

# 健康医疗 HEALTH AI INDEX REPORT 人工智能指数报告 20



中国医院协会健康医疗大数据应用管理专业委员会  
北京大学健康医疗大数据国家研究院  
北京大学信息技术高等研究院（浙江）  
北京大学人工智能研究院智慧公众健康研究中心  
爱思唯尔

二零二零年十月





## 专家指导委员会

### 主席：

詹启敏院士

董尔丹院士

### 成员：

邓志鸿 北京大学

段会龙 浙江大学

付君 哈尔滨医科大学

蒋云 北京大学信息技术高等研究院（浙江）

金海 华中科技大学

孔桂兰 北京大学

匡铭 中山大学

李姣 北京协和医学院

马婷 鹏城实验室

施秉银 西安交通大学

汤步洲 哈尔滨工业大学（深圳）

王海波 中山大学

王耀刚 天津医科大学

王志锋 北京大学

詹思延 北京大学

张路霞 北京大学

赵歌喃 爱思唯尔

赵明辉 北京大学

---

## 执笔委员会

杜建 北京大学健康医疗大数据国家研究院

吴静依 北京大学信息技术高等研究院（浙江）

赵璐 爱思唯尔

Baas, Jeroen 爱思唯尔（数据科学家）

Schalkwijk, Jan-Maarten Van 爱思唯尔（数据科学家）

白永梅 北京大学健康医疗大数据国家研究院

---

## 协调委员会

王迈、崔娜 北京大学健康医疗大数据国家研究院

李鹏飞 北京大学信息技术高等研究院（浙江）

张丹丹、王巍 爱思唯尔



## 目 录

<b>引言</b>	<b>01</b>
<b>第一章 界定与分类</b>	<b>03</b>
1. 数据集界定 .....	03
2. 研究领域分类 .....	04
<b>第二章 科学研究概览</b>	<b>06</b>
1. 数据与指标 .....	06
1.1. 数据来源 .....	06
1.2. 分析指标 .....	06
2. 分析结果 .....	08
2.1. 主要国家 .....	08
2.2. 研究机构 .....	08
2.3. 科研合作 .....	10
2.4. 研究主题分布 .....	11
2.5. 高科学影响力论文 .....	14
2.6. 全球 Health AI 细分领域分析 .....	14
2.7. 中国 Health AI 细分领域分析 .....	15
<b>第三章 科学技术交叉</b>	<b>20</b>
1. 数据与指标 .....	20
1.1. 数据来源 .....	20
1.2. 学术界与产业界的知识流动指标 .....	20
2. 分析结果 .....	20
2.1. 学术界与产业界的知识流动总体状况 .....	20
2.2. 对 Health AI 技术产生影响的基础性研究 .....	21
2.3. Health AI 细分领域科学 - 技术交叉 .....	23
<b>第四章 科学社会交互</b>	<b>24</b>
1. 数据与指标 .....	24
1.1. 数据来源 .....	24
1.2. 科学社会交互指标 .....	24
2. 分析结果 .....	25
2.1. 多媒介提及指数较高的话题 .....	25
2.2. 社交媒体传播指数较高的话题 .....	25

# 目录

<b>第五章 人类 - 机器协同 (AI 临床试验)</b>	<b>29</b>
1. 引言 .....	29
2. 数据与指标 .....	31
2.1. 数据来源 .....	31
2.2. 分析指标 .....	31
3. 分析结果 .....	32
3.1. 临床试验数量 .....	32
3.2. 发起机构分布 .....	32
3.3. 临床试验分期 .....	34
3.4. 研究类型分布 .....	34
3.5. 干预措施类型 .....	34
3.6. 目标人群 (疾病谱) .....	35
3.7. 样本量分布 .....	36
3.8. 招募状态分布 .....	37
3.9. 临床试验结果报道 .....	37
<b>第六章 主要结论</b>	<b>39</b>
1. 前沿科学技术与医学的深度融合是健康医疗人工智能发展的基础，未来将在公共卫生和临床诊疗中发挥更大作用。 .....	39
2. 中国已成为健康医疗人工智能科学研究与临床试验的最主要贡献者之一，但在学术影响力和技术转化方面仍有待提升。 .....	39
3. 健康医疗人工智能技术谱的核心是机器学习 (含深度学习) 和医疗机器人，疾病谱以慢病和神经系统疾病为主，传染病、罕见病等与医疗 AI 的深度结合仍有空间。 .....	40
4. 与科学共同体的热点研究主题相比，健康医疗人工智能研发与应用的伦理学问题成为社交媒体关注的焦点。 .....	40
5. 健康医疗人工智能全球临床研究仍处于早期阶段，中国以大学 / 医院为发起主体，侧重疾病智能诊断，企业参与仍有待加强。 .....	41
6. 建议将循证范式引入健康医疗人工智能安全性和有效性评价，以促进其落地应用。 ...	41

## 引言

人口老龄化、社会环境因素变化、慢性疾病预防负担加重、新发突发传染病等对我国人民健康带来了巨大挑战。健康医疗人工智能（Health Artificial Intelligence, Health AI）的核心是利用人工智能等前沿科学技术赋能医疗健康，构建最优化的大健康生态体系，提供优质、高效、经济的新型医疗服务，为解决我国医疗供需矛盾、推动医学发展提供有效的手段。

2019年4月26日，《英国医学杂志》（The BMJ）在线发表北京大学健康医疗大数据国家研究院詹启敏院士等学者题为“Can AI fulfil its medical promise?（健康医疗——人工智能的应许之地？）”的观点文章。<sup>1</sup>该文肯定了人工智能（AI）技术在健康医疗领域颇具潜力的应用场景，例如医学影像和病理学诊断，以及作为辅助手段用于常见疾病一般状况的医疗决策。然而，由于AI算法的内在运行逻辑难以探究、被研究者称为“黑匣子”，而这“会给使用AI系统的医生带来迟疑和困惑”。作者提出，AI系统的有效性和安全性必须得到科学的评估，并建议运用流行病学及医学研究的思路来验证基于AI的预测模型。此外，人文关怀是医学中至关重要、并且无法被任何技术系统替代的关键内容。最后，作者将健康医疗领域“划归”为AI的“应许之地”，但同时也表示，为充分发挥其潜质，“医生、科研人员和AI科学家应当紧密合作；基于可靠的方法、遵循伦理的准则，力争在医疗实践中应用、评估和改进AI技术”。

自2017年以来，斯坦福大学联合麻省理工学院、哈佛大学等机构，每年发布AI指数报告（AI Index Report），从学术、产业、政策等多个角度介绍全球AI的最新进展，至今已连续3年发布报告<sup>2</sup>。2018年，信息分析公司爱思唯尔发布《人工智能：知识的创造、转移与应用》报告，对AI领域进行了全面定义，勾勒出该领

<sup>1</sup> <https://blogs.bmj.com/bmj/2019/04/26/can-ai-fulfill-its-medical-promise/>

<sup>2</sup> <https://hai.stanford.edu/research/ai-index-2019>

域在全球范围内的研究趋势，同时聚焦 AI 在中国、欧洲和美国的发展态势<sup>3</sup>。2019 年 12 月，美国国家医学院发表《医疗人工智能：希望、炒作、浮夸承诺、危险》综述报告，汇总现有 AI 知识和技术应用，包括成功案例和失败教训，为医疗健康领域的人工智能研发、应用和维护，提供了指导建议和实践指南<sup>4</sup>。2020 年 4 月，经济合作与发展组织（Organization for Economic Co-operation and Development, OECD）发布《值得信赖的医疗人工智能》报告，讨论了人工智能在医疗保健领域的前景和风险，以及政策制定者在不确定的环境下需要解决的关键政策问题<sup>5</sup>。

人工智能在健康医疗领域的研究与应用主要集中于 4 个方面<sup>6</sup>：一是疾病诊断，例如电子病历（EMR）/ 电子健康档案（EHR）和专家知识整合，基于 AI 的常见病诊断和评估；二是疾病治疗，例如通过将 EHR 和临床指南整合，基于 AI 治疗常见病、基于药物基因组学指导临床用药等；三是人群健康管理，例如建立以患者为中心的信息系统，开展健康生活方式监测与干预，疾病早期

检测与发现，健康知识教育等；四是管理与监管，例如医疗服务质量评估、药物不良反应监测等。与金融、电信等行业相比，AI 在健康医疗领域发展相对滞后且未充分落地，研究和应用还有很大空间<sup>7</sup>。但目前尚未见从客观数据角度系统分析健康医疗人工智能研究与开发状况的报告。

北京大学健康医疗大数据国家研究院致力于促进和引领人工智能在健康医疗领域的研究和应用，并研发一个科研发展复合指数以供该领域国内科研机构参考。指数（Index），即反映某个领域发展状况的多维、综合指标，为该领域的相关决策提供依据。本报告将从科学研究概览、科学技术交叉、科学社会交互、人类机器协同四个方面，以已发表的科学出版物和已注册的临床试验为基础数据，回顾分析健康医疗人工智能领域最近 5 年（2015-2019）的全球科学研究和临床试验的规模、结构和发展趋势，并分析中国的表现，以期为国内健康医疗人工智能领域的战略规划、研发布局和临床应用管理提供参考。

---

<sup>3</sup> <https://www.elsevier.com/zh-cn/research-intelligence/artificial-intelligence>

<sup>4</sup> <https://nam.edu/artificial-intelligence-special-publication/>

<sup>5</sup> <https://www.oecd.org/health/trustworthy-artificial-intelligence-in-health.pdf>

<sup>6</sup> He, J., Baxter, S. L., Xu, J., Xu, J., Zhou, X., & Zhang, K. (2019). The practical implementation of artificial intelligence technologies in medicine. *Nature Medicine*, 25(1), 30–36.

<sup>7</sup> <https://www.wish.org.qa/wp-content/uploads/2018/11/IMPJ6078-WISH-2018-Data-Science-181015.pdf>



# 第一章 界定与分类

目前，学界对健康医疗人工智能的界定尚未建立共识。通过对科学出版物的分析可帮助我们系统且清晰地描述该领域及其子领域所涵盖的内容和知识结构，本报告尝试提出该领域科学出版物数据集的界定方案。

## 1. 数据集界定

### 定义的数据集由两部分出版物构成：

第一部分由北京大学健康医疗大数据国家研究院提供。采用医学领域权威的知识组织体系——医学主题词表 (Medical Subject Headings, MeSH)，通过 MEDLINE 数据库对健康医疗人工智能科学出版物进行界定。为减少数据噪声，本报告采用主要主题词 (MeSH Major Topic, 该文章最核心的研究内容) 检索出版物。一般情况下，每篇 MEDLINE 论文会标注 10 条左右的主题词 (MeSH)，但会从中再遴选出 3~5 个最能代表这篇论文核心内容的主题词，标注为主要主题词。如果一篇论文被标注的主要主题词中，同时含有人工智能和疾病健康两个方面，则视为健康医疗人工智能出版物。其中，“人工智能”采用 "artificial intelligence" 及其所有下位术语来表示；“疾病或健康”采用 "diseases category"[C] 或 "mental disorders"[F03] 或 "health"[N01.400] 或 "public health"[N06.850] 及其所有下位术语来表示。

按如下纳入排除标准：①仅纳入针对人的研究，检索表达式为 “human[MeSH]”；②排除“撤稿”类出版物。检索表达式为 “Retracted Publication[Publication Type]” 或 “Retraction of Publication[Publication Type]”；③纳入发表时间窗为 2015-2019 共计 5 个完整年份，检索表达式为 “2015”[Date - Publication] : “2019”[Date - Publication]。

第二部分由爱思唯尔提供。爱思唯尔 2018 年《人工智能：知识的创造、转移与应用》报告通过机器学习的手段对人工智能领域的科学出版物进行了界定。该部分文集通过对爱思唯尔已构建 AI 数据集进行学科分类，主要提取 “Medicine”、“Health Professions”、“Dentistry”、“Nursing” 和 “Multidisciplinary” 这五个学科的论文。其中对于 “Multidisciplinary” 学科 (例如 Science、Nature 等综合性期刊) 将基于该学科内论

文的施引和被引分布进行学科重新归类，对施引和被引主要分布于 “Medicine”，“Health Professions”，“Dentistry” 和 “Nursing” 的论文进行提取。

两部分数据集的组合示意图 1-1，融合之后的数据集作为本报告界定的健康医疗人工智能科学出版物语料库，进而展开分析。

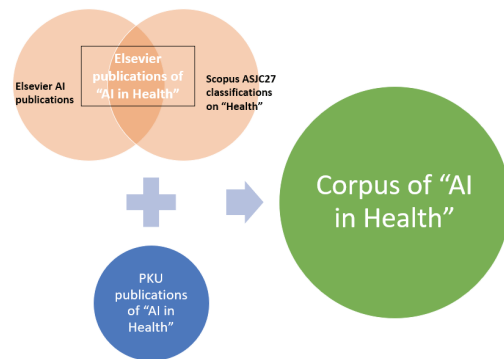


图 1-1 健康医疗人工智能科学出版物数据集界定方法

融合后数据集为 25,717 篇科学出版物，其中源自 Elsevier 已构建 AI 数据集有 18,390 篇，占 71%；北京大学健康医疗大数据国家研究院补充的剩余 29% 的数据集中，10% (2,496 篇) 同时出现在双方的数据集中；19% (4,831 篇) 为独立补充数据 (图 1-2)。

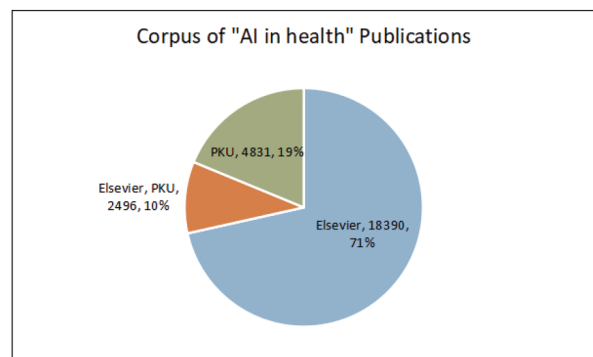


图 1-2 健康医疗人工智能科学出版物数据集的来源

## 2. 研究领域分类

由于本报告的数据集以科学出版物为主，且聚焦健康医疗领域，因此我们考虑仍采用医学主题词表这一术语体系对健康医疗人工智能研究领域进行分类。在总数据集中，14,820 篇被 MEDLINE 收录，含自动标注的 MeSH 术语；剩余 10,897 篇科学出版物并未被 MEDLINE 收录，未标注 MeSH 主题词；对于这部分科学出版物，采用文本挖掘工具——Medical Text Indexer (MTI) 将标题和摘要文本自动映射并标注 MeSH 主题词；利用 python 3.7 通过爬虫将文件列表中的摘要抓取并储存，随后分配一个识别编号，用于后续 MeSH 主题词映射的追踪与分析。对上述数据进行清洗，数据符合输入要求后输入 MTI 系统。在得到 MTI 结果后，利用 python 程序通过识别编号将主题词与文献之间建立映射。对于已有 PMID 的文献 (n=14,820)，我们利用 PubMed 获取数据进行分析。

### 本报告重点关注两个方面的分类：

- (1) 健康医疗人工智能技术领域的分类；
- (2) 健康医疗人工智能涉及的疾病分类，考虑采用世界卫生组织 (WHO) ICD-10 分类标准，并与疾病负担数据进行映射。

在 MeSH 树状结构表中，人工智能 (Artificial Intelligence) 位于信息科学 (Information Science) 大类中，具体层级结构为：

- Information Science [L]
- Information Science [L01]
- Computing Methodologies [L01.224]
- Algorithms [L01.224.050]
- Artificial Intelligence [L01.224.050.375]

“Artificial Intelligence”这一术语的 MeSH 树状结构编码为 L01.224.050.375，其一级下位术语有 8 个，包括：计算机启发式决策、专家系统、模糊逻辑、知识组织系统、机器学习、自然语言处理、计算机神经网络和机器人。这些一级下位术语的所有下位术语包括：

- Artificial Intelligence [L01.224.050.375]
- Computer Heuristics [L01.224.050.375.095]
- Expert Systems [L01.224.050.375.190]
- Fuzzy Logic [L01.224.050.375.250]
- Knowledge Bases [L01.224.050.375.480]
- Biological Ontologies [L01.224.050.375.480.500]
- Gene Ontology [L01.224.050.375.480.500.500]
- Machine Learning [L01.224.050.375.530]
- Deep Learning [L01.224.050.375.530.250]

- Supervised Machine Learning [L01.224.050.375.530.500]
- Support Vector Machine [L01.224.050.375.530.500.500]
- Unsupervised Machine Learning [L01.224.050.375.530.750]
- Natural Language Processing [L01.224.050.375.580]
- Neural Networks, Computer [L01.224.050.375.605]
- Deep Learning [L01.224.050.375.605.500]
- Robotics [L01.224.050.375.630]

由于上述 8 个一级下位术语之间存在交叉，经咨询医学信息学和医疗人工智能领域专家意见，按如下规则对其进行重组分类：①将 Computer Heuristics 和 Fuzzy Logic 合并，在算法层面，都是先验知识的形式化表示。Expert Systems 单独列出，是应用层面的信息系统。② Neural networks, computer 归并到 Machine Learning 下面。处理后，共计 6 个大类：

表 1- 1 健康医疗人工智能技术领域分类

类别	类别名称	MeSH 术语与编码
1	决策规则	-Computer Heuristics [L01.224.050.375.095] -Fuzzy Logic [L01.224.050.375.250]
2	专家系统	-Expert Systems [L01.224.050.375.190]
3	知识库	-Knowledge Bases [L01.224.050.375.480] --Biological Ontologies [L01.224.050.375.480.500] ---Gene Ontology [L01.224.050.375.480.500.500]
4	机器学习	-Machine Learning [L01.224.050.375.530] --Deep Learning [L01.224.050.375.530.250] --Supervised Machine Learning [L01.224.050.375.530.500] ---Support Vector Machine [L01.224.050.375.530.500.500] ---Unsupervised Machine Learning [L01.224.050.375.530.750] -Neural Networks, Computer [L01.224.050.375.605] --Deep Learning [L01.224.050.375.605.500]
5	自然语言处理	-Natural Language Processing [L01.224.050.375.580]
6	机器人	-Robotics [L01.224.050.375.630]

