队归属需要重新评估。团队层面指标包括团队平均一致率、团队内一致性水平分布和团队内集中类型的分布结构。

机构方向偏离率（T3%）为 T3（方向偏离层）成果数除以个人成果总数，衡量科研人员研究方向与机构整体学科定位和发展方向的相符程度。设定两级预警阈值：T3%（方向偏离率）不低于 10%为黄色预警，不低于 20%为红色预警。

16个研究方向的成果分布显示（表 3），D07（心血管流行病学）成果最多，共 304 项（16.0%），其次为 D06（主动脉与大血管疾病）204 项（10.7%）、D02（心力衰竭与心脏病）200 项（10.5%）、D03（脂代谢与动脉粥样硬化）183 项（9.6%）、D12（心血管免疫与炎症）147 项（7.7%）。这 5 个方向合计 1038 项，占全部成果的 54.6%，研究力量向少数优势方向集中的特征明显。其余 11 个方向合计 862 项，占 45.4%，其中 D15（心血管病其他领域）158 项（8.3%），为未归入 D01 至 D14 具体研究方向的心血管领域成果；D16（非心血管领域）66 项（3.5%），是被识别为偏离机构整体研究方向的成果。科研成果类型分布中，学术论文 1365 项（71.8%）为最主要的产出形式，其次是科研项目 254 项（13.4%）和专利 117 项（6.2%），指南、获奖、药械、软著、转化等其他 6 种类型合计占 8.6%。

表 3 各研究方向成果统计

		单位	%
			%
			%
			%
			%
			%
			%
			%
			%
			%
			%
			%
			%
			%
			%
			%
			%
			%
			%
			%
			%

表 4 成果分布与 TCS 分层构成

维度	类别	成果数（项）	占比（%）
TCS 分层	T1（团队方向层）	1018	53.6
	T2（领域交叉层）	816	42.9
	T3（方向偏离层）	66	3.5

TCS 分层的构成显示了三个值得注意的特征（表 4）。第一，T1（团队方向层）占 53.6%，超过半数的成果直接归属于研究人员所在团队的核心方向，说明团队作为学术生产

10.12201/bmr.202610.00022V1

的基本组织单元，对成员的研究方向具有实质性的引领作用。第二，T2（领域交叉层）占42.9%，说明领域内的跨方向合作是该机构学术研究活动的常态。第三，T3（方向偏离层）仅占3.5%，表明该机构整体的研究方向偏离风险处于较低水平。

表 4 成果分布与 TCS 分层构成

维度	类别	成果数 (项)	占比 (%)
TCS 分层	T1 (团队方向层)	1018	53.6
	T2 (领域交叉层)	816	42.9
	T3 (方向偏离层)	66	3.5

全机构 CR₁（第一大方向占比）均值为 58.8%，中位数为 53.9%，半数以上研究人员的最大研究方成果占比超过 50%。76 位研究人员中，高度集中型 45 人（59.2%），相对集中型 19 人（25.0%），分散型 12 人（15.8%）（表 5）。全机构 CR₁（第一大方向占比）值的分布跨度较大，表明该机构内部研究人员的聚焦程度存在较大的个体差异。高度集中型占比接近 2/3，表明该机构的研究模式以深耕特定领域为主导，但分散型也占据 15.8% 比例，说明有接近六分之一的研究人员的学术产出在方向上较为发散。

表 5 研究人员主题集中类型分布

集中类型	人数 (n)	占比 (%)	判定标准
高度集中型	45	59.2	CR ₁ ≥ 80%
相对集中型	19	25.0	60% ≤ CR ₁ < 80%
分散型	12	15.8	CR ₁ < 60%
合计	76	100.0	

有固定团队归属的 51 位研究人员，其 T1%（个人与团队一致率）均值为 71.9%，中位数为 75.0%。一致率水平分布如下：T1%（个人与团队一致率）不低于 80% 的 22 人（43.1%），50%至~80%之间的 19 人（37.3%），低于 50% 的 10 人（19.6%）。12 个团队的个人与团队一致性水平如表 6 所示。团队间 T1% 的差异较为显著，说明团队的组织模式是形成其成员研究聚焦特征的重要因素。

