

DOI: 10.13376/j.cblls/2025157

文章编号: 1004-0374(2025)12-1615-09



李华婷, 国家优青, 研究员, 主治医师, 上海市糖尿病研究所副所长, 上海市糖尿病重点实验室副主任。致力于内分泌代谢病的临床与科研, 发表SCI论著 50 余篇, 代表性论文发表于 *Nature Medicine*、*Cell Metabolism*、*Nature Metabolism*、*Science Translational Medicine* 等; 入选国家人力资源和社会保障部香江学者计划等人才计划; 获得国家重点研发计划、国家自然科学基金重大研究计划等课题资助; 获上海市卫生健康行业青年五四奖章, 作为主要完成人获国家自然科学基金二等奖、上海市科技进步奖一等奖等。

循证视角下多模态人工智能在慢性病中的应用

黄惠洁^{1,2}, 于淑洁^{1,2}, 管洲榆^{1,2}, 吴倩^{1,2}, 盛斌³, 贾伟平^{1,2}, 李华婷^{1,2*}

(1 上海交通大学医学院附属第六人民医院内分泌代谢科, 上海市糖尿病研究所, 上海市糖尿病重点实验室, 上海市糖尿病临床医学中心, 上海 200233; 2 上海交通大学医学院, 上海 200025; 3 上海交通大学电子信息与电气工程学院计算机系, 上海 200240)

摘要: 慢性病 (如心血管疾病、糖尿病、癌症及慢性呼吸系统疾病) 已成为全球公共卫生的重大挑战, 严重威胁人类健康并造成沉重经济负担。传统筛查和管理方式存在识别效率低、依从性不足及资源分布不均等局限。随着医学人工智能的发展, 多模态人工智能模型通过整合临床资料、影像学检查、组学数据、生理监测及生活方式信息, 为慢性病的早期筛查、风险预测、辅助诊断及个体化干预提供了新路径。尽管如此, 当前证据循证强度有限, 未来需通过大规模、多中心、前瞻性研究和随机对照试验提升证据等级, 解决模型推广性、可解释性及临床转化等挑战。总体而言, 多模态人工智能在循证医学框架下有望成为慢性病防控的重要工具, 推动精准医学与公共健康的协同发展。

关键词: 循证医学; 多模态模型; 慢性病; 人工智能; 筛查与防治

中图分类号: R319; TP18 **文献标志码:** A

Application of multimodal artificial intelligence models in chronic disease management: an evidence-based perspective

HUANG Hui-Jie^{1,2}, YU Shu-Jie^{1,2}, GUAN Zhou-Yu^{1,2}, WU Qian^{1,2}, SHENG Bin³, JIA Wei-Ping^{1,2}, LI Hua-Ting^{1,2*}

(1 Shanghai Clinical Center for Diabetes, Shanghai Key Laboratory of Diabetes Mellitus, Shanghai Diabetes Institute, Department of Endocrinology and Metabolism, Shanghai Sixth People's Hospital, Shanghai Jiao Tong University School of Medicine, Shanghai 200233, China; 2 Shanghai Jiao Tong University School of Medicine, Shanghai 200025, China; 3 Department of Computer Science and Engineering, Shanghai Jiao Tong University, Shanghai 200240, China)

收稿日期: 2025-09-02; 修回日期: 2025-10-15

基金项目: 深圳市医学研究专项资金项目(E250200610); 深圳市医学研究专项资金(C2406002); 四大慢病重大专项(2024ZD0523200); 国家重点研发计划(2022YFA1004804)

*通信作者: E-mail: huarting99@sjtu.edu.cn

Abstract: Chronic diseases, including cardiovascular disease, diabetes, cancer, and chronic respiratory disorders, have become major global public health challenges, causing substantial morbidity, mortality, and economic burden. Traditional approaches to screening and management are limited by inefficiency, poor adherence, and unequal distribution of healthcare resources. With the rapid advancement of artificial intelligence, multimodal models integrating clinical data, medical imaging, omics profiles, physiological monitoring, and lifestyle information have provided new opportunities for chronic disease prevention and control. Nevertheless, the current evidence remains limited in strength, and future efforts should focus on large-scale, multi-center, prospective studies and randomized controlled trials to enhance the evidence quality and address challenges such as model generalizability, interpretability, and clinical applicability. Overall, under the framework of evidence-based medicine, multimodal AI holds great promise as a transformative tool for chronic disease prevention and management, advancing both precision medicine and public health.