注意以上 6 个大类并非完全相互独立，也存在交叉，例如机器学习的下位术语深度学习，往往被用于自然语言处理；但基本能反映健康医疗人工智能涉及的主要技术领域。

## 第二章 科学研究概览

### 1. 数据与指标

#### 1.1. 数据来源

采用由北京大学健康医疗大数据国家研究院和爱思唯尔双方融合后的数据集，该数据集含有 25,717 篇与 Health AI 主题有关的科学出版物，基于爱思唯尔的 Scopus 和 Scival 平台进行统计分析。

#### 1.2. 分析指标

##### (1) 科研产出及影响力指标

**发文量 (Scholarly Output):** 发文量统计了被评估主体发表的包含期刊论文、会议文集、综述文章、发表丛书的所有文章的数量，代表了被评估主体在某一个固定时间段内的科研产出。

**归一化引文影响力 (Field-weighted Citation Impact, FWCI):** FWCI 是指被评估主体发表文章所收到的总被引次数相比于与其同类型发表文章 (相同发表年份、相同发表类型和相同学科领域) 所收到的平均被引次数的比值。FWCI 在一定程度上反映了被评估主体发表文章的学术影响力。相比于总被引次数，FWCI 能够更好的规避不同规模的发表量、不同学科被引特征、不同发表年份带来的被引数量差异。如果 FWCI 为 1 意味着被评估主体的文章被引次数正好等于整个 Scopus 数据库同类型文章的平均水平。本次分析的文章被引用次数统计截至 2020 年 9 月 18 日。

其中，归属国家 / 地区定义为：只要在一篇文章的所有作者中，其隶属研究机构从属于某一国家 / 地区，则该文章会被归为该国家 / 地区的一篇文章。一篇国际合作型文章会由于署有多个国家 / 地区的隶属机构而同时属于多个国家 / 地区。

##### (2) 机构科研产出及影响力指标

机构科研产出及影响力的评价指标包括：发文量、FWCI 和被引次数。被引次数 (Citation) 是指在某一个固定时间段内被评估主体所发表文章的所有被引用次数，在一定程度上反映了被评估主体发表文章的学术影响力。但同时也需考虑到，发表时间较近的文章相比于年份较久的文章，会由于积累时间较少而导致总被引次数较少。本次分析的文章的所有被引次数统计截至 2020 年 9 月 18 日。

##### (3) 科研合作指标

**学术科研合作文章分为三类：**国际合作 (International Collaboration)、国内合作 (National Collaboration) 和机构内合作 (Institutional Collaboration)，其中：

**国际合作文章：**是指文章的发表作者为多位作者，且至少有两位作者的署名机构来源于不同的国家 / 地区，表明该类文章源于国际合作的成果。

**国内合作文章：**是指文章的发表作者为多位作者，且作者中没有隶属于国外研究机构，但至少有一位作者隶属于国内不同的研究机构，表明该类文章源于国内合作的成果。

**机构内合作文章：**是指文章的发表作者为多位作者，且所有作者全部隶属于国内同一机构，表明该类文章源于机构内合作的成果。

**无合作文章：**是指文章发表作者为一人。该类别文章作为对照组进行展示。

##### (4) 产学合作指标

**产学合作 (Academic-Corporate Collaboration):** 该指标考察了学术机构和企业的合作程度。产学合作文章源于产学合作的成果，是指文章的发表作者为多位，其中至少有一位作者的隶属单位属于学术机构，且至少有一位作者的隶属单位属于产业界。

### (5) 研究主题分析

**研究主题 (Topic)**<sup>8</sup>: 研究主题是指一群具有共同研究兴趣的文章所研究内容的共同焦点。在 Scopus 数据库中, 所有文章通过直接被引算法归类于约 96,000 个研究主题中, 每篇文章只能属于一个研究主题。

**研究主题群 (Topic Cluster)**: 研究主题群是指将具有相似研究焦点的研究主题 (topic) 聚集在一起, 形成更广泛、更高层次的研究领域。在深入研究更为细分的研究主题之前, 这些研究主题群可用于更广泛地了解一个国家 / 地区、机构或研究人员正在进行的研究。96,000 个研究主题分别被匹配到 1,500 个研究主题群中, 同样地, 每篇文章只能属于一个研究主题群。研究主题群同样通过直接被引算法计算而得, 当多个研究主题间的引文链接强度达到某个阈值时, 就形成了一个研究主题群。

**研究主题 (群) 显著度得分 (Topic Prominence Score)**: 每个研究主题 (群) 的显著度得分是根据该研究主题 (群) 内所有文章的被引次数、在 Scopus 中的被浏览数和期刊平均引用分 (CiteScore) 等三个指标的线性计算得到。根据显著度得分从高到低排名可得到本研究主题 (群) 的全球显著度百分位数。已有相关研究表明<sup>9</sup>, 研究主题 (群) 显著度得分代表了该研究主题 (群) 被全球学者的关注程度、热门程度和发展势头, 并且研究主题 (群) 的显著度得分与其所获研究资助呈现正相关关系。

### (6) 高科学影响力论文分析

采用论文的被引次数评价其相应的科学影响力, 分析被引次数排名前 20 位的论文的特征。

### (7) 疾病分类方法

根据国际疾病分类 (International Classification of Diseases 10th Revision, ICD-10) 编码的分类结构对疾病的 MeSH 主题词进行了重新分类。采用荷兰莱顿大学学者构建的 ICD-MeSH 映射词表<sup>10</sup>, 在此基础上制作了包含该表所列疾病主题词的所有下位主题词的映射词表。使用该映射词表可直接基于 MeSH 主题词实现对疾病的分类和统计。

<sup>8</sup> [https://service.elsevier.com/app/answers/detail/a\\_id/28428/supporthub/scival/](https://service.elsevier.com/app/answers/detail/a_id/28428/supporthub/scival/)

<sup>9</sup> Klavans, R., & Boyack, K. W. (2017). Research portfolio analysis and topic prominence. *Journal of Informetrics*, 11(4), 1158–1174.

<sup>10</sup> Yegros-Yegros, A., Van de Klippe, W., Abad-Garcia, M. F., & Rafols, I. (2020). Exploring why global health needs are unmet by research efforts: the potential influences of geography, industry and publication incentives. *Health research policy and systems*, 18, 1-14.

## 2. 分析结果

### 2.1. 主要国家

2015-2019 年间，全球有关 Health AI 的文章总量为 25,717 篇，FWCI 均值为 2.0。全球及发文量前五位国家的科研产出及影响力的分布情况详见图 2-1。发文量排名前五位的国家依次是美国、中国、印度、英国和德国，FWCI 由高到低依次是美国、英国、德国、中国和印度。2015-2019 年，中国文章的平均 FWCI 为 2.0，与全球平均水平持平（FWCI=2.0），但低于美国（FWCI=2.8）。美国、英国和德国的 FWCI 均在同一水平线上，印度（FWCI=1.2）则显著低于全球平均水平。

发文量排名前五位的国家科研产出及影响力趋势变化见图 2-2。全球的文章数量从 2015 年的 2,573 篇增加到 2019 年的 10,018 篇，年复合增长率为 31.2%；中国的文章数量从 2015 年较低的 269 篇增长到 2019 年的 1,819 篇，年复合增长率达到 46.6%。前五位国家中，文章产出年复合增长率最高的国家是印度，其年复合增长率达到 54.8%。

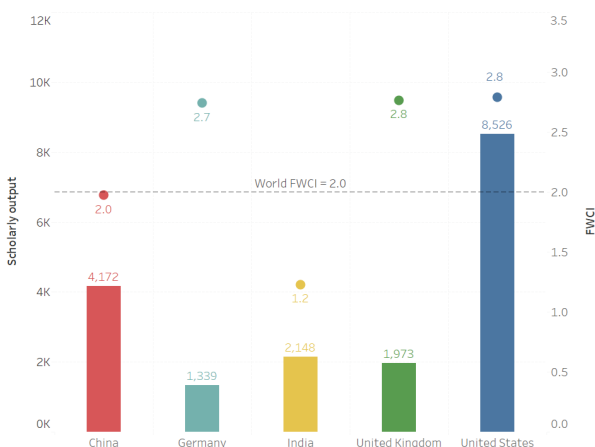


图 2-1 发文量排名前五位的国家的科研产出及影响力对比

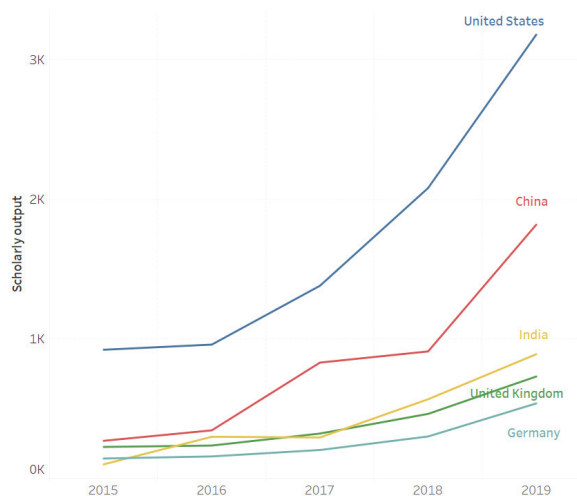


图 2-2 发文量排名前五位的国家发表科学出版物的年度分布

### 2.2. 研究机构

2015-2019 年间发文量排名前十位的中国学术机构的科研产出及影响力情况见图 2-3。其中，发文量排名最高的学术机构为上海交通大学（206 篇），其文章被引次数排名第三（2,360 次），同时 FWCI 较低。香港中文大学在前十位学术机构中发文量最低，但其文章的被引次数最高（3,235 次），且 FWCI 达到 5.6，明显高于其他学术机构。

2015-2019 年间 FWCI 排名前十位的中国学术机构的科研产出及影响力情况见图 2-4。FWCI 排名最高的学术机构为香港中文大学，其 FWCI 均值为 5.6，其文章的被引次数同样最高（3,235 次）。中山大学在 FWCI 排名前十位的学术机构中发文量最高（148 篇），且其被引次数排名第二（2,509 次）。

2015-2019 年间被引次数排名前十位的中国学术机构的科研产出及影响力情况见图 2-5。被引次数排名最高的学术机构为香港中文大学，其被引次数达到 3,235 次，其文章的 FWCI 同样最高（5.6）。上海交通大学在被引次数排名前十位的学术机构中发文量最高（206 篇），但其 FWCI 较低。

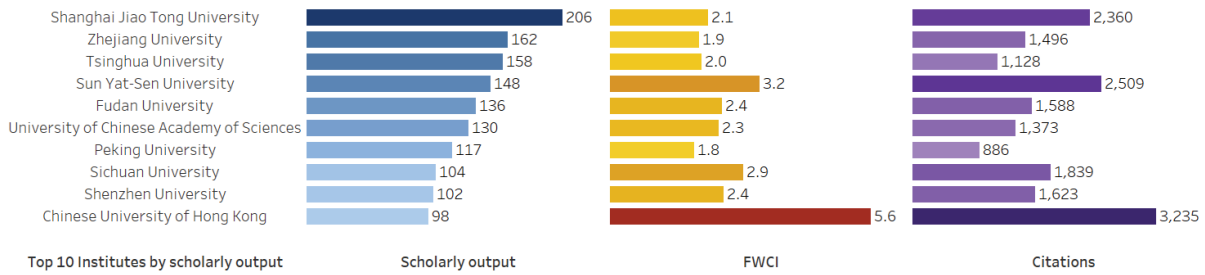


图 2-3 发文量排名前十位的中国学术机构的科研产出及影响力

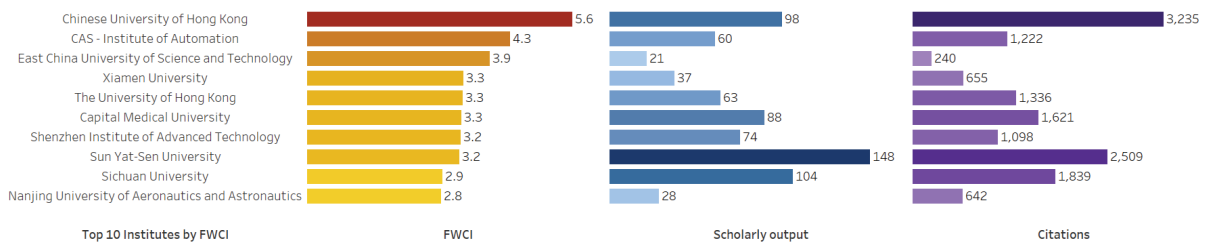


图 2-4 FWCI 排名前十位的中国学术机构的科研产出及影响力

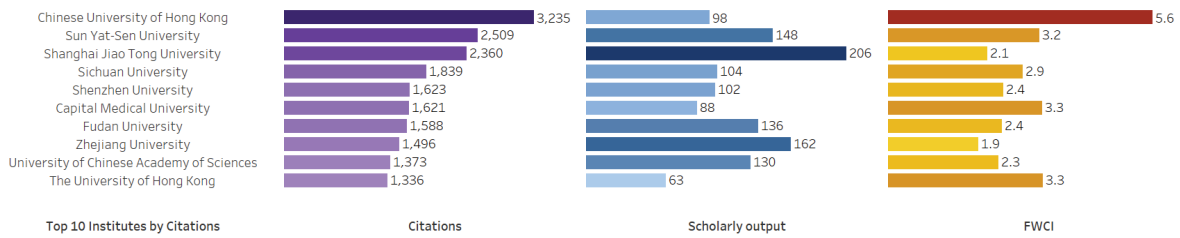


图 2-5 被引次数排名前十位的中国学术机构的科研产出及影响力

### 2.3. 科研合作

中国、美国和全球科研合作类文章的发文量占比及 FWCI 分布情况见图 2-6。中国和美国在各个类型的科研合作中发文量占比均较为相近，体现为国际合作占比最多，机构内合作最少；国际合作的 FWCI 最高，机构内合作的 FWCI 最低。中国的国际合作类文章的 FWCI (2.9) 高于全球平均 FWCI (2.8)，其余合作类型均低于相应的全球水平。

中国、美国和全球产学研合作类文章的发文量占比及 FWCI 对比见图 2-7，美国的产学研合作类文章的发文量占比与 FWCI 均高于中国。中国的产学研合作类文章虽然占比低于全球平均水平，但 FWCI (3.7) 高于全球平均水平 (3.4)。

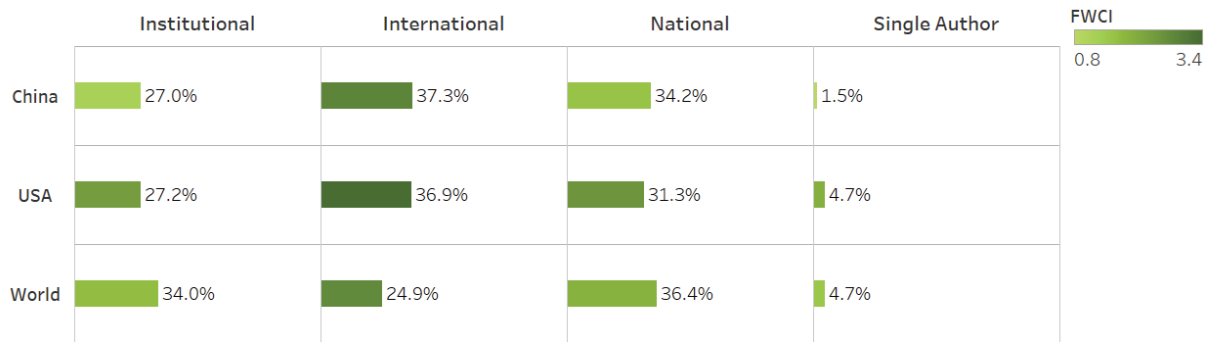


图 2-6 中国、美国和全球科研合作类文章的发文量占比及 FWCI

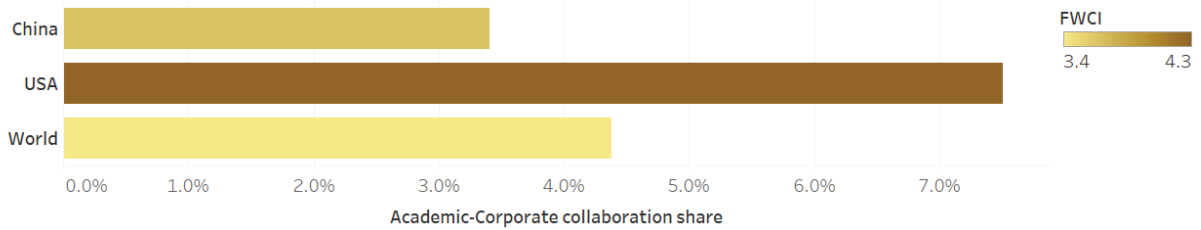


图 2-7 中国、美国和全球产学研合作类文章的发文量占比及 FWCI



## 2.4. 研究主题分布

2015-2019 年间健康医疗人工智能文章的研究主题群分布情况见图 2-8。该类研究共涉及 987 个研究主题群，涉及学科领域包括医学、计算机科学、物理学、生物化学、遗传学和分子生物学、环境科学、工程学及社会科学等。包含文章数最多的研究主题群为“算法、计算机视觉、模型”（Algorithms, Computer Vision, Models），含 1,965 篇文章。该研究主题群显著度百分位达到 99.8%，其下包含 270 个研究主题，其关键词包括“神经网络”（Neural Network）、“深度学习”（Deep Learning）、“计算机视觉”（Computer vision）、“面部识别”（Facial Recognition）和“目标检测”（Object Detection）等。由于在 Scopus 数据库中，每一篇文章都会根据直接被引算法被归类为一个研究主题群，即每一个研究主题群中都包含一定数量的文章，称为该研究主题群的全球文章数量。在本研究的数据集中，被归为某个特定研究主题的文章数占该研究主题所有文章数量的比例体现了健康医疗人工智能对于该研究主题群的内容贡献。其中占比最大的研究主题群，即包含健康医疗人工智能文章比例最高的研究主题群为“细胞学；图像分割；医学影像学”（Cytology; Image Segmentation; Medical Imaging），含 636 篇文章。该研究主题群显著度百分位为 67.9%，其下包含 71 个研究主题，其关键词包括“深度学习”（Deep Learning）、“神经网络”（Neural Network）、“流式细胞仪”（Flow Cytometry）、“组织病理学”（Histopathology）和“图像分割”（Image Segmentation）等。