表 6 12 个研究团队聚焦度指标对比

团队编号	团队人数	团队 CR ₁ 均值 (%)	T1% (%)	高度集中型占比	分散型占比
1	3	50.8	37.9	1/3	1/3
2	15	37.2	—	3/15	5/15
3	11	52.3	33.1	7/11	1/11
4	7	67.8	66.7	5/7	0/7
5	6	46.3	29.9	5/6	1/6
6	10	46.8	—	5/10	2/10
7	6	79.6	79.6	5/6	0/6
8	3	74.9	41.6	2/3	1/3
9	4	78.3	78.3	3/4	0/4
10	5	69.4	49.4	3/5	0/5
11	3	96.1	96.1	3/3	0/3
12	3	79.2	66.7	2/3	0/3

团队的 T1%（个人与团队一致率）均值为 54.3%。团队 11（96.1%）和团队 9（78.3%）的 T1%（个人与团队一致率）较高，这些团队组织规模较小（3~4 人），具有明确的单一研究方向，成员围绕该方向形成了高度集中的学术产出格局。团队 11 的三名成员全部为高度集中型且无分散型，团队 9 的四名成员中三名高度集中、无分散型，两支团

10.12201/bmr.202610.00022V1

队均实现了方向上的高度统一。团队 1（37.9%）、团队 5（29.9%）和团队 3（33.1%）的 T1%（个人与团队一致率）偏低，这些团队的组织特征是多方向覆盖，成员的学术产出广泛分布于多个心血管亚领域，团队内部的研究方向呈现出分散特征。

全机构 T2（领域交叉率）成果 816 项，占 42.9%。如果聚焦于有固定团队归属的 51 人，其 T2%（领域交叉率）均值为 24.7%，中位数为 20.0%，说明这部分研究人员相当部分的产出属于领域内的跨方向合作。这一比例在各类团队之间波动较大，从单一方向团队到多方向综合团队差异明显。

全机构 T3%（方向偏离率）均值为 3.1%，多数研究人员未出现任何方向偏离，总体偏离程度较低。6 位研究人员的 T3%（方向偏离率）超过 10%的黄色预警线，其中 3 人超过 20%的红色预警线。对红色预警人员的成果方向进行审查后发现，偏离主要涉及肿瘤学基础研究和药物化学领域，与机构的心血管疾病研究定位存在一定偏差。这 3 人的 T3%（方向偏离率）分别达到 29.2%、42.9%和 50.0%，需要重点评估其研究方向是否需要调整或是否应调整至相应学科方向的机构。

研究

结合成果数量分布与专设团队情况分析，4 个方向被识别为相对薄弱的研究方向。例如 D10（心律失常）有 29 项成果，占全机构总成果的 1.5%。该方向无独立专设研究团队，缺乏主导性的学科带头人，研究成果分散于多个团队之中，难以形成持续的学术积累。D13（睡眠呼吸暂停与肺血管）有 44 项成果，占全机构总成果的 2.3%，分散在多个团队。D11（遗传性心血管病）有 34 项成果，占全机构总成果的 1.8%，有专设团队但规模较小。随着基因组学技术发展，该方向存在较大的人才需求空间，团队在湿实验、干实验和临床遗传咨询等领域的覆盖可能存在缺口。

与上述薄弱方向形成对比的是，D07（心血管流行病学，304 项）、D06（主动脉与大血管疾病，204 项）和 D02（心力衰竭与心肌病，200 项）三大方向的成果总量占全机构的 37.2%，且均拥有独立的专设团队，研究人员的学科背景和学术产出在这些方向上形成了稳固的积累。

研究方向聚焦不是一个可以在短期内任意调整的策略选择，而是在科研职业生涯早期形成后将持续产生影响的学术行为模式。

④对于高度集中型研究人员，关键是保障其在核心方向上的持续和深度投入。例如本研究中 59.2%的科研人员为高度集中型，反映了该机构科研人员普遍是在某一研究方向上开展深耕的研究模式。频繁的研究方向调整和过短周期的科研绩效考核，都可能会打断科研人员专业知识积累进程，影响其长期学术竞争力。尤其对于资历较浅的研究人员，在尚未建立足够专业深度的情况下过早推动多方向发展，也不利于其学术根基的形成。从机构层面看，识别并持续支持这类人员，也是维护和强化科研机构核心方向竞争力的需要。