中国学术机构参与的健康医疗人工智能文章的研究主题群分布情况如图 2-9 所示，该部分文集共涉及到 505 个研究主题群，涉及到的学科领域包括医学、计算机科学、生物化学、遗传学和分子生物学、环境科学、工程学及社会科学等。包含文章数最多的研究主题群同样是“算法、计算机视觉、模型”（Algorithms, Computer Vision, Models），含 477 篇文章。该研究主题群显著度百分位达到 99.8%，其下包含 270 个研究主题，关键词包括“神经网络”（Neural Network）、“深度学习”（Deep Learning）、“计算机视觉”（Computer Vision）、“面部识别”（Facial Recognition）和“目标探测”（Object Detection）等。该文集被收录在每个研究主题群的全球所有文章中占比最高的研究主题群

为“磁共振成像、图像分割、医学成像”（Magnetic Resonance Imaging; Image Segmentation; Medical Imaging），含 175 篇文章。该研究主题群显著度百分位为 74.6%，其下包含 99 个研究主题，其关键词包括“脑肿瘤”（Brain Neoplasm），“影像配准”（Image Registration）和“图像分割”（Image Segmentation）等。

图 2-10 至 2-13 分别为按照发文量、发文量占比、FWCI 和显著度百分位进行排名的前十个研究主题群。

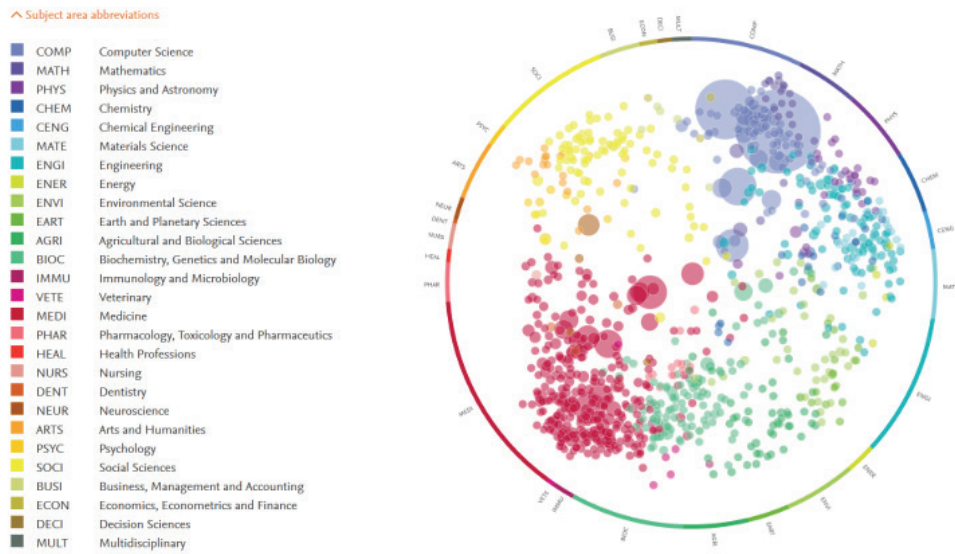


图 2-8 全球健康医疗人工智能文章的研究主题群圈图

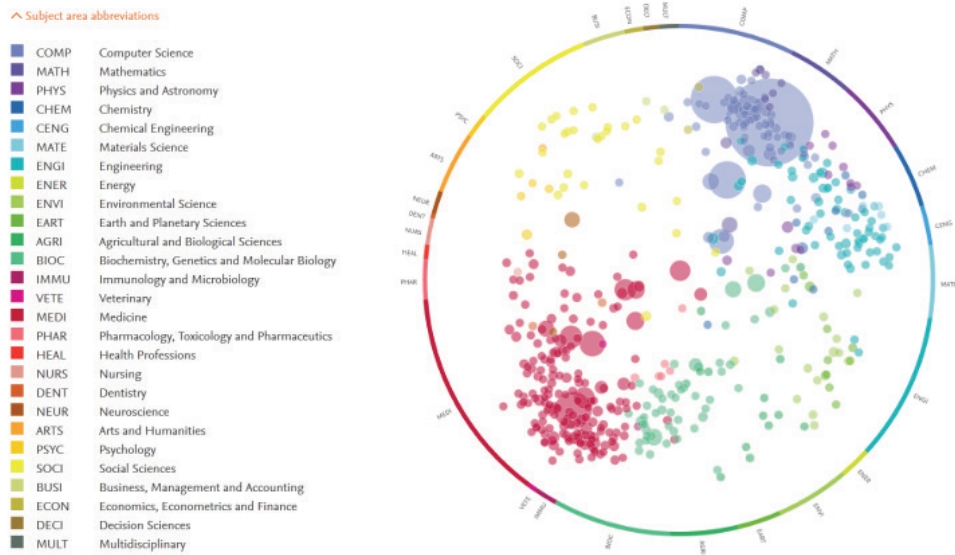


图 2-9 中国健康医疗人工智能研究文章的研究主题群圈图

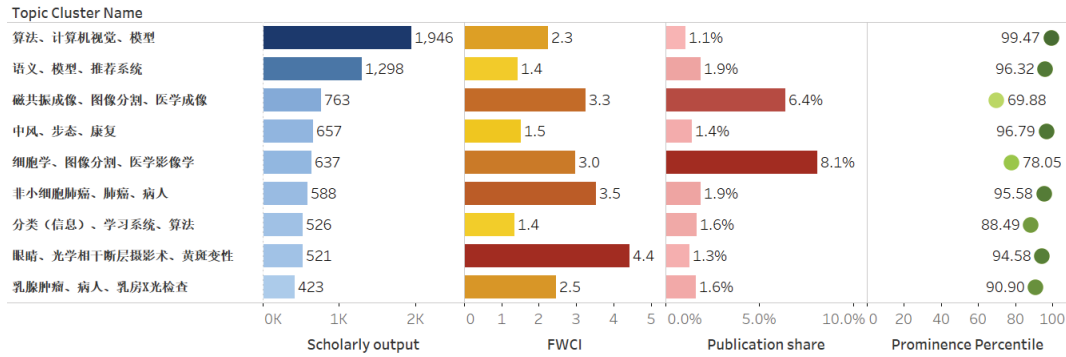


图 2-10 发文量排名前十位的研究主题群

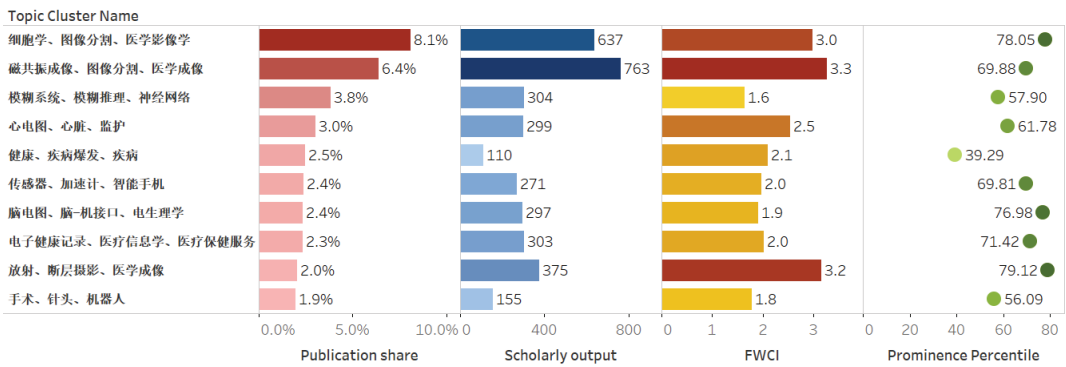


图 2-11 发文量占比排名前十位的研究主题群

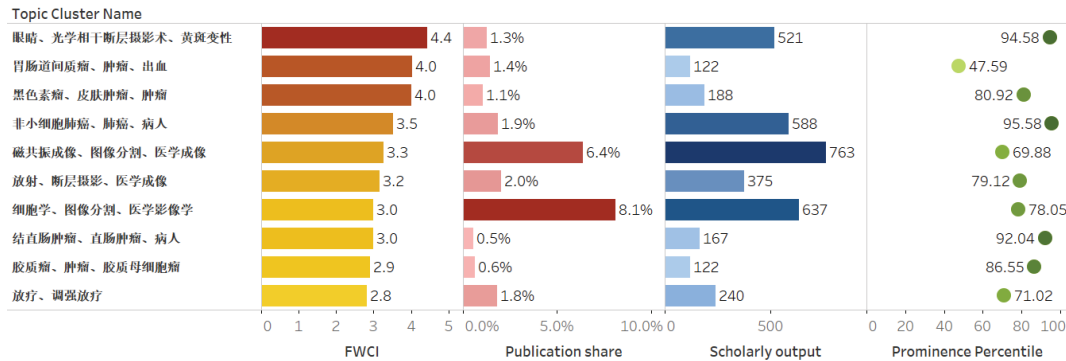


图 2-12 FWCI 排名前十位的研究主题群

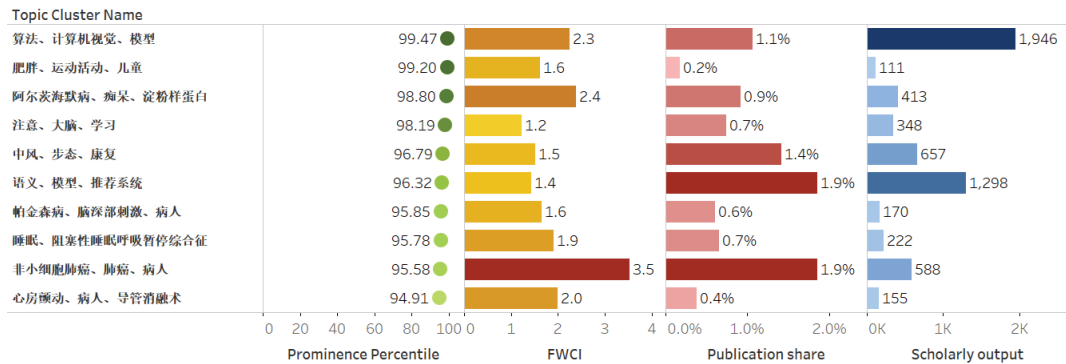


图 2-13 显著度百分位排名前十位的研究主题群

## 2.5. 高科学影响力论文

科学影响力排名前 20 位的研究如表 2-1 所示，包括 2 项社论材料，2 项回顾性研究和 16 项试验性研究。其研究内容提示 Health AI 技术被广泛应用于“辅助检测”、“疾病诊断和分类”、“遗传学和基因组学”等医学领域，涉及疾病主要是“癌症和脑部疾病”；“卷积神经网络”是其主要算法。被引次数排名前 20 位的研究中，17 项研究均涉及图像识别技术，主要用于疾病的诊断和辅助检测。如一项哈佛大学的研究 (Esteva et al., 2017) 采用多中心、大样本的试验方法，将不同水平皮肤科医生和卷积神经网络 (Convolutional Neural Networks, CNNs) 对皮肤癌诊断的结果进行了比较，结果显示 CNNs 对皮肤癌的漏诊率较医生更低、良性肿瘤误诊率更低。截至 2020 年 9 月，该研究在 Scopus 中被引次数高达 2631 次。此外，深度学习算法应用于计算机辅助检测 (Shin et al., 2016) 和糖尿病视网膜病变 (Gulshan et al., 2016) 的研究的被引次数均达到 1500 次左右 (表 1)，远高于其他同类研究。表明医学图像分析是健康医疗人工智能的研究热点，从被引规模上看，也显示了该领域学术共同体的规模较大，吸引了大批研究者。

参与高科学影响力的 Health AI 研究项目的机构共 26 所，其中除 Google 为企业以外，其余 25 家机构均为大学 / 医院。美国的哈佛医学院和荷兰的内梅亨大学均参与 2 项研究，其余机构则均为 1 项。其中，来自中国的机构有广州医科大学、四川大学和广州再生医学与健康广东省实验室。高科学影响力的研究来自 12 个国家 / 地区，主要来源为欧美国家。其中，美国参与了 9 项 (45%) 研究，荷兰参与了 3 项 (15%)，加拿大参与了 2 项 (10%)，其余国家 / 地区均为 1 项，亚洲国家仅包含中国、以色列、卡塔尔和韩国。20 项研究共发表在 11 类学术期刊上，其中 IEEE Transactions on Medical Imaging 是发表相关研究最多的期刊。

## 2.6. 全球 Health AI 细分领域分析

### (1) 全球总体情况

2015—2019 年间，除机器学习之外的五类 AI 细分领域中，文章发表数量总体相对稳定，部分年份的发表数量有小幅波动。而机器学习这一细分领域的文章发表数量在过去五年中则呈指数增长，增长速度迅速 (图 2-14)。

以国家为单位统计 AI 各细分领域的发文量时，美国、中国和印度位列前三 (表 2-2)。各个国家均对机器学习这一细分领域表现出较高关注，此外，中国和美国还较为关注自然语言处理领域，英国和日本则较为关注机器人领域。

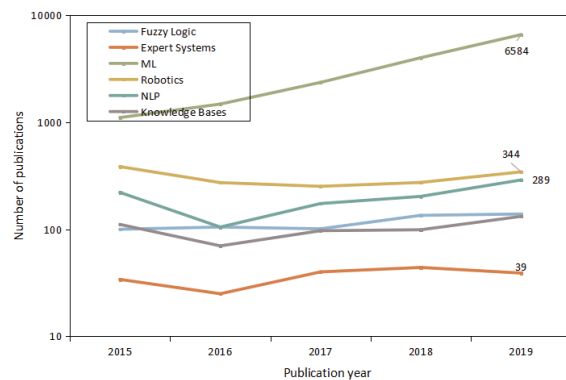


图 2-14 2015—2019 年全球 Health AI 细分领域发文量

## (2) 各医学研究方向中 Health AI 细分领域的发文情况

基于 SciVal 平台内嵌的经济合作与发展组织的 FORD (Fields of Research and Development) 学科分类标准, 将医学研究方向分为五类, 分别是“Clinical Medicine”、“Health Science”、“Basic Medicine”、“Medical Biotechnology”和“Medical Engineering”。在各个医学研究方向中, 最受研究人员关注的方向是 Health Science 和 Clinical Medicine, 这两个方向提供了大量供人工智能学习或训练的数据。

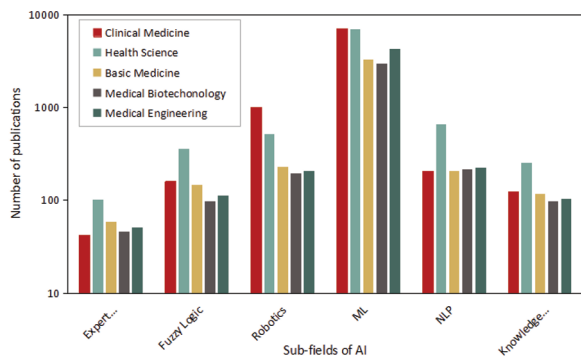


图 2-15 各医学研究方向中 Health AI 细分领域的发文情况

## (3) Health AI 研究的疾病谱分布

不同疾病谱与人工智能技术结合的深度也存在一定差异。相关文章发表数量最多的疾病分类包括各类癌症、以糖尿病、中风和心血管疾病为代表的慢性心脑血管疾病和以癫痫、帕金森和痴呆为代表的神经系统相关疾病。传染病领域的发文量较少, 仅疟疾、结核、艾滋病、下呼吸道感染等少部分疾病受到研究人员的关注。

Health AI 研究的疾病谱分布存在着长尾效应, 大部分疾病的 AI 相关研究发文量较低。这些疾病可能尚未被科研人员所重视, 未来可能成为有较高潜力的研究方向。全球 ICD 疾病与 Health AI 细分领域的发文量映射表见表 2-4。

疾病负担排名前 20 位的疾病中, 发文量与疾病负担呈现正相关关系 (图 2-16)。母婴疾病、先天性疾病等部分疾病负担较大但发文量少的原因还有待进一步研究。

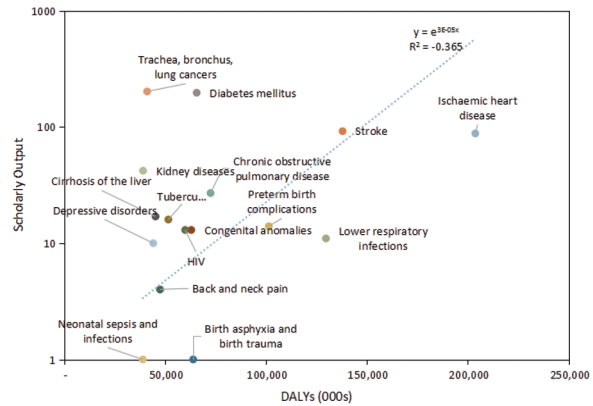


图 2-16 疾病负担排名前 20 位的疾病的发文量与疾病负担相关图

## 2.7. 中国 Health AI 细分领域分析

中国在 Health AI 领域表现较为活跃 (图 2-17)。2015-2019 年间, 中国 Health AI 领域总体的发文量增长迅速, 其中增长势头最强劲的领域同样是机器学习这一细分领域。

中国与全球不同疾病 Health AI 类文章发文量排序比较见图 2-18。中国与全球 Health AI 发文量最多的疾病均为支气管及肺部癌症。此外, 国内获得更多关注的疾病包括肝癌、甲状腺癌和胃癌; 全球获得较多关注的疾病则包括心肌症、结直肠癌和肾脏疾病。这可能与不同国家的高发疾病谱的差异以及所能获得的医疗大数据的人群分布差异存在关联。

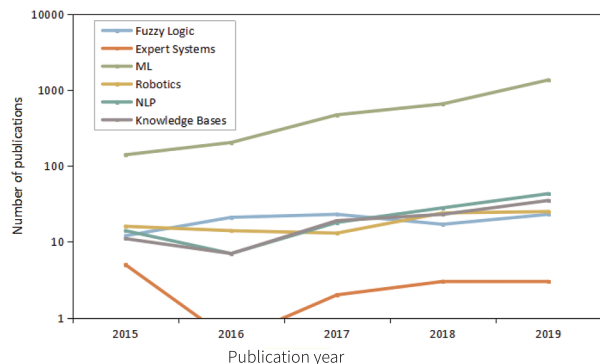


图 2-17 2015—2019 年中国 Health AI 各细分领域发文量对比

Rank	Global	Rank	China
1	Trachea, bronchus, lung cancers	1	Trachea, bronchus, lung cancers
2	Diabetes mellitus	2	Alzheimer's disease and other dementias
3	Brain and nervous system cancers	3	Diabetes mellitus
4	Breast cancer	4	Brain and nervous system cancers
5	Melanoma and other skin cancers	5	Breast cancer
6	Alzheimer's disease and other dementias	6	Epilepsy
7	Prostate cancer	7	Ischaemic heart disease
8	Stroke	8	Liver cancer
9	Ischaemic heart disease	9	Melanoma and other skin cancers
10	Epilepsy	10	Parkinson's disease
11	Parkinson's disease	11	Skin diseases
12	Cardiomyopathy, myocarditis, endocarditis	12	Prostate cancer
13	Skin diseases	13	Stroke
14	Colon and rectum cancers	14	Thyroid cancer
15	Kidney diseases	15	Stomach cancer
.....	.....	.....	.....
16	Liver cancer	21	Cardiomyopathy, myocarditis, endocarditis
22	Thyroid cancer	22	Kidney diseases
24	Stomach cancer	25	Colon and rectum cancers

图 2-18 中国与全球各疾病领域 Health AI 类文章发文章量排序比较

表 2-1 高科学影响力论文 (Top20)

序号	主题	国家 / 地区	机构	出版物来源	被引次数
1	深卷积神经网络诊断皮肤癌达到皮肤科专家的水平 (Esteva et al., 2017)	USA	Stanford Univ	Nature	2631
2	深卷积神经网络用于计算机辅助检测 (Shin et al., 2016)	USA	NIH	IEEE Transactions on Medical Imaging	1567
3	深度学习算法用于检测糖尿病视网膜病变 (Gulshan et al., 2016)	USA	Google Res	JAMA - Journal of the American Medical Association	1484
4	深层神经网络用于脑肿瘤分割 (Havaei et al., 2017)	Canada	Univ Sherbrooke	Medical Image Analysis	945
5	3D 卷积神经网络和全连接 CRF 用于精确分割脑损伤 (Kamnitsas et al., 2017)	UK	Imperial Coll London	Medical Image Analysis	925
6	卷积神经网络用于医学图像分析 (Tajbakhsh et al., 2016)	USA	Arizona State Univ	IEEE Transactions on Medical Imaging	859
7**	深度学习用于医学图像分析 (Shen et al., 2017)	USA; South Korea	Univ N Carolina; Korea Univ	Annual Review of Biomedical Engineering	787
8	卷积神经网络用于 MRI 脑肿瘤分割 (Thaha et al., 2019)	Portugal	University of Minho	IEEE Transactions on Medical Imaging	778
9*	预测未来大数据、机器学习和临床医学 (Obermeyer and Emanuel, 2016)	USA	Harvard Med Sch	New England Journal of Medicine	618
10*	深度学习在医学影像学的应用 (Greenspan et al., 2016)	Israel	Tel-Aviv University	IEEE Transactions on Medical Imaging	616
11	机器学习用于全球网格化土壤信息预测的研究 (Hengl et al., 2017)	Netherlands	ISRIC World Soil Informat	PLoS ONE	581
12	计算放射学系统用于解码射线表现型的研究 (van Griethuysen et al., 2017)	USA	Harvard Med Sch	Cancer Research	510
13	机器学习在人类的剪接编码对于疾病遗传决定因素中的应用 (Xiong et al., 2015)	Canada	Univ Toronto	Science	502
14	图像的深度学习用于识别医学诊断和可治疗疾病 (Kermary et al., 2018)	China; USA	Guangzhou Med Univ; Univ Calif San Diego; Sichuan Univ; Guangzhou Regenerat Med & Hlth Guangdong Lab; Vet Adm Healthcare Syst	Cell	493
15	深度学习算法用于诊断乳腺癌女性淋巴结转移 (Bejnordi et al., 2017)	Netherlands	Radboud Univ Nijmegen	JAMA - Journal of the American Medical Association	488
16**	机器学习用于遗传学和基因组学注释的研究 (Libbrecht and Noble, 2015)	USA	Univ Washington	Nature Reviews Genetics	456
17	多视图卷积网络减少 CT 图像中的肺结节检测假阳性 (Setio et al., 2016)	Netherlands	Radboud Univ Nijmegen	IEEE Transactions on Medical Imaging	449
18	深卷积神经网络对间质性肺疾病的肺模式分类 (Anthimopoulos et al., 2016)	Switzerland	Univ Hosp Bern	IEEE Transactions on Medical Imaging	441
19	局部敏感深度学习在常规结肠癌组织学图像中检测和分类细胞核 (Sirinukunwattana et al., 2016)	Qatar; England	Qatar Univ; Univ Warwick	IEEE Transactions on Medical Imaging	406
20	基于 DNA 甲基化机器学习的中枢神经系统肿瘤分类 (Capper et al., 2018)	Germany	Univ Hosp Heidelberg; German Canc Res Ctr; NCT Heidelberg KiTZ;	Nature	400

注: \*type of Editorial Material; \*\*type of Review.

表 2-2 发文量前十国家在 Health AI 细分领域的发文量

	Expert Systems	Fuzzy Logic/ Computer Heuristics	Robotics	Machine Learning	Natural Language Processing	Knowledge Bases	Total
美国	28	55	464	5095	515	193	6350
中国	13	96	92	2806	110	95	3212
印度	9	152	33	1442	42	25	1703
英国	12	21	106	1107	40	47	1333
德国	20	8	92	767	35	39	961
韩国	4	14	97	625	14	13	767
日本	2	15	114	581	29	11	752
意大利	11	11	171	434	13	18	658
伊朗	18	67	9	388	4	5	491
土耳其	8	28	26	308	4	2	376

表 2-3 疾病负担前 20 位的疾病的发文量 \*

Rank	Causes	DALYs (000s)	Scholarly Output
1	Ischaemic heart disease	203,700	88
2	Stroke	137,941	92
3	Lower respiratory infections	129,690	11
4	Preterm birth complications	101,397	14
5	Chronic obstructive pulmonary disease	72,512	27
6	Diabetes mellitus	65,666	197
7	Birth asphyxia and birth trauma	63,928	1
8	Congenital anomalies	62,980	13
9	HIV/AIDS	59,951	13
10	Tuberculosis	51,643	16
11	Back and neck pain	47,515	4
12	Cirrhosis of the liver	45,287	17
13	Depressive disorders	44,175	10
14	Trachea, bronchus, lung cancers	41,121	202
15	Kidney diseases	39,079	42
16	Neonatal sepsis and infections	39,009	1

\* 注：ICD 疾病负担前 20 名的原因中包含车祸、摔倒等非疾病原因，在本表中被排除。



表 2-4 全球 Health AI 领域发文量前 20 位的疾病

	Fuzzy Logic/ Computer Heuristic	Expert Systems	ML	Robotics	NLP	Knowledge Bases	Total
Trachea, bronchus, lung cancers	5	1	188	5	0	3	202
Diabetes mellitus	5	3	171	2	6	10	197
Brain and nervous system cancers	17	0	119	1	2	0	139
Breast cancer	3	3	120	1	2	0	129
Melanoma and other skin cancers	0	1	118	3	1	0	123
Alzheimer's disease and other dementias	5	0	113	0	2	2	122
Prostate cancer	0	0	97	0	0	1	98
Stroke	4	0	73	13	2	0	92
Ischaemic heart disease	4	3	79	0	1	1	88
Epilepsy	2	0	83	0	0	2	87
Parkinson's disease	2	1	71	0	1	0	75
Cardiomyopathy, myocarditis, endocarditis	0	1	58	0	0	0	59
Skin diseases	0	0	41	1	1	0	43
Colon and rectum cancers	0	0	39	2	1	0	42
Kidney diseases	2	0	37	2	1	0	42
Liver cancer	4	0	30	2	0	2	38
Glaucoma	1	1	34	0	0	0	36
Macular degeneration	0	0	33	0	0	0	33
Autism and Asperger syndrome	2	0	23	1	0	1	27
Chronic obstructive pulmonary disease	0	0	27	0	0	0	27

## 第三章 科学技术交叉

### 1. 数据与指标

#### 1.1. 数据来源

采用由北京大学健康医疗大数据国家研究院和爱思唯尔双方融合后的数据集，该数据集含有 25,717 篇与 Health AI 主题有关的科学出版物。该部分基于爱思唯尔的 Scival 数据平台对学术论文被专利引用的识别和数据统计进行数据分析。从学术发表物的角度来看，这是“前向引证”，表明研究成果是否随后被用于专利领域。其中专利数据库包含了对全球五大专利数据库的数据统计，分别为：全球专利局 WIPO、美国专利商标局 USPTO、欧洲专利局 EPO、日本专利局 JPO、英国专利局 UKPO。

#### 1.2. 学术界与产业界的知识流动指标

施引专利数 (Citing-patent count)：施引专利数统计了被评估的文集作为一个整体被专利引用的专利数量，体现了该文集对于专利产出的贡献。

被专利引用的文章数 (Patent-cited scholarly output)：被专利引用的文章数统计了被评估的文集中被专利引用的文章数量，体现了该文集的技术转化程度。

### 2. 分析结果

#### 2.1. 学术界与产业界的知识流动总体状况

该领域内学术界与产业界知识流动详见图 3-1，美国文章的施引专利数量 (419) 远高于中国 (83)，约为全球总量 (863) 的一半；被专利引用的文章数 (172) 也远高于中国 (44)，同样约达到了全球总量 (354) 的一半。中国文章的施引专利数量 (83) 约为美国 (419) 的 1/5；被专利引用的文章数 (44) 约为美

国 (172) 的 1/4。中国的 44 篇被专利引用的文章中，参与贡献文章数排名前五的中国机构为中国科学院、清华大学、浙江大学、香港中文大学和深圳先进技术研究院。出现频率最高的五个关键词为“神经网络” (Neural Network)、“深度学习” (Deep Learning)、“心律失常” (Heart Arrhythmia)、“人工智能” (Artificial Intelligence) 和“计算机辅助诊断” (Computer-aided Diagnose)。美国的 172 篇被专利引用的文章中，参与贡献文章数排名前五的美国机构为哈佛大学、美国国家卫生研究院 (National Institutes of Health, NIH)、麻省理工学院、范德堡大学和加州大学洛杉矶分校。出现频率最高的五个关键词“深度学习” (Deep Learning)、“机器学习” (Machine Learning)、“神经网络” (Neural Network)、“乳腺钼靶” (Mammography) 和“阿尔茨海默病” (Alzheimer's Disease)。

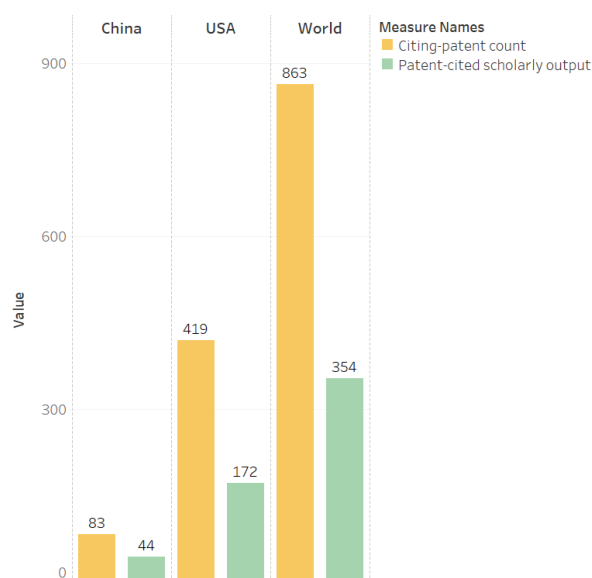


图 3-1 学术界与产业界的知识流动

## 2.2. 对 Health AI 技术产生影响的基础性研究

对 Health AI 技术产生较高影响的前 10 项基础性研究见表 3-1。2015-2020 年专利高被引前 10 位的研究包括 9 项试验性研究和 1 项回顾性研究，其中 6 项研究处于高科学影响力排名前 20 位，这表明 Health AI 领域的研究在科学影响力和专利影响力上具有一定的正相关关系。

从研究内容来看，被专利引用前 10 位的研究应用领域包括计算机辅助检测、疾病诊断和分类、遗传学和基因组学等内容，其中 7 项研究涉及图像识别技术。被专利引用最多的研究为 NIH 于 2016 年将深度卷积神经网络用于计算机辅助检测 (Shin et al., 2016) 的研究，被引次数高达 23 次；奈梅根大学在 2016 年有关 CT 图像肺结节假阳性检测 (Setio et al., 2016)，卡塔尔大学和华威大学在 2016 年开展的有关结肠癌检测 (Sirinukunwattana et al., 2016) 等辅助检测研究也分别被专利引用 14 和 12 次；2015 年华盛顿大学 (Libbrecht and Noble, 2015) 和多伦多大学 (Xiong et al., 2015) 将机器学习用于遗传学和基因组学的研究，被引次数分别达到了 20 次和 13 次。另外还有 50% 的研究将 Health AI 技术应用于疾病的诊断和分类上，包括冠状动脉相关研究 2 项，细胞核分类、皮肤癌诊断、组织病理学诊断相关研究各 1 项。其中，飞利浦公司在 2017 年做的有关深度学习在冠状动脉病变血流动力学评估中的准确性研究 (Freiman et al., 2017) 被引用次数高达 15 次。

参与专利高被引前 10 位的研究机构共 10 家，其中有 2 家机构是企业，分别是被引排名第 3 的飞利浦公司 (Freiman et al., 2017) 和被引排名第 5 的西门子公司 (Itu et al., 2016)，其开展的研究均与机器学习应用于冠状动脉检测相关，且被引用次数较高，分别是 15 次和 13 次，这在一定程度上说明该研究领域是公司近期产品研发的重点。除荷兰的奈梅根大学有 2 项研究被引次数位于前 10 位外，其余单位仅有 1 项研究。50% 的专利引用量前 10 位的研究来自美国，除卡塔尔和以色列 2 个亚洲国家外，其他国家均为欧洲国家。被专利引用次数前 10 位的研究中有 2 项 2015 年的研究，6 项 2016 年的研究，2 项 2017 年的研究，近两年研究成果未出现在高被引前 10。专利高被引研究出现的年份均较早，有一定延迟性，可能与专利申请周期较长有关。

表 3-1 对 Health AI 技术产生影响的基础性研究 (Top 10)

序号	主题	国家 / 地区	机构	出版物来源	被专利引用的次数
1	深卷积神经网络用于计算机辅助检测 (Shin et al., 2016)	USA	NIH	IEEE Transactions on Medical Imaging	23
2*	机器学习用于遗传学和基因组学研究 (Libbrecht and Noble, 2015)	USA	Univ Washington	Nature Reviews Genetics	20
3	部分容积效应的冠状动脉腔自动分割算法在基于 CCTA 的冠状动脉病变血流动力学评估的准确性 (Freiman et al., 2017)	Israel	Philips Med Syst Technol Ltd	Medical Physics	15
4	多视图卷积网络减少 CT 图像中的肺结节检测假阳性 (Setio et al., 2016)	Netherlands	Radboud Univ Nijmegen	IEEE Transactions on Medical Imaging	14
5	机器学习用于冠状动脉断层扫描 (Itu et al., 2016)	USA	Siemens AG	Journal of Applied Physiology	13
6	机器学习在人类的剪接编码对于疾病遗传决定因素中的应用 (Xiong et al., 2015)	Canada	Univ Toronto	SCIENCE	13
7	自动学习卷积神经网络用于核分割 (Xing et al., 2016)	USA	Univ Florida	IEEE Transactions on Medical Imaging	12
8	深度神经网络诊断皮肤癌达到皮肤病专家的水平 (Esteva et al., 2017)	USA	Stanford Univ	Nature	12
9	局部敏感深度学习在常规结肠癌组织学图像中检测和分类细胞核 (Sirinukunwattana et al., 2016)	Qatar; England	Qatar Univ; Univ Warwick	IEEE Transactions on Medical Imaging	12
10	深度学习提高组织病理学诊断准确性和效率 (Litjens et al., 2016)	Netherlands	Radboud Univ Nijmegen	SCIENTIFIC REPORTS	11

注: \* Review.