10.12201/bmr.202610.00022V1

如果高度集中型人员因短期考核压力而转向热点方向或追逐多类型产出，机构将面临核心方向人才流失的风险。

(2)对领域内多方向的研究人员，关键是认可其跨方向研究的学术合理性和成果价值。如在本研究中，T2（领域交叉层）成果占全机构的42.9%，说明跨方向合作在该心血管研究机构中是较为普遍的学术行为，并且心血管疾病研究本身具有跨亚领域的特征。高血压与心力衰竭共享血流动力学机制，动脉粥样硬化与血管炎症共享免疫通路，遗传学发现需要流行病学数据验证其人群意义。这些交叉是知识生产的内在要求，不应简单等同于研究兴趣的分散。Chakraborty等发现高被引研究者倾向于采取散聚策略，在整段职业生涯中涉猎多个领域，但在特定时间段内高度聚焦于1至2个领域^[11]，为T2（领域交叉层）的学术合理性提供了实证参照。Madsen和Nielsen对10475名英国基金获得者的研究发现，主题性基金资助会引导研究者进入新方向，但资助结束后研究者倾向于回归原有方向，研究方向调整受外部资源牵引与内在学术惯性的共同作用^[12]。从管理角度看，T2%（领域交叉率）本身不构成判断优劣的依据。对于T2%（领域交叉率）持续偏高（如超过60%）的个别人员，可以重新评估其团队归属是否合适，分析当科研人员的大部分成果产出都不在所属团队的核心方向上时，现有团队能否为其提供充分的学术支持与协作环境。

(3)对跨领域的研究人员，关键是识别其研究方向与所在机构核心方向偏离的性质并评估其研究方向是否需要调整。如在本研究中，全机构研究成果的T3（方向偏离层）仅3.5%，方向偏离是少数现象。对少数出现方向偏离的人员，应首先识别偏离的性质，是主动的学术方向转向还是学科交叉的合理延伸；对核实后确实属于方向偏离的人员，可提示管理者重点评估其研究方向是否需要调整，或是否需要调整至相应学科方向的团队或机构。

通过聚焦度测度，结合成果数量分布与科研团队情况进行综合判断，可为精准识别机构薄弱的研究方向提供客观可量化的依据。

对缺乏独立团队的前沿方向，关键是进行前瞻性布局并引进交叉背景青年人才。以D10（心律失常）为例，该方向当前成果量相对少、无独立专设团队且缺乏学科带头人，但人工智能心电信号智能判读、可穿戴设备远程监测和房颤早期筛查等技术的应用正处于快速增长期，在这一方向建立研究力量，关系到该机构在心血管学科布局中的竞争力。建议可以优先引进具有人工智能与心脏电生理交叉背景的人才，待有一定学术积累后再设立独立团队。

(2)对有一定积累但缺乏独立团队的方向，关键是遴选骨干并组建团队。以D10（心律失常）和D13（睡眠呼吸暂停与肺血管）为例，两个方向的成果量分别为29项和44项，科研活动主要由多个团队的成员以分散的个体方式开展。当一个研究方向达到一定规模而缺乏独立团队时，研究活动的系统性和持续性会受到影响，也难以形成稳定的人才培养和课题申报体系。建议在这两个方向上遴选具备课题负责人潜力的中青年骨干，给予团队组建支持和初期资源倾斜，使其从分散的个体探索转向组织化的团队研究。

(3)对已有团队但规模不足的方向，关键是扩充团队规模。以D11（遗传性心血管病）为例，该方向在基因组学技术快速迭代的背景下，规模不足使其对多组学、基因编辑

10.12201/bmr.202610.00022V1

等新技术分支的覆盖能力受限，建议通过博士后专项岗位、联合培养机制弥补短期人才缺口，逐步健全该方向的技术覆盖面和人才梯队。

本研究从研究主题集中度、个人与团队一致性、个人与机构定位一致性三个层级，构建了医学科研人员研究聚焦度的测度框架。这一框架将聚焦度从科学计量学中的一个描述性指标，转化为可嵌入医学科研管理流程的操作性工具，为实施精细化人才评价提供了方法支撑。