### 2.3. Health AI 细分领域科学 - 技术交叉

在专家系统及决策规则这两个 AI 子领域中，被专利引用的文献及引用文献的专利数量均为 0，这两个领域的产研融合有待进一步发展。在其余四个 AI 子领域之中，机器学习依然是被引及引用最多的学科领域，机器人领域的被引数量也相对较高（表 3-2）。但总体被专利引用文献数量较每年发表的文献数量仍然有两个数量级的差距，这表明产研结合的空间和前景仍然十分广阔。

表 3-2 全球文献与专利引用情况

	Robotics		Machine Learning		Natural Language Processing		Knowledge Bases	
	施引专利数	被专利引用文献数	施引专利数	被专利引用文献数	施引专利数	被专利引用文献数	施引专利数	被专利引用文献数
2015	6	3	77	38	3	1	0	0
2016	5	2	206	78	0	0	0	0
2017	1	1	151	74	2	2	3	1
2018	2	1	152	92	2	1	0	0
2019	0	0	15	11	0	0	0	0

## 第四章 科学社会交互

### 1. 数据与指标

#### 1.1. 数据来源

采用由北京大学健康医疗大数据国家研究院和爱思唯尔提供双方融合后的数据集，该数据集含有 25,717 篇与 Health AI 主题有关的科学出版物。该部分基于爱思唯尔旗下的数据平台 PlumX 进行数据统计分析。PlumX 平台对各类学术研究成果的交流、分享以及互动进行广泛的数据统计和研究，进而对学术研究成果的社会影响力进行评估。

#### 1.2. 科学社会交互指标

**利用率指数 (Usage)**<sup>11</sup>：利用率指标体现了一篇文章被阅读或者以其他方式被研究的程度，包含该文章被点击、下载、阅读、摘要阅读、图书馆收录等数据。利用率是研究人员继被引次数之后另一个想要了解的统计指标。

**注意力指数 (Capture)**：注意力指标体现该研究工作被引起注意并被反复研究的程度，包含一篇文章的“书签”、“最喜欢”、“读者”、“导出”、“订阅”等数据。

**多媒介提及指数 (Mention)**：提及指标体现了研究工作被各种媒体所提及的程度，包含一篇文章被博客提及、评论、归属于论坛主题、新闻提及、列为参考文献、归为综述内容等数据。

**社交媒体传播指数 (Social Media)**：社交媒体指标体现了研究工作在社交媒体上被传播的程度，包含 YouTube、Facebook、新浪微博、Reddit、Twitter 等媒体的数据。

PlumX 分别给出每篇文章在上述 4 个指标上的分值，但并未建立一个综合指数。为重点分析研究成果的社会影响力，本报告仅纳入多媒介提及指数和社交媒体传播指数这两个指标进行分析。

<sup>11</sup> <https://plumanalytics.com/learn/about-metrics/>



以下详细分析全球和中国学者发表的社交媒体传播指数排名前 10 位的研究，见表 4-1 和表 4-2。全球范围内被社交媒体报道量最高的 10 项研究包括 2 项回顾性研究和 8 项试验性研究；中国被大众媒体报道量最高的 10 项研究包括 1 项回顾性研究和 9 项试验性研究。全球社交媒体报道量 Top10 的研究与高科学影响力 Top20、高技术影响力 Top10 的研究均未重复，在一定程度上体现了社交媒体关注的内容与科学共同体关注重点不尽相同。

从报道内容来看，由于统计的社交媒体源侧重 Twitter、Facebook 原因，中国的研究被社交媒体传播频次显著较低，全球范围内被传播次数最多的研究为北卡罗来纳大学在 2017 年做的“有关自闭症谱系障碍高危婴儿的早期大脑发育” (Hazlett et al., 2017) 的研究，被报道量高达 100,260 次，远高于其他研究被报道的次数，全球范围内其他被高频报道的研究还涉及医疗支出、精神疾病、临床结局预测、放射学、肿瘤学以及常见病多发病的评估等研究方向。而中国被大众媒体报道次数最多的研究为深圳大学所做的有关“厌恶感知神经基础在种族偏见中的作用” (Liu et al., 2015) 的研究，被报道 182 次，其余 9 项被报道次数均未超过 100 次，内容涉及心理疾病、医疗关系、癌症、帕金森病和心血管等常见病、多发病。相比较科学共同体的关注点，社交媒体报道内容更多涉及常见病、多发病以及医疗卫生领域公众所关心的问题，且与当下的社会需求紧密关联，如美国国家经济研究局和麻省理工学院合作的有关晚年医疗支出预测模型 (Einav et al., 2018) 的研究与美国老龄化加剧的社会现实相联系。

我国被社交媒体高频报道和传播的 10 项研究均来自各大高校的研究结果，其中中国电子科技大学参与度较其他高校更高；而全球范围内被高频报道的研究所涉及的机构类型较为多元化，除大学外还包括 IBM 公司 (Bedi et al., 2015)、美国放射学学会和犹太民族健康协会 (Geis et al., 2019) 等企业或社会机构。从相关研究的发表时间看，全球范围内被高频报道前 10 位的研究中有 4 项研究于 2019 年被发表，我国被高度报道前 10 的研究中有 5 项研究于 2019 年发表，相对于高科学影响力和高技术影响力的研究仅有 1 项于 2019 年发表（存在较长的引用时滞），提示社交媒体对科学研究的反馈具有及时性。



表 4-1 全球学者发表的社交媒体传播指数排名前 10 位的研究

序号	主题	报道量	机构
1	自闭症谱系障碍高危婴儿的早期大脑发育 (Hazlett et al., 2017)	100,260	Univ N Carolina
2	美国晚年医疗支出预测模型 (Einav et al., 2018)	25,984	Natl Bur Econ Res; MIT
3	自动分析预测高危青年的精神病发作 (Bedi et al., 2015)	23,004	IBM TJ Watson Res Ctr
4	机器学习和逻辑回归在临床预测模型的应用 (Christodoulou et al., 2019)*	10,584	Katholieke Univ Leuven; Leiden Univ
5	人工智能伦理在放射学中的应用 (Geis et al., 2019)	5,177	Amer Coll Radiol; Natl Jewish Hlth
6	胰腺囊肿患者治疗的多模态试验 (Springer et al., 2019)	3,537	Johns Hopkins Univ
7	机器学习和应用程序预测埃博拉患者的预后 (Colubri et al., 2016)	2,896	Harvard Univ; Broad Inst MIT & Harvard
8	机器学习用于糖尿病风险分层 (Maniruzzaman et al., 2018)	2,220	AtheroPoint LLC, Stroke Monitoring & Diagnost Div; Global Biomed Technol
9	机器学习“应对伦理挑战” (Char et al., 2018)*	1,734	Stanford Univ
10	判别分析和机器学习方法在 DLBCL COO 分类免疫组化算法中的应用 (Perfecto-Avalos et al., 2019)	1,344	Ctr Med Dr Ignacio Chavez ISSSTESON

表 4-2 中国学者发表的社交媒体传播指数排名前 10 位的研究

序号	主题	报道量	机构
1	厌恶感知神经基础在种族偏见中的研究 (Liu et al., 2015)	182	Shenzhen Univ
2	抗精神病药物与心源性猝死的预测、管理和未来挑战 (Zhu et al., 2019)*	80	Jining Med Univ
3	卷积神经网络在临床文本中的医疗关系分类应用 (He et al., 2019)	74	Harbin Inst Technol
4	帕金森病患者皮质回缩的研究 (Xu et al., 2017)	48	University of Electronic Science & Technology of China Univ Elect Sci & Technol China; Chinese Academy of Sciences Shenzhen Institute of Advanced Technology
5	人工智能在抑郁症与慢性疲劳综合征相关性中的应用 (Zhang et al., 2019)	42	Beijing Univ Chinese Med
6	催产素在相互感知信号和外部社会线索之间的作用 (Yao et al., 2018)	36	Univ Elect Sci & Technol China
7	深度学习与非深度学习在磁共振成像分类寻找前列腺癌中的应用 (Wang et al., 2017)	36	Huazhong Univ Sci & Technol
8	深度学习识别癌症特异性结合位点 (Wang et al., 2019)	30	Shaanxi Normal Univ
9	机器学习预测心血管事件 (Chen et al., 2019)	24	South China Univ Technol Guangzhou
10	深度投票模型在自动地理萎缩分割的应用 (Ji et al., 2018)	24	Nanjing Univ Sci & Technol

## 第五章 人类 - 机器协同 (AI 临床试验)

### 引言

医疗领域的信息化进程累积了海量的人类健康数据，正有越来越多的临床医生与计算机科学家合作致力于利用这些宝贵的健康数据挖掘信息、开发产品，以提升人类健康水平并减轻医疗卫生体系现有的沉重负担。据估计，全球医疗人工智能市场的价值将从 2018 年的 20 亿美元增长到 2025 年的 36 亿美元，年增长率将达到 50%<sup>12</sup>。人工智能相关的健康管理设备与临床决策支持系统已经成为当前医疗领域的研究热点之一。现有的人工智能应用于医疗领域的研究涵盖了多种应用场景，包括疾病筛检，疾病严重程度分类，辅助诊断，疾病预后预测，临床决策支持和治疗方案推荐等。

深度学习是人工智能的一个分支，在医学成像领域表现出落地应用的前景。随着越来越多研究成果的发表，各界对医学成像等领域的深度学习研究兴趣日益浓厚。美国斯坦福大学自 2017 年首份 AI 指数报告列出“人类级表现里程碑” (Human-Level

Performance Milestones) 清单后，健康医疗领域人工智能每年均有入选，包括 2017 年入选的“人工智能诊断皮肤癌”、2018 年入选的“人工智能用于前列腺癌的分级”和 2019 年入选的“人工智能以专家级的准确性检测糖尿病视网膜病变”。在《自然》杂志 2017 年发表的一篇文章中，Esteva 等人描述了一个基于数据集的人工智能系统，包含 2032 种不同疾病的 129,450 张临床图像，并比较其与 21 名通过认证的皮肤科医生的诊断水平。比较结果发现人工智能系统有能力对皮肤癌进行分类，其能力可比肩皮肤科医生<sup>14</sup>。2018 年，谷歌开发了一个深度学习系统，对前列腺癌进行自动分级的总体准确率可达 70%，而美国委员会认证的病理学家在研究中的平均准确率为 61%<sup>15</sup>。2019 年的一项研究表明，深度学习算法能够以专家级的准确性检测糖尿病视网膜病变 (DR)，临床验证显示，其准确度明显高于专家<sup>16</sup>。

<sup>12</sup> Nagendran M, Chen Y, Lovejoy C A, et al. Artificial intelligence versus clinicians: systematic review of design, reporting standards, and claims of deep learning studies [J]. *bmj*, 2020, 368:

<sup>13</sup> Raymond Perrault, Yoav Shoham, Erik Brynjolfsson, Jack Clark, John Etchemendy, Barbara Grosz, Terah Lyons, James Manyika, Saurabh Mishra, and Juan Carlos Nieves, “The AI Index 2019 Annual Report”, AI Index Steering Committee, Human-Centered AI Institute, Stanford University, Stanford, CA, December 2019.

<sup>14</sup> Esteva, A., Kuprel, B., Novoa, R. et al. Dermatologist-level classification of skin cancer with deep neural networks. *Nature* 542, 115–118 (2017).

<sup>15</sup> <https://ai.googleblog.com/2018/11/improved-grading-of-prostate-cancer.html>

<sup>16</sup> Ruamviboonsuk, P., Krause, J., Chotcomwongse, P. et al. Deep learning versus human graders for classifying diabetic retinopathy severity in a nationwide screening program. *npj Digit. Med.* 2, 25 (2019).

近两年，人们可经常看到一些媒体新闻出现诸如“研究发现，谷歌人工智能比医生早一年发现肺癌”以及“人工智能比医生更擅长诊断皮肤癌”这样的标题。媒体宣传极大增加了公众和商业对健康医疗人工智能的兴趣，也促进了技术的发展和运用。但实际上，背后的研究方法和偏倚风险尚未得到详细的检验。根据《英国医学杂志》(The BMJ)2020年发表的一项分析<sup>17</sup>，伦敦帝国理工学院的研究人员回顾了过去10年发表的研究结果，系统地检查研究设计、报告标准、偏倚风险，并将深度学习算法在医学成像方面的表现与临床专家进行比较。结果显示，目前很少有前瞻性的深度学习研究和随机试验。大多数非随机化试验不具有前瞻性，存在较高的偏倚风险，并偏离现有的报告标准。大多数研究缺乏数据和代码可用性，而且人类对照组通常很小。目前存在着许多关于AI与临床医生比肩或优于临床医生诊断能力的夸大说法，这在社会层面上对患者安全和人口健康构成了潜在风险。过分的承诺，会使研究容易被媒体和公众曲解，结果可能不符合患者的最佳利益，也无法最大限度地保障患者的安全，而最佳策略是确保有高质量和透明度的报告作为证据基础。

当前面向人工智能相关的医疗应用设备与系统的研究、评价与审批制度并不完善。2020年Nature同时发表了两篇人工智能相关的临床试验研究的报告规范指南<sup>18,19</sup>，以推进人工智能相关的医疗应用设备与系统研究的规范化进程。国务院办公厅于2016年发布了《关于促进和规范健康医疗大数据应用发展的指导意见（国办发[2016]47号）》，提出通过“互联网+健康医疗”探索服务新模式培育发展新业态的目标。自此之后，我国人工智能相关的医疗应用设备与系统的临床试验迅速发展，临床试验研究机构承接及牵头试验的能力也大幅提升。

对于“研究条件”下人机可媲美能否转化为“真实世界”人机可媲美尚有争议，本部分重点介绍全球以及我国人工智能医疗设备与系统开展临床试验的数量、人群、干预措施及研究设计等，旨在描述全球以及我国正在开展的健康医疗人工智能临床试验的基本特征及变化趋势。

---

<sup>17</sup> <https://www.bmj.com/content/368/bmj.m689>

<sup>18</sup> Cruz Rivera S, Liu X, Chan A-W, et al. Guidelines for clinical trial protocols for interventions involving artificial intelligence: the SPIRIT-AI extension [J]. Nature Medicine, 2020, 26(9): 1351-1363.

<sup>19</sup> Liu X, Cruz Rivera S, Moher D, et al. Reporting guidelines for clinical trial reports for interventions involving artificial intelligence: the CONSORT-AI extension [J]. Nature Medicine, 2020, 26(9): 1364-1374.

## 2. 数据与指标

### 2.1. 数据来源

本部分的数据来源于国际通用的临床试验登记与信息公示平台 ClinicalTrials.gov (<https://clinicaltrials.gov/>)。基于该数据库提取人工智能相关的临床试验数据的流程如图 5-1 所示：首先，以“Deep Learning”、“Artificial Intelligence”、“Machine Learning”、“AI”为关键词对临床试验的干预措施及标题进行初筛 (n=782)，然后由医疗领域的专业人士进行人工复筛，排除非人工智能相关的临床试验 (n=304)；其次根据研究疾病、干预措施、试验分期、研究机构均相同筛选出可能的重复临床试验，专业人士人工判断是否为重复临床试验，排除重复临床试验 (n=1)；排除招募状态为暂停、终止、撤销的临床试验 (n=15)。最终，筛选出 462 例人工智能相关的临床试验纳入分析。

### 2.2. 分析指标

从最终筛选的人工智能相关的临床试验登记数据中提取信息，主要包括：(1) 基本信息：题目、研究机构、申报日期、所在国家等；(2) 试验设计信息：目的、研究类型、干预措施、研究人群、试验分期、样本量等；(3) 试验实施信息：招募状态、试验结果等。

数据处理与统计分析使用 Python 3.7。根据名称中是否包含“university/college”、“hospital”、“company/Co., Ltd/Inc.”等标识，将机构分为两类：大学 / 医院类研究机构和企业。统计学方法包括描述性统计，使用数值（百分比）描述计数型数据。结果描述包括临床试验的数量、招募状态、研究机构、试验分期、研究类型、干预措施、研究人群、样本量的分布特征、时间变化趋势与国家分布比较，并对已有试验结果的临床试验进行总结描述。

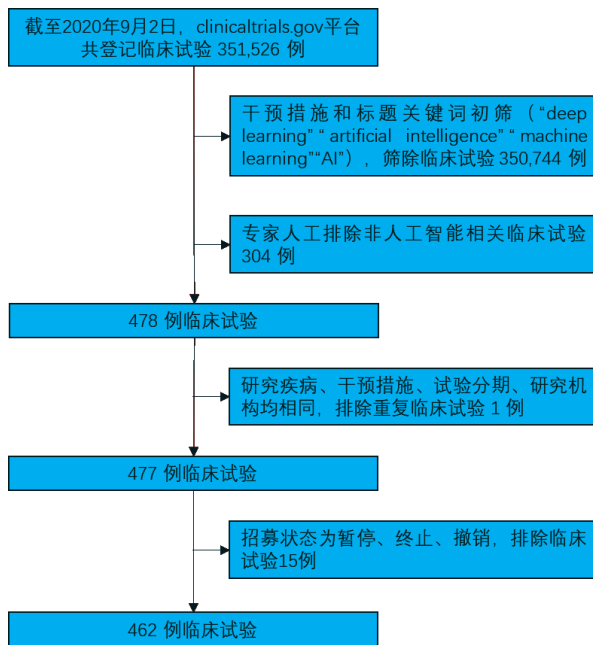


图 5-1 数据提取流程

### 3. 分析结果

#### 3.1. 临床试验数量

自 2006 年到 2020 年，ClinicalTrials.gov 平台共登记来自全球的 462 例人工智能相关的临床试验。全球及数量占比前十国家每年发起的人工智能临床试验数量的变化趋势见图 5-2。自 2017 年起，全球人工智能相关临床试验的新增数量呈快速递增趋势，其中主要发起国为中国和美国。中国人工智能相关临床试验的新增数量在 2017 年超越美国，成为全球开展人工智能相关临床试验新增数量最多的国家。2020 年，中国人工智能相关临床试验的新增数量达到 48 例，在全球新增人工智能相关临床试验中占比 27.3%（2020 年的数据统计截至 2020 年 9 月）。

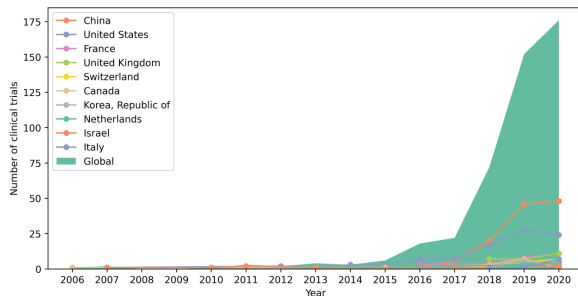


图 5-2 2006-2020 全球及数量占比前十国家发起的人工智能临床试验新增数量变化趋势

#### 3.2. 发起机构分布

2006-2020 年全球及中国发起人工智能临床试验的研究机构数量的时间变化趋势见图 5-3 和图 5-4。2017 至 2019 年，全球及中国从事过人工智能相关临床试验的研究机构数量均呈快速递增趋势。截至 2020 年，全球从事过人工智能相关临床试验的研究机构为 126 所，其中中国的研究机构为 27 所（21.4%）。

全球及中国人工智能相关临床试验数量排名前十位的研究机构见表 5-1。全球人工智能相关临床试验数量最多的前十位研究机构中，中国的研究机构占到 6 位。全球范围内人工智能临床试验发起量最多的研究机构为中国的中山大学（20 例）；排在第二位的是中国的山东大学（14 例）。

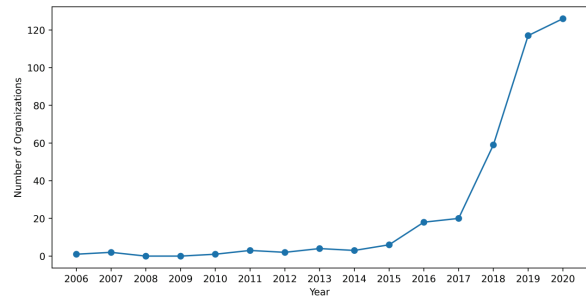


图 5-3 2006-2020 全球人工智能相关临床试验的研究机构数量的时间变化趋势

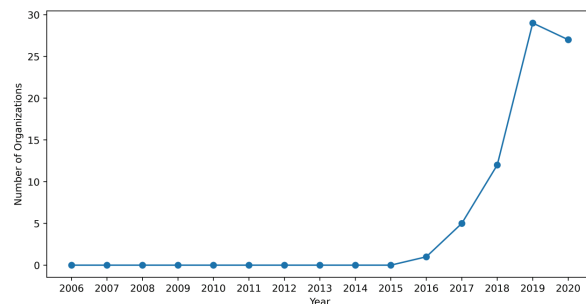


图 5-4 2006-2020 中国人工智能相关临床试验的研究机构数量的时间变化趋势

全球、中国及美国人工智能临床试验发起机构的类型分布见图 5-5。全球人工智能相关临床试验的发起机构中，大学/医院类机构有 171 所（59.2%），企业类机构有 118 所（40.8%）。中国人工智能相关临床试验的发起机构中，绝大部分为大学/医院类机构，有 48 所（88.9%）；企业类机构仅有 6 所（11.1%）。美国人工智能临床试验的发起机构中，大学/医院类机构有 39 所（60.0%），企业类机构有 26 所（40.0%）。

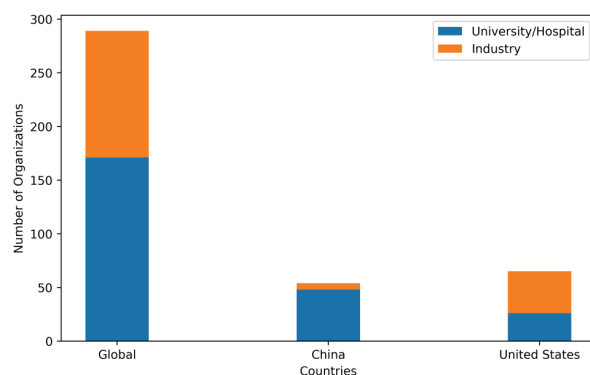


图 5-5 研究机构的类型

表 5-1 全球及中国人工智能相关临床试验数量排名前十位的研究机构

Global Organizations	Number of Clinical Trials	Chinese Organizations	Number of Clinical Trials
Sun Yat-sen University	20	Sun Yat-sen University	20
Shandong University	14	Shandong University	14
Dascena	12	The First Affiliated Hospital of Zhengzhou University	6
Mayo Clinic	7	Changhai Hospital	5
University Hospital, Basel, Switzerland	6	Sun Yat-Sen Memorial Hospital of Sun Yat-Sen University	5
The First Affiliated Hospital of Zhengzhou University	6	The University of Hong Kong	4
Changhai Hospital	5	Sixth Affiliated Hospital, Sun Yat-sen University	4
Sun Yat-Sen Memorial Hospital of Sun Yat-Sen University	5	Shanghai 10th People's Hospital	3
The University of Hong Kong	4	Chinese PLA General Hospital	3
Maastricht University Medical Center	4	Second Affiliated Hospital, School of Medicine, Zhejiang University	3

### 3.3. 临床试验分期

2006-2020 年全球人工智能相关临床试验分期的时间变化趋势见图 5-6。全球范围内，96.8% 的临床试验的分期被研究者划分为“Not Applicable”或者缺失，不能归入类似药物临床试验的四个分期的类别中。已有明确分期的 15 项临床试验中，10 项处于临床早期阶段（1-2 期），处于临床 3 期和 4 期的只有 5 项。

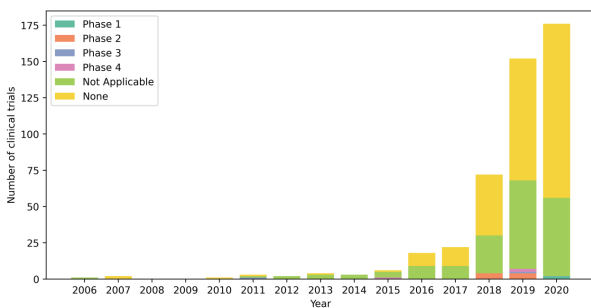


图 5-6 2006-2020 年全球人工智能相关临床试验分期的时间变化趋势

### 3.4. 研究类型分布

全球人工智能相关临床试验的研究类型分布见图 5-7。其中，观察型的研究为 272 例（58.9%），干预型的研究为 190 例（41.1%）。

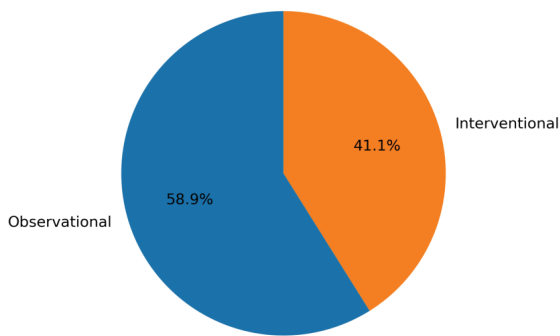


图 5-7 临床试验的研究类型

### 3.5. 干预措施类型

2006-2020 年全球人工智能临床试验干预措施的时间变化趋势见图 5-8。其中，干预措施以设备类干预（22.7%）和诊断试验类干预（19.3%）为主，其次为行为干预（5.4%）。自 2017 年起，设备类干预与诊断试验类干预的临床试验数量大幅增加。

2006-2020 年中国人工智能相关临床试验干预措施的时间变化趋势见图 5-9。其中，干预措施以诊断试验类干预（24.8%）和设备类干预（20.7%）为主，其次为操作类干预（5.8%）。2020 年，中国诊断试验类干预的临床试验数量大幅增加。

2006-2020 年美国人工智能相关临床试验干预措施的时间变化趋势见图 5-10。美国人工智能相关临床试验中，干预措施以设备类干预（32.3%）和行为干预（16.1%）为主。自 2017 年起，美国设备类干预的临床试验的数量大幅增加。

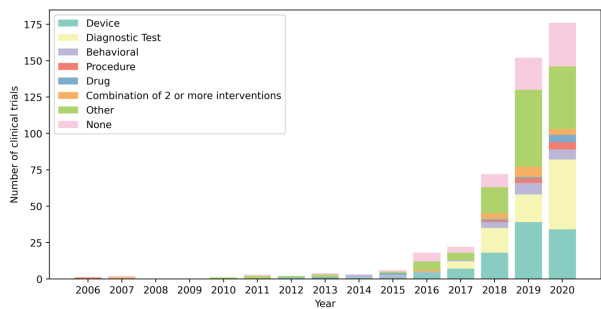


图 5-8 2006-2020 年全球人工智能相关临床试验的干预措施的时间变化趋势

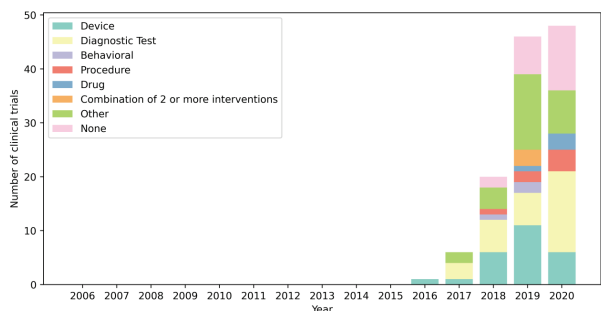


图 5-9 2006-2020 年中国人工智能相关临床试验的干预措施的时间变化趋势

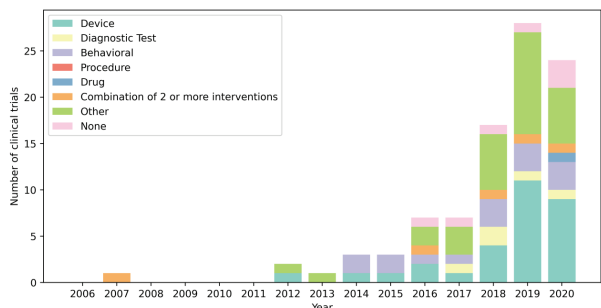


图 5-10 2006-2020 年美国人工智能相关临床试验的干预措施的时间变化趋势



### 3.6. 目标人群（疾病谱）

全球、中国以及美国人工智能临床试验的目标人群见表 52。根据临床试验研究人群的疾病词云图和具体疾病频数表，全球范围内人工智能临床试验主要关注的疾病包括：肿瘤（尤其是乳腺癌），糖尿病和心血管疾病（尤其是冠心病）。此外，抑郁症与 2020 年爆发的 COVID-19 也受到较多关注。中国人工智能临床试验主

要关注的疾病包括肿瘤，眼部疾病（尤其是青光眼与糖尿病引起的视网膜病变），心血管疾病（尤其是冠心病）和糖尿病。美国人工智能临床试验主要关注的疾病包括肿瘤，神经系统病变，心血管疾病（尤其是心衰），抑郁症和慢性疼痛。

表 5-2 全球、中国以及美国人工智能临床试验的目标人群分布

Global		China		United States	
Condition	Number of Clinical Trials	Condition	Number of Clinical Trials	Condition	Number of Clinical Trials
Breast Neoplasms	20	Diabetic Retinopathy	6	Heart Failure	6
Polyps	19	Breast Neoplasms	6	Breast Neoplasms	5
Adenoma	17	Glioma	6	Depression	4
Colonic Polyps	14	Adenoma	5	Polyps	4
Coronary Artery Disease	12	Glaucoma	5	Heart Murmurs	3
Diabetes Mellitus	12	Colonic Polyps	5	Hypertension	3
Heart Failure	11	Eye Diseases	5	Diabetes Mellitus	3
Sepsis	10	Polyps	5	Heart Diseases	3
Depression	10	Myocardial Ischemia	5	Adenoma	3
Myocardial Ischemia	10	Coronary Artery Disease	5	Colonic Neoplasms	2
COVID-19	10	Lung Neoplasms	4	Hypotension	2
Lung Neoplasms	10	Rectal Neoplasms	4	Chronic Pain	2
Stroke	9	Carcinoma	4	Diabetes Mellitus, Type 2	2
Retinal Diseases	9	Carcinoma, Hepatocellular	4	Sepsis	2
Diabetic Retinopathy	9	Colitis, Ulcerative	2	Carcinoma	2
Toxemia	9	Neoplasm Metastasis	2	Coronavirus Infections	2
Heart Diseases	8	Ulcer	2	Anxiety Disorders	2
Cardiovascular Diseases	8	Epilepsy	2	Peripheral Vascular Diseases	2
Carcinoma	8	Diabetes Mellitus	2	Prostatic Neoplasms	2
Glioma	8	Stroke	2	Toxemia	2

### 3.7. 样本量分布

全球人工智能临床试验的样本量分布见图 5-11。纳入样本量 >1000 人的临床试验为 134 例 (29.1%)，纳入样本量 >5000 人的临床试验为 63 例 (13.7%)。

其中，样本量 >5000 人的人工智能相关临床试验

的干预措施的分布见图 5-12，研究人群的疾病分布见图 5-13 和表 5-3。样本量 >5000 人的临床试验中，干预措施以诊断试验类干预 (32.9%) 为主，研究人群的疾病类型主要包括：肿瘤，毒血症，心血管疾病和 COVID-19。

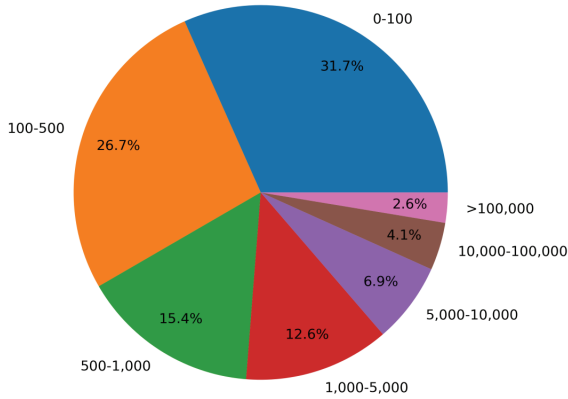


图 5-11 样本量分布

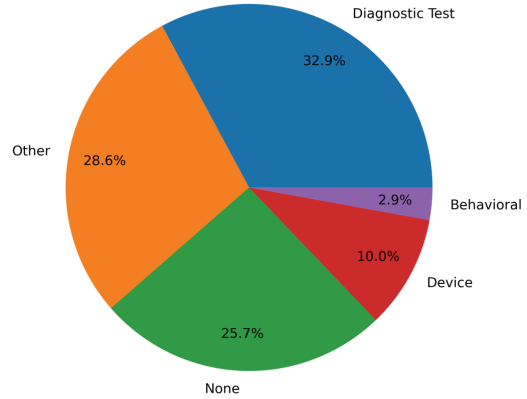


图 5-12 样本量 >5000 人的人工智能相关临床试验的干预措施

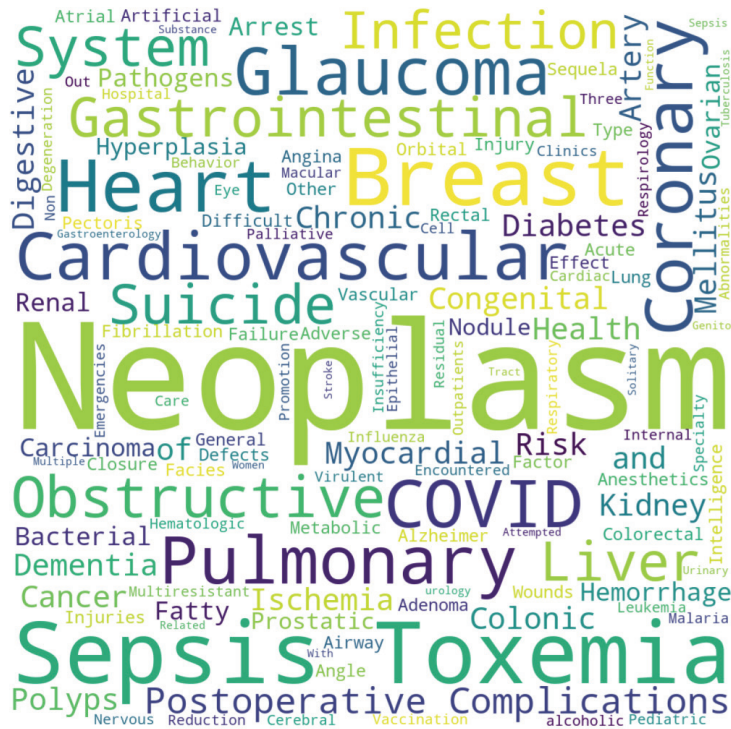


图 5-13 样本量 >5000 人的人工智能相关临床试验的研究人群词云图

表 5-3 样本量 >5000 人的人工智能相关临床试验的研究人群

Condition	Number of Clinical Trials
Sepsis	8
Toxemia	7
Breast Neoplasms	5
COVID-19	4
Glaucoma	3
Postoperative Complications	3
Cardiovascular Diseases	3
Heart Diseases	2
Coronary Artery Disease	2
Pulmonary Disease, Chronic Obstructive	2

### 3.8. 招募状态分布

全球人工智能临床试验的招募状态分布见图 5-14。其中，已有 102 例（22.1%）临床试验完成受试者招募，有 185 例（40.0%）临床试验正在进行受试者招募。已完成受试者招募的临床试验中，有 5 例（4.9%）临床试验已报告了部分或全部试验结果。