(1) CR_n 集中度指数通过量化研究主题集中度，为识别科研人员的研究特征提供了直观依据。该指标自 Siegfried 在产业组织领域确立方法论地位以来^[13]，已在产业集聚测度^[14]和专利集中度分析^[15]等领域得到应用，但此前尚未引入医学科研人才评价。在聚焦度测度中选择 CR_n ，主要基于它的准确性和可理解性。Kvålseth 基于优超理论证明， CR_n 与 HHI 存在确定的边界对应^[16]，可由 CR_n 近似推知 HHI 的取值，说明 CR_n 所反映的集中度信息与 HHI 基本一致。同时， CR_n 以百分比形式直接呈现，相比熵值、HHI 等技术指标更直观易懂，更适合科研管理者和决策者在管理中使用。

(2) TCS 分层模型将聚焦度评价从个人层面拓展至团队和机构层面，为判断个人研究成果与团队方向、机构定位的匹配程度提供了量化依据。已有研究多聚焦于个人层面的专业化或跨学科测度，对个人与团队、机构之间的匹配关系关注不足。TCS 模型通过 T1（团队方向层）、T2（领域交叉层）、T3（方向偏离层）的三层标注，将一致性分析分解为两个维度：T1 和 T2 共同刻画个人与团队的关系，T3 单独衡量个人与机构定位的相符程度。模型的一个关键设计选择是引入 T2（领域交叉层）作为独立的中间类别，在心血管疾病这类具有跨亚领域特征的研究领域，可将跨方向合作作为独立类别处理，避免将合理的学科交叉误判为研究兴趣分散。

(3) 聚焦度指标在科研管理中应作为辅助工具加以运用，不宜作为独立的硬性考核指标。指标设计需契合人才发展阶段特征^[17]，且更适合作为课题负责人和团队评估中的参考信息，为管理者判断研究方向规划、团队归属调整等提供线索。

本研究存在以下局限。首先，本研究以心血管疾病领域为方法学设计与验证的起点，后续需要在肿瘤、代谢病、慢性呼吸系统疾病等其他慢病领域测试方法的适用性。本研究的分类体系是针对心血管疾病特定研究机构建立，尚未在其他慢病疾病及其亚专科层面开展分类探索。其次，理论上科研人员在职业生涯不同阶段的聚焦程度可能呈现不同模式，比如早期广泛探索，后期选定主攻深入方向^[18]，在某一领域深耕建立学术声望后，会扩展至更广泛领域。本研究对每位研究人员多年的成果合并计算聚焦度，但尚未利用成果的时间维度识别其科研生涯不同阶段的动态演变特征。最后，各研究方向词典的构建依赖领域专家知识，推广至其他学科领域的医学科研机构时需重新构建适配的词典。不同学科方向之间的语义距离差异，可能影响 TCS 分层精度。

后续研究可在以下方向拓展。建立包括心血管、肿瘤、代谢病等不同慢病领域的人员和成果数据集，检验聚焦度测度方法在不同慢病学科语境下的适用性与稳定性。引入年度

10.12201/bmr.202610.00022V1

纵向数据，追踪 CR_n （研究聚焦度）和 $T1$ （团队方向层）、 $T2$ （领域交叉层）、 $T3$ （方向偏离层）占比的时序变化轨迹，分析研究主题在各阶段转换的幅度与频次^[19]，识别典型的聚焦度演化模式。使用 Rao-Stirling 指数等综合多样性指标^[20]弥补 CR_n 仅反映占比分布而忽略方向间语义距离的缺陷，并建立覆盖全学术生涯链条的人才画像体系。