### 3.9. 临床试验结果报道

ClinicalTrials.gov 平台已报告了部分结果或全部结果的 5 项临床试验的研究设计与试验结果见表 5-4。其中，4 项临床试验的结果均支持人工智能相关的设备或行为干预对受试者的健康状况有正向积极的影响，1 项临床试验因未设计对照组，试验结果不可比。

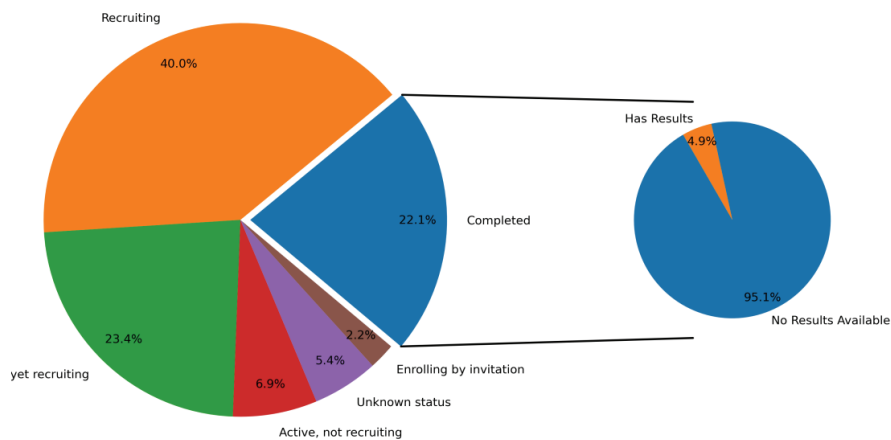


图 5-14 招募状态

Trial registration	Title	Country	Organization	Completion date	Conditions	Phase	Enrollment	Intervention	Results
NCT02176226	Artificial Intelligence in a Mobile Intervention for Depression and Anxiety (AIM)	United States	Northwestern University	2016.6	Depression, Anxiety	Not Applicable	105	Behavioral: a mobile phone application, IntelliCare	Compared with the start of treatment, the mobile intervention decreased the degree of depression severity and anxiety severity in patients with Major Depression and/or Anxiety.
NCT02801877	Artificial Intelligence in a Mobile (AIM) Intervention for Depression	United States	Northwestern University	2018.1	Depression, Anxiety	Not Applicable	301	Behavioral: IntelliCare Behavioral: Hub App with the Recommender System Behavioral: Coaching	Compared with only use of IntelliCare, IntelliCare combined with Hub Recommender and/or Coaching helped to increase patients' adherence, but neither the degree of depression severity or anxiety severity in patients with Major Depression and/or Anxiety.
NCT02988193	Piloting Healthcare Coordination in Hypertension	United States	Optima Integrated Health, University of California	2017.2	Hypertension	Not Applicable	28	Device: a clinical reasoning expert system, optima4BP Medication Management	Most medication treatment optimization recommended by the clinical reasoning expert system were adopted by the treating physicians. No adverse outcome happened both in the intervention group and the control group.
NCT03633825	Randomized Controlled Trial of an Online Machine Learning-Driven Risk Assessment and Intervention Platform for Increasing the Use of Crisis Services	United States	Harvard University	2017.9	Suicide and Depression	Not Applicable	39450	Behavioral: Brief help-seeking barrier reduction intervention	Individuals assigned to the brief Barrier Reduction Intervention (BRI) condition would report using crisis resources at higher rates than individuals in the control condition.
NCT03643692	Adaptive, Real-time, Intelligent System to Enhance Self-care of Chronic Disease	United Kingdom	Imperial College London	2019.7	Diabetes Mellitus, Type 1	Not Applicable	12	Device: Adaptive, Real-time, Intelligent System to Enhance Self-care of chronic diseases (ARISEs)	Participants with ARISES intervention achieved 64 minutes in target range (3.9-10mmol/L) without insulin dose increase.

## 第六章 主要结论

本报告综合发挥北京大学健康医疗大数据国家研究院和爱思唯尔双方各自优势，分别通过医学术语组配检索和综合性 AI 数据集自动分类的方式，提出了健康医疗人工智能科学出版物的界定方案。回顾性分析了最近 5 年健康医疗人工智能科学研究和临床试验的规模、结构和趋势。

各维度指标构建的综合指数数据分析显示：

### 1. 前沿科学技术与医学的深度融合是健康医疗人工智能发展的基础，未来将在公共卫生和临床诊疗中发挥更大作用。

全球和我国健康医疗人工智能最热的研究主题均为“计算机视觉算法与模型”，主要涉及人工智能用于疾病的影像学诊断，最近 5 年高科学影响力的论文绝大多数属于这个主题，该领域也是目前健康医疗人工智能落地应用最有前景的领域，美国斯坦福大学自 2017 年起发布的人工智能指数报告中连续 3 年将人工智能用于皮肤癌、前列腺癌、糖尿病视网膜病变影像诊断列为人类表现级里程碑式成果。其次为语义模型与推荐系统，主要涉及计算机辅助诊疗与临床决策支持系统。

前沿科学技术与医学的深度融合是智慧医疗发展的基础，涉及面向健康医疗领域特征的技术创新，包括以下方面。一是，多模态健康医疗数据的融合。多样性、多模态是健康医疗数据的特征，对于这一技术瓶颈的攻关对于盘活海量健康医疗数据的应用价值至关重要。二是面向健康医疗领域数据的人工智能技术攻关。目前人工智能技术具有依赖大数据、需要大量人工标注、可解释性差等不足，对于健康医疗领域应用尤其构成瓶颈，需要针对这些瓶颈开展技术攻关、促进在健康医疗领域的应用。三是面向健康医疗领域数据的区块链技术。区

块链技术对于打破健康医疗领域的数据孤岛、激发数据共享具有重要的价值，但需要结合健康医疗领域数据采集与应用的特点进行技术开发与模式打造。

人工智能将在公共卫生和临床诊疗中发挥更大作用，主要表现为：前沿信息技术手段助力公共卫生体系建设：运用大数据、人工智能、云计算等数字技术，在重大公共卫生危机以及重大慢病防控中发挥重要的作用；医学知识的可计算化与快速应用转化：基于前沿科学技术手段，将海量的医学知识转化为可计算和可大规模共享的医学知识，促进大数据驱动的医学知识转化和医疗质量提升；基于人工智能等前沿信息技术的临床辅助决策和个人健康管理：经过人工智能技术与优质医疗资源的深度结合，构建面向临床的疾病风险预测、影像诊断、辅助治疗等应用，以及面向个人健康管理的健康检测、智能辅助决策系统；人工智能技术辅助新药研发：人工智能技术可大大缩短药物研发时间、提高研发效率并控制研发成本、降低失败率，并能够辅助新药发现。

### 2. 中国已成为健康医疗人工智能科学研究与临床试验的最主要贡献者之一，但在学术影响力和技术转化方面仍有待提升。

2015-2019 年，全球健康医疗人工智能领域发表科学出版物最多的五个国家为：美国、中国、印度、英国和德国。中国和美国的科学出版物数量均表现为指数型增长趋势，且双方呈现出竞争态势。从平均引用角度学术影响力指标看，中国健康医疗人工智能科学出版物的影响力与全球平均水平持平，美国、英国、德国超过全球平均水平；印度低于全球平均水平。科学所产生的技术影响力可在一定程度上反映科学向技术的转化，全球范围内对健康医疗人工智能专利技术产生影响的科学出

出版物中，美国贡献了一半的份额；中国在该领域的科学产出向专利技术的转化还有提升空间。

自 2017 年起，全球人工智能相关临床试验数量急速增长，其主要增长来源为中国和美国。截至 2020 年 9 月，中国已经成为全球开展人工智能相关临床试验数量最多的国家。同时，中国从事人工智能相关临床试验的研究机构也迅速增多，全球人工智能相关临床试验数量最多的前十位研究机构中，中国的研究机构占到 6 位。由此可见，在国家近些年来一系列推进“互联网+健康医疗”的宏观政策和国内医疗行业改革需求的推动下，中国的智慧医疗行业正进入急速发展的新时期。

中国健康医疗人工智能领域主要的研究机构包括：上海交通大学、浙江大学、清华大学、中山大学、复旦大学、中国科学院大学、北京大学、四川大学等。从产学合作的角度，中国学术界-工业界合作论文占比低于美国，也低于全球平均水平；提示我国在健康医疗人工智能产学合作方面还有较大提升空间。

### 3. 健康医疗人工智能技术谱的核心是机器学习（含深度学习）和医疗机器人，疾病谱以慢病和神经系统疾病为主，传染病、罕见病等与医疗 AI 的深度结合仍有空间。

本报告对健康医疗人工智能研究领域进行了划分，包括 6 个子领域，分别为：决策规则（包括计算机启发式决策、模糊逻辑）、专家系统、知识组织系统、机器学习、自然语言处理和机器人。数据显示，机器学习（含深度学习）是最受关注和增长最快的领域；其次为医疗机器人、自然语言处理、决策规则、知识组织系统、专家系统，这 5 个子领域的发展相对平稳。

对健康医疗人工智能研究涉及的疾病谱分析显示，医疗 AI 的研究量和疾病负担总体上呈正相关，主要集中于各类癌症，以糖尿病、中风和心血管疾病等疾病为代表的慢性心脑血管疾病和以癫痫、帕金森和阿尔茨海默病为代表的神经系统疾病。传染病领域的疾病负担重，但与健康医疗 AI 的结合相对较少，这与本报告该部分数

据集截至 2019 年底有关。本次新冠病毒肺炎疫情已使我们感受到人工智能在重大公共卫生事件应急和传染病监测预警领域的应用前景。另外，母婴疾病、先天性疾病的负担较重，与医疗 AI 的结合研究也较少，与这类疾病的数据量较少有关。我国人口众多，具有患者人群优势，将人工智能用于罕见病、先天性疾病将是一大优势。

从科学-技术交叉的角度分析，机器学习和医疗机器人两个研究领域的科学出版物被专利引用较多，表现出紧密的科学-技术交叉特征。

### 4. 与科学共同体的热点研究主题相比，健康医疗人工智能研发与应用的伦理学问题成为社会媒体关注的焦点。

本报告基于多媒介提及指数和社交媒体传播指数选出的社会关注度较高的研究，在其高频关键词中明显出现了“伦理学”；但有关“伦理学”的研究，并未出现在高科学影响力的文章列表中，说明社会更加关注健康医疗人工智能研发与应用的伦理学问题。建议加强健康医疗人工智能、智慧医疗的伦理规范研究。智慧医疗涉及个人隐私数据的采集与共享、医疗数据的所有权及归属、新型医疗服务模式的应用等，需要在这些新兴领域开展医学伦理的研究，守住行业底线。同时，加强政策与法规建设。从立法立规的角度，规范和引导智慧医疗相关产品与服务的研发、应用、以及监管等，保障行业的健康稳定发展。

关于健康医疗应该要求可解释的人工智能，还是接受“黑箱”的问题，应该继续鼓励这方面的创新研究，以揭开“黑箱”的神秘面纱，但在临床应用之前坚持要求高水平的 AI 透明性可能会扼杀创新。建议需要做出临床和监管方面的配合或平衡。首先，在所有情况下，在将 AI 广泛应用于临床实践之前，都需要进行严格的质量和安全性评估。黑匣子和相对透明的算法在被 FDA 或被临床医生和患者接受之前，应表现出与现有诊疗标准相比等效或更高的性能。其次，卫生系统应确保 AI 模型在其特定患者人群中有效。缺乏模型可解释性使得局部测试尤为重要。第三，部署后应密切监测 AI 模型，特别是

如果它们随着时间的推移继续表现出适应性，应持续监测。最后，应该对医学生和临床医生进行有关AI的益处、风险和局限性的教育。医师有责任使用新技术，这些新技术可能会给医疗保健带来意义深远的改变。

## 5. 健康医疗人工智能全球临床研究仍处于早期阶段，中国以大学/医院为发起主体，侧重疾病智能诊断，企业参与仍有待加强。

全球人工智能相关的临床试验中，95%以上不能划分入传统临床试验的四个分期。传统临床试验的划分规则不适用于大部分人工智能相关的临床试验。这提示，人工智能相关临床试验的研究设计及评估与传统临床试验可能有较大差别，人工智能相关临床试验尚缺乏一个通用的、符合其应用需求的试验设计及评估规范。目前，人工智能相关的临床试验的研究设计规范、报告规范都处于起步摸索阶段。2020年Nature同时发表了两篇人工智能相关的临床试验研究的报告规范指南：SPIRIT-AI (Standard Protocol Items: Recommendations for Interventional Trials–Artificial Intelligence) 和 CONSORT-AI (Consolidated Standards of Reporting Trials–Artificial Intelligence)，两篇指南分别是基于传统临床试验的国际通用标准 SPIRIT 2013 报告指南与 CONSORT 2010 报告指南，结合文献研究与专家意见制定而成。

全球人工智能临床试验主要采取的干预措施为设备类干预与诊断试验类干预。其中，中国与美国略有差异，中国人工智能临床试验的干预措施以诊断试验类干预（人工智能辅助诊断）和设备类（人工智能诊疗设备）干预为主，而美国则以设备类（人工智能诊疗设备）干预和行为干预（如健康行为监测）为主。中美人工智能临床试验的发起机构分布也有较大差异。中国80%以上为大学/医院发起临床试验，这类机构主要接受政府研发投入，研究重点更倾向于应用范围更广的诊断试验一类的基础应用类医学研究。而美国从事人工智能相关临床试验研究的机构有40%为企业类机构，该类机构的研究重点更倾向于面向用户的行为干预类的应用研究，

如健康监测、促进治疗依从性一类的可穿戴设备等。

全球人工智能临床试验主要目标人群是患有肿瘤、重大慢性疾病或心理疾病的人群，同时2019年爆发的COVID-19也受到较多关注。中国与美国发起的人工智能临床试验的关注人群略有差异。中国临床试验对青光眼和糖尿病引起的视网膜病变给予了较多关注，这类疾病的特点是防重于治，因此早期的诊断筛查对该类疾病较为重要，这与上述提到的中国人工智能相关的临床试验的干预措施以诊断试验为主相一致。而美国人工智能相关的临床试验则对抑郁症、慢性疼痛等疾病给予了较多关注，这类疾病重视行为干预治疗，关注患者的长期依从性，这与上述提到的美国人工智能相关的临床试验的干预措施以行为干预为主相一致。

目前，大部分人工智能相关的临床试验尚处于临床研究早期阶段，未能报告相关的试验结果。此外，现有的人工智能相关的临床试验设计仍呈现一定的局限性，如约70%的临床试验的样本量小于1000人，超过半数的临床试验为观察型研究等。基于少数已报告的人工智能相关的临床试验的结果来看，人工智能相关的设备或行为干预对受试者的健康状况呈现正向、积极的影响。然而，人工智能相关的医疗应用设备在多种应用场景中的广泛的、长远的效果及影响仍有待更多临床试验结果的报告及综合分析。

## 6. 建议将循证范式引入健康医疗人工智能安全性和有效性评价，以促进其落地应用。

本报告通过多维数据展示了人工智能在健康医疗领域科学研究的热度和趋势。但是，要实现落地应用，一个至关重要的问题是：AI系统产生的信息是否值得信赖？如果我们需要依靠AI系统来辅助决策，我们就必须考虑其可靠性和有效性。基于循证医学的理念，对于智慧医疗的应用开展严谨的真实世界研究、对于安全性和有效性进行科学评价至关重要。

当前，基于深度学习的AI算法就像一个“黑匣子”：大多数机器学习模型的内在逻辑很难解释，也很难为医生提供决策建议的前因后果。由于探究因果关系和循因

治病是医疗的根本，这种不确定性会给使用 AI 系统的医生带来迟疑和困惑。研究人员也对基于 AI 的智能手机诊断程序的使用表达了担忧。正如一项针对皮肤癌智能手机诊断程序的研究中指出的，如果漏诊可疑表现，患者可能不会在疾病早期寻求专业建议，从而错过早期诊断治疗。

医学是一门不确定性的科学，“循证医学”的概念同样适用于应对 AI 在医学中应用中的不确定性。如同医学领域中其他新的干预手段，AI 系统的效力和安全性必须得到科学的评估，方能为医患所用。目前，已有倡议应用循证医学的思路来验证 AI 系统提供的医学建议；这一做法应该成为通用规则。随着技术的进步，AI 算法将变得更加稳健和成熟；应用比较效果研究评价 AI 算法在真实世界中的表现、以及评估其对患者疾病健康结局的影响是至关重要的。同样，基于 AI 的预测模型也需要在流行病学或医学研究中进行评估。

总之，健康医疗领域可能成为 AI 的应许之地，但也为 AI 技术提出了诸多挑战。为了充分发挥 AI 的潜力，医生、科研人员和 AI 科学家应当紧密合作；基于可靠的方法、遵循伦理的准则，力争在医疗实践中应用、评估和改进 AI 技术，共创健康美好未来！

**未来计划：**本部《健康医疗人工智能指数 2020》报告，侧重以已发表的科学出版物和预注册的临床试验为基础数据，未来计划将在中国医院协会健康医疗大数据应用管理专业委员会的平台上继续深入和拓展。



## Appendix: 高科学影响力、技术影响力和社媒影响力论文列表

Anthimopoulos, M., Christodoulidis, S., Ebner, L., Christe, A., and Mougiakakou, S. (2016). Lung Pattern Classification for Interstitial Lung Diseases Using a Deep Convolutional Neural Network. *IEEE Transactions on Medical Imaging* 35, 1207-1216.

Bedi, G., Carrillo, F., Cecchi, G. A., Slezak, D. F., Sigman, M., Mota, N. B., Ribeiro, S., Javitt, D. C., Copelli, M., and Corcoran, C. M. (2015). Automated analysis of free speech predicts psychosis onset in high-risk youths. *Npj Schizophrenia* 1.