作者贡献：钟华负责研究设计、数据分析和论文撰写；范少萍、单连慧负责研究资料整理和论文审校；安新颖负责论文整体思路与内容审阅。

参考文献

10.12201/bmr.202610.00022V1

- 1[] 国家心血管病中心,中国心血管健康与疾病报告编写组.中国心血管健康与疾病报告 2023 概要[J].中国循环杂志,2024,39(7):625-660.
- 2[] De Rassenfosse G, Higham K, Penner O. Scientific rewards for biomedical specialization are large and persistent[J]. BMC Biology, 2022, 20(1): 211.
- 3[] 宋培彦,龙晨翔,李怡然,等.基于冰山模型的科研人员学术专长识别方法研究[J].数据分析与知识发现,2023,7(6):50-60.
- 4[] Hladík R, Renisio Y. Mapping knowledge: topic analysis of science locates researchers in disciplinary landscape[J]. Poetics, 2025, 108: 101950.
- 5[] Kwiek M, Szymula Ł. Quantifying lifetime productivity changes: a longitudinal study of 320,000 late-career scientists[J]. Quantitative Science Studies, 2025, 6: 1002-1038. DOI: 10.1162/qss.a.16
- 6[] Banerjee S, Ghosh A, Basu B. Exploring citation diversity in scholarly literature: an entropy-based approach[J]. Scientometrics, 2025, 130(5): 2673-2704. DOI: 10.1007/s11192-025-05313-2.
- 7[] Nandy A, Lathabai HH, Singh VK. x_d-index and its variants: a set of overall scholarly expertise diversity indices for the research portfolio management of institutions[J]. Scientometrics, 2024, 129(10): 5937-5962. DOI: 10.1007/s11192-024-05131-y
- 8[] 科技部,教育部,人力资源社会保障部,中科院,工程院.关于开展清理“唯论文、唯职称、唯学历、唯奖项”专项行动的通知[Z].国科发政〔2018〕210号,2018-10-15.
- 9[] Minaee S, Kalchbrenner N, Cambria E, et al. Deep learning-based text classification: a comprehensive review[J]. ACM Computing Surveys, 2021, 54(3): 1-40.
- 10[] Fu W, Akbar S. Expert profile identification from community detection on author-publication-keyword graph with keyword extraction[J]. IEEE Access, 2024, 12: 27918-27930. DOI: 10.1109/ACCESS.2024.3368003.
- 11[] Chakraborty T, Tammana V, Ganguly N, et al. Understanding and modeling diverse scientific careers of researchers[J]. Journal of Informetrics, 2015, 9(1): 69-78.
- 12[] Madsen EB, Nielsen MW. Do thematic funding instruments lead researchers in new directions? Strategic funding priorities and topic switching among British grant recipients[J]. Research Evaluation, 2024, 33: rvae015. DOI: 10.1093/reseval/rvae015.
- 13[] Siegfried JJ. In defense of the average concentration ratio[J]. Journal of Political Economy, 1975, 83(6): 1269-1276.
- 14[] 王子龙,谭清美,许箫迪.高技术产业集聚水平测度方法及实证研究[J].科学学研究,2006,24(5): 706-714.
- 15[] [1]胡海容,石冰琪,周孟蓉.基于 HHI 指数和 E 指数的我国区块链技术专利集中度测算研究[J].世界科技研究与发展,2021,43(5):523-534.
- 16[] Kvålseth TO. Relationship between concentration ratio and Herfindahl-Hirschman index: A re-examination based on majorization theory[J]. Heliyon, 2018, 4(10): e00846. DOI: 10.1016/j.heliyon.2018.e00846.
- 17[] 张洋,黄子萱,朱嘉麒.青年科技人才画像及鉴识方法研究[J].情报学报,2024,43(11):1283-1296.
- 18[] Gautheron L. Balancing specialization and adaptation in a transforming scientific landscape[J]. EPJ Data Science, 2025, 14(1): 2. DOI: 10.1140/epjds/s13688-024-00516-8
- 19[] Yin D. Quantitative analysis of scholars' topic switching behavior in computer science: a two-dimensional metric approach[J]. IEEE Access, 2024, 12: 104263-104271. DOI: 10.1109/ACCESS.2024.3434546.
- 20[] 许海云,尹春晓,郭婷,等.学科交叉研究综述[J].图书情报工作,2015,59(5):119-127.