Bejnordi, B. E., Veta, M., van Diest, P. J., van Ginneken, B., Karssemeijer, N., Litjens, G., van der Laak, J., and Consortium, C. (2017). Diagnostic Assessment of Deep Learning Algorithms for Detection of Lymph Node Metastases in Women With Breast Cancer. *Jama - Journal of the American Medical Association* 318, 2199-2210.

Capper, D., Jones, D. T. W., Sill, M., Hovestadt, V., Schrimpf, D., Sturm, D., Koelsche, C., Sahm, F., Chavez, L., Reuss, D. E., Kratz, A., Wefers, A. K., Huang, K., Pajtler, K. W., Schweizer, L., Stichel, D., Olar, A., Engel, N. W., Lindenberg, K., Harter, P. N., Braczynski, A. K., Plate, K. H., Dohmen, H., Garvalov, B. K., Coras, R., Holsken, A., Hewer, E., Bewerunge-Hudler, M., Schick, M., Fischer, R., Beschoner, R., Schittenhelm, J., Staszewski, O., Wani, K., Varlet, P., Pages, M., Temming, P., Lohmann, D., Selt, F., Witt, H., Milde, T., Witt, O., Aronica, E., Giangaspero, F., Rushing, E., Scheurlen, W., Geisenberger, C., Rodriguez, F. J., Becker, A., Preusser, M., Haberler, C., Bjerkvig, R., Cryan, J., Farrell, M., Deckert, M., Hench, J., Frank, S., Serrano, J., Kannan, K., Tsirogos, A., Bruck, W., Hofer, S., Brehmer, S., Seiz-Rosenhagen, M., Hanggi, D., Hans, V., Rozsnoki, S., Hansford, J. R., Kohlhof, P., Kristensen, B. W., Lechner, M., Lopes, B., Mawrin, C., Ketter, R., Kulozik, A., Khatib, Z., Heppner, F., Koch, A., Jouvett, A., Keohane, C., Muhleisen, H., Mueller, W., Poh, U., Prinz, M., Benner, A., Zaparka, M., Gottardo, N. G., Driever, P. H., Kramm, C. M.,

Muller, H. L., Rutkowski, S., von Hoff, K., Fruhwald, M. C., Gnekow, A., Fleischhack, G., Tippelt, S., Calaminus, G., Monoranu, C. M., Perry, A., Jones, C., et al. (2018). DNA methylation-based classification of central nervous system tumours. *Nature* 555, 469-+.

Char, D. S., Shah, N. H., and Magnus, D. (2018). Implementing Machine Learning in Health Care - Addressing Ethical Challenges. *New England Journal of Medicine* 378, 981-983.

Chen, R., Lu, A. J., Wang, J. J., Ma, X. H., Zhao, L., Wu, W. J., Du, Z. C., Fei, H. W., Lin, Q. W., Yu, Z. L., and Liu, H. (2019). Using machine learning to predict one-year cardiovascular events in patients with severe dilated cardiomyopathy. *European Journal of Radiology* 117, 178-183.

Christodoulou, E., Ma, J., Collins, G. S., Steyerberg, E. W., Verbakel, J. Y., and Van Calster, B. (2019). A systematic review shows no performance benefit of machine learning over logistic regression for clinical prediction models. *Journal of Clinical Epidemiology* 110, 12-22.

Colubri, A., Silver, T., Fradet, T., Retzepi, K., Fry, B., and Sabeti, P. (2016). Transforming Clinical Data into Actionable Prognosis Models: Machine-Learning Framework and Field-Deployable App to Predict Outcome of Ebola Patients. *Plos Neglected Tropical Diseases* 10, 17.

Einav, L., Finkelstein, A., Mullainathan, S., and Obermeyer, Z. (2018). Predictive modeling of US health care spending in late life. *Science* 360, 1462-+.

Esteva, A., Kuprel, B., Novoa, R. A., Ko, J., Swetter, S. M., Blau, H. M., and Thrun, S. (2017). Dermatologist-level classification of skin cancer with deep neural networks. *Nature* 542, 115-+.

Freiman, M., Nickisch, H., Prevrhal, S., Schmitt, H., Vembar, M., Maurovich-Horvat, P., Donnelly, P., and Goshen, L. (2017). Improving CCTA-based lesions' hemodynamic significance assessment by accounting for partial volume modeling in automatic coronary lumen segmentation. *Medical Physics* 44, 1040-1049.

Geis, J. R., Brady, A. P., Wu, C. C., Spencer, J., Ranschaert, E., Jaremko, J. L., Langer, S. G., Kitts, A. B., Birch, J., Shields, W. F., van Genderen, R. V., Kotter, E., Gichoya, J. W., Cook, T. S., Morgan, M. B., Tang, A., Safdar, N. M., and Kohli, M. (2019). Ethics of Artificial Intelligence in Radiology: Summary of the Joint European and North American Multisociety Statement. *Radiology* 293, 436-440.

Greenspan, H., Van Ginneken, B., and Summers, R. M. (2016). Guest Editorial Deep Learning in Medical Imaging: Overview and Future Promise of an Exciting New Technique. *IEEE Transactions on Medical Imaging* 35, 1153-1159.

Gulshan, V., Peng, L., Coram, M., Stumpe, M. C., Wu, D., Narayanaswamy, A., Venugopalan, S., Widner, K., Madams, T., Cuadros, J., Kim, R., Raman, R., Nelson, P. C., Mega, J. L., and Webster, R. (2016). Development and Validation of a Deep Learning Algorithm for Detection of Diabetic Retinopathy in Retinal Fundus Photographs. *Jama-Journal of the American Medical Association* 316, 2402-2410.

Havaei, M., Davy, A., Warde-Farley, D., Biard, A., Courville, A., Bengio, Y., Pal, C., Jodoin, P. M., and Larochelle, H. (2017). Brain tumor segmentation with Deep Neural Networks. *Medical Image Analysis* 35, 18-31.

Hazlett, H. C., Gu, H. B., Munsell, B. C., Kim, S. H., Styner, M., Wolff, J. J., Elison, J. T., Swanson, M. R., Zhu, H. T., Otteron, K. N. B., Collins, D. L., Constantino, J. N., Dager, S. R., Estes, A. M., Evans, A. C., Fonov, V. S., Gerig, G., Kostopoulos, P., McKinstry, R. C., Pandey, J., Paterson, S., Pruett, J. R., Schultz, R. T., Shaw, D. W., Zwaigenbaum, L., Piven, J., and Network, I. (2017). Early brain development in infants at high risk for autism spectrum disorder. *Nature* 542, 348-+.

He, B., Guan, Y., and Dai, R. (2019). Classifying medical relations in clinical text via convolutional neural networks. *Artificial Intelligence in Medicine* 93, 43-49.

Hengl, T., de Jesus, J. M., Heuvelink, G. B. M., Gonzalez, M. R., Kilibarda, M., Blagotic, A., Shangguan,

W., Wright, M. N., Geng, X. Y., Bauer-Marschallinger, B., Guevara, M. A., Vargas, R., MacMillan, R. A., Batjes, N. H., Leenaars, J. G. B., Ribeiro, E., Wheeler, I., Mantel, S., and Kempen, B. (2017). SoilGrids250m: Global gridded soil information based on machine learning. *Plos One* 12, 40.

Itu, L., Rapaka, S., Passerini, T., Georgescu, B., Schwemmer, C., Schoebinger, M., Flohr, T., Sharma, P., and Comaniciu, D. (2016). A machine-learning approach for computation of fractional flow reserve from coronary computed tomography. *Journal of Applied Physiology* 121, 42-52.

Ji, Z. X., Chen, Q., Niu, S. J., Leng, T., and Rubin, D. L. (2018). Beyond Retinal Layers: A Deep Voting Model for Automated Geographic Atrophy Segmentation in SD-OCT Images. *Translational Vision Science & Technology* 7, 21.

Kamnitsas, K., Ledig, C., Newcombe, V. F. J., Simpson, J. P., Kane, A. D., Menon, D. K., Rueckert, D., and Glocker, B. (2017). Efficient multi-scale 3D CNN with fully connected CRF for accurate brain lesion segmentation. *Medical Image Analysis* 36, 61-78.

Kermany, D. S., Goldbaum, M., Cai, W. J., Valentim, C. C. S., Liang, H. Y., Baxter, S. L., McKeown, A., Yang, G., Wu, X. K., Yan, F. B., Dong, J., Prasadha, M. K., Pei, J., Ting, M., Zhu, J., Li, C., Hewett, S., Dong, J. S., Ziyar, I., Shi, A., Zhang, R. Z., Zheng, L. H., Hou, R., Shi, W., Fu, X., Duan, Y. O., Huu, V. A. N., Wen, C., Zhang, E. D., Zhang, C. L., Li, O. L., Wang, X. B., Singer, M. A., Sun, X. D., Xu, J., Tafreshi, A., Lewis, M. A., Xia, H. M., and Zhang, K. (2018). Identifying Medical Diagnoses and Treatable Diseases by Image-Based Deep Learning. *Cell* 172, 1122-+.

Libbrecht, M. W., and Noble, W. S. (2015). Machine learning applications in genetics and genomics. *Nature Reviews Genetics* 16, 321-332.

Litjens, G., Sanchez, C. I., Timofeeva, N., Hermsen, M., Nagtegaal, I., Kovacs, I., Hulsbergen-van de Kaa, C., Bult, P., van Ginneken, B., and van der Laak, J. (2016). Deep learning as a tool for increased accuracy and efficiency of histopathological diagnosis. *Scientific Reports* 6, 11.

- Liu, Y. Z., Lin, W. J., Xu, P. F., Zhang, D. D., and Luo, Y. J. (2015). Neural Basis of Disgust Perception in Racial Prejudice. *Human Brain Mapping* 36, 5275-5286.
- Maniruzzaman, M., Rahman, M. J., Al-MehediHasan, M., Suri, H. S., Abedin, M. M., El-Baz, A., and Suri, J. S. (2018). Accurate Diabetes Risk Stratification Using Machine Learning: Role of Missing Value and Outliers. *Journal of Medical Systems* 42, 17.
- Obermeyer, Z., and Emanuel, E. J. (2016). Predicting the Future - Big Data, Machine Learning, and Clinical Medicine. *New England Journal of Medicine* 375, 1216-1219.
- Perfecto-Avalos, Y., Garcia-Gonzalez, A., Hernandez-Reynoso, A., Sanchez-Ante, G., Ortiz-Hidalgo, C., Scott, S. P., Fuentes-Aguilar, R. Q., Diaz-Dominguez, R., Leon-Martinez, G., Velasco-Vales, V., Cardenas-Escudero, M. A., Hernandez-Hernandez, J. A., Santos, A., Borbolla-Escoboza, J. R., and Villela, L. (2019). Discriminant analysis and machine learning approach for evaluating and improving the performance of immunohistochemical algorithms for COO classification of DLBCL. *Journal of Translational Medicine* 17, 12.
- Setio, A. A. A., Ciompi, F., Litjens, G., Gerke, P., Jacobs, C., van Riel, S. J., Wille, M. M. W., Naqibullah, M., Sanchez, C. I., and van Ginneken, B. (2016). Pulmonary Nodule Detection in CT Images: False Positive Reduction Using Multi-View Convolutional Networks. *Ieee Transactions on Medical Imaging* 35, 1160-1169.
- Shen, D. G., Wu, G. R., and Suk, H. I. (2017). Deep Learning in Medical Image Analysis. In "Annual Review of Biomedical Engineering, Vol 19" (M. L. Yarmush, ed.), Vol. 19, pp. 221-248. Annual Reviews, Palo Alto.
- Shin, H. C., Roth, H. R., Gao, M. C., Lu, L., Xu, Z. Y., Nogues, I., Yao, J. H., Mollura, D., and Summers, R. M. (2016). Deep Convolutional Neural Networks for Computer-Aided Detection: CNN Architectures, Dataset Characteristics and Transfer Learning. *Ieee Transactions on Medical Imaging* 35, 1285-1298.
- Sirinukunwattana, K., Raza, S. E. A., Tsang, Y. W., Snead, D. R. J., Cree, I. A., and Rajpoot, N. M. (2016). Locality Sensitive Deep Learning for Detection and Classification of Nuclei in Routine Colon Cancer Histology Images. *Ieee Transactions on Medical Imaging* 35, 1196-1206.
- Springer, S., Masica, D. L., Dal Molin, M., Douville, C., Thoburn, C. J., Afsari, B., Li, L., Cohen, J. D., Thompson, E., Allen, P. J., Klimstra, D. S., Schattner, M. A., Schmidt, C. M., Yip-Schneider, M., Simpson, R. E., Fernandez-Del Castillo, C., Mino-Kenudson, M., Brugge, W., Brand, R. E., Singhi, A. D., Scarpa, A., Lawlor, R., Salvia, R., Zamboni, G., Hong, S. M., Hwang, D. W., Jang, J. Y., Kwon, W., Swan, N., Geoghegan, J., Falconi, M., Crippa, S., Doglioni, C., Paulino, J., Schulick, R. D., Edil, B. H., Park, W., Yachida, S., Hijioka, S., van Hooft, J., He, J., Weiss, M. J., Burkhart, R., Makary, M., Canto, M. I., Goggins, M. G., Ptak, J., Dobbyn, L., Schaefer, J., Sillman, N., Popoli, M., Klein, A. P., Tomasetti, C., Karchin, R., Papadopoulos, N., Kinzler, K. W., Vogelstein, B., Wolfgang, C. L., Hruban, R. H., and Lennon, A. M. (2019). A multimodality test to guide the management of patients with a pancreatic cyst. *Science Translational Medicine* 11, 14.
- Tajbakhsh, N., Shin, J. Y., Gurudu, S. R., Hurst, R. T., Kendall, C. B., Gotway, M. B., and Liang, J. M. (2016). Convolutional Neural Networks for Medical Image Analysis: Full Training or Fine Tuning? *Ieee Transactions on Medical Imaging* 35, 1299-1312.
- Thaha, M. M., Kumar, K. P. M., Murugan, B. S., Dhanasekeran, S., Vijayakarhick, P., and Selvi, A. S. (2019). Brain Tumor Segmentation Using Convolutional Neural Networks in MRI Images. *Journal of Medical Systems* 43, 10.
- van Griethuysen, J. J. M., Fedorov, A., Parmar, C., Hosny, A., Aucoin, N., Narayan, V., Beets-Tan, R. G. H., Fillion-Robin, J. C., Pieper, S., and Aerts, H. (2017). Computational Radiomics System to Decode the Radiographic Phenotype. *Cancer Research* 77, E104-E107.

Wang, X. G., Yang, W., Weinreb, J., Han, J., Li, Q. B., Kong, X. C., Yan, Y. L., Ke, Z., Luo, B., Liu, T., and Wang, L. (2017). Searching for prostate cancer by fully automated magnetic resonance imaging classification: deep learning versus non-deep learning. *Scientific Reports* 7, 8.

Wang, Z. F., Lei, X. J., and Wu, F. X. (2019). Identifying Cancer-Specific circRNA-RBP Binding Sites Based on Deep Learning. *Molecules* 24, 13.

Xing, F. Y., Xie, Y. P., and Yang, L. (2016). An Automatic Learning-Based Framework for Robust Nucleus Segmentation. *Ieee Transactions on Medical Imaging* 35, 550-566.

Xiong, H. Y., Alipanahi, B., Lee, L. J., Bretschneider, H., Merico, D., Yuen, R. K. C., Hua, Y. M., Gueroussov, S., Najafabadi, H. S., Hughes, T. R., Morris, Q., Barash, Y., Krainer, A. R., Jovic, N., Scherer, S. W., Blencowe, B. J., and Frey, B. J. (2015). The human splicing code reveals new insights into the genetic determinants of disease. *Science* 347, 9.

Xu, J. P., Zhang, J. Q., Zhang, J. L., Wang, Y., Zhang, Y. L., Wang, J., Li, G. L., Hu, Q. M., and Zhang, Y. C. (2017). Abnormalities in Structural Covariance of Cortical Gyri-fication in Parkinson's Disease. *Frontiers in Neuroanatomy* 11, 9.

Yao, S. X., Becker, B., Zhao, W. H., Zhao, Z. Y., Kou, J., Ma, X. L., Geng, Y. Y., Ren, P., and Kendrick, K. M. (2018). Oxytocin Modulates Attention Switching Between Interoceptive Signals and External Social Cues. *Neuropsychopharmacology* 43, 294-301.

Zhang, F. L., Wu, C. H., Jia, C. X., Gao, K., Wang, J. P., Zhao, H. H., Wang, W., and Chen, J. X. (2019). Artificial intelligence based discovery of the association between depression and chronic fatigue syndrome. *Journal of Affective Disorders* 250, 380-390.

Zhu, J. J., Hou, W. H., Xu, Y., Ji, F., Wang, G. W., Chen, C., Lin, C. G., Lin, X. D., Li, J., Zhuo, C. J., and Shao, M. J. (2019). Antipsychotic drugs and sudden cardiac death: A literature review of the challenges in the prediction, management, and future steps. *Psychiatry Research* 281, 7.



健康医疗 HEALTH  
AI INDEX REPORT  
人工智能指数报告 20



■ 友情链接：

中国医院协会：<http://www.cha.org.cn/>

北京大学健康医疗大数据国家研究院：<http://www.nihds.pku.edu.cn/>

北京大学信息技术高等研究院（浙江）：<http://aiit.org.cn/>

北京大学人工智能研究院智慧公众健康研究中心：<http://www.ai.pku.edu.cn/yjzx1/zhgzjkyjzx.htm>

爱思唯尔公司：<https://www.elsevier.com/>

爱思唯尔《人工智能：知识的创造、转移与应用》报告：<https://www.elsevier.com/research-intelligence/resource-library/ai-report>