Key words: evidence-based medicine; multimodal model; chronic diseases; artificial intelligence; screening and management

随着经济的快速发展和居民生活方式的改变,慢性病已成为全球范围内严重威胁人类健康的重要公共卫生问题,对社会造成沉重负担。慢性病主要包括心血管疾病 (cardiovascular disease, CVD)、癌症、糖尿病 (diabetes, DM) 及慢性呼吸系统疾病,每年导致约 4 100 万人死亡,占全球总死亡人数的近四分之三^[1]。预计到 2030 年,慢性病直接医疗费用将高达 3 018 亿美元,死亡人数可能上升至 5 600 万^[2]。世界卫生组织将 30~70 岁人群的死亡定义为“过早死亡”,并提出到 2030 年将慢性病导致的过早死亡率降低三分之一的全球目标^[3],强调在预防、筛查、治疗和长期管理方面加大投入。

然而,慢性病早期通常缺乏典型临床表现,加之患者基数庞大,使得全面筛查既不经济也不可行,因而漏诊率较高,导致疾病进入中晚期时方才确诊,增加了治疗难度并影响预后^[4]。目前亟需建立早期、便捷且经济可行的慢性病筛查和预警方法。但目前的慢性病防治体系面临多重挑战^[5]:其一,大量健康档案、实时监测数据(如血糖、血压)、基因组学及多组学数据亟待整合,信息碎片化严重,增加了个体化治疗与整体管理的难度;其二,慢性病管理高度依赖患者长期自我管理,但约 50% 的患者无法坚持药物治疗、饮食控制和运动干预,导致疾病进展与并发症发生;其三,医疗资源分配不均和基层专业医生不足,尤其在低收入和中等收入国家更为突出,进一步限制了慢性病管理效果。

在此背景下,医学人工智能 (artificial intelligence, AI) 的快速发展为慢性病防治提供了新的契机^[6]。深度学习 (deep learning, DL) 等方法能够从图像、文本和生理信号中自动提取潜在模式与表征,展现

出在疾病风险预测和辅助诊断中的独特优势,从而推动数字技术与医疗的深度融合^[7]。特别是涵盖电子病历、实验室与影像学检查、生物样本分析、患者自我报告、可穿戴设备监测乃至社交媒体健康信息等多源多模态数据的应用,为慢性病研究和临床实践提供了更为全面的健康画像。这一数据整合有助于建立更加精准的慢性病模型,实现疾病的早期筛查、风险预测、辅助诊断及个体化干预。

医学 AI 落地于临床实践需以循证医学作为核心支撑。循证医学强调将高质量研究证据、临床专业经验和患者价值观三者结合^[8],以确保多模态模型开发的科学性和临床实用性,并为模型评价、临床转化及决策优化提供规范框架^[9]。根据牛津循证医学中心 (Oxford Centre for Evidence-Based Medicine, OCEBM) 2025 v2.1 的证据分级体系^[10],多模态 AI 模型开发的早期研究(如单中心回顾性分析)多属于 OCEBM 4-5 级证据;若具备多中心大样本及外部验证,可提升至 OCEBM 2b-3b 级证据^[11];而若能在前瞻性随机对照试验 (randomized controlled trial, RCT) 中证实其对诊断决策或临床结局的改善,则可达到 OCEBM 1a-1b 级证据^[12],并为纳入临床指南提供坚实依据。对于多模态 AI 模型而言,外部验证、跨人群的可推广性以及长期随访终点的验证尤为关键^[13]。基于此,本文将从循证医学的视角出发,系统综述多模态 AI 模型在慢性病筛查、诊断与防治中的应用与发展(图 1)。

1 心血管疾病

CVD 是全球首位死亡原因,约占全部死亡的三分之一,严重威胁全球人类健康与生命。据统计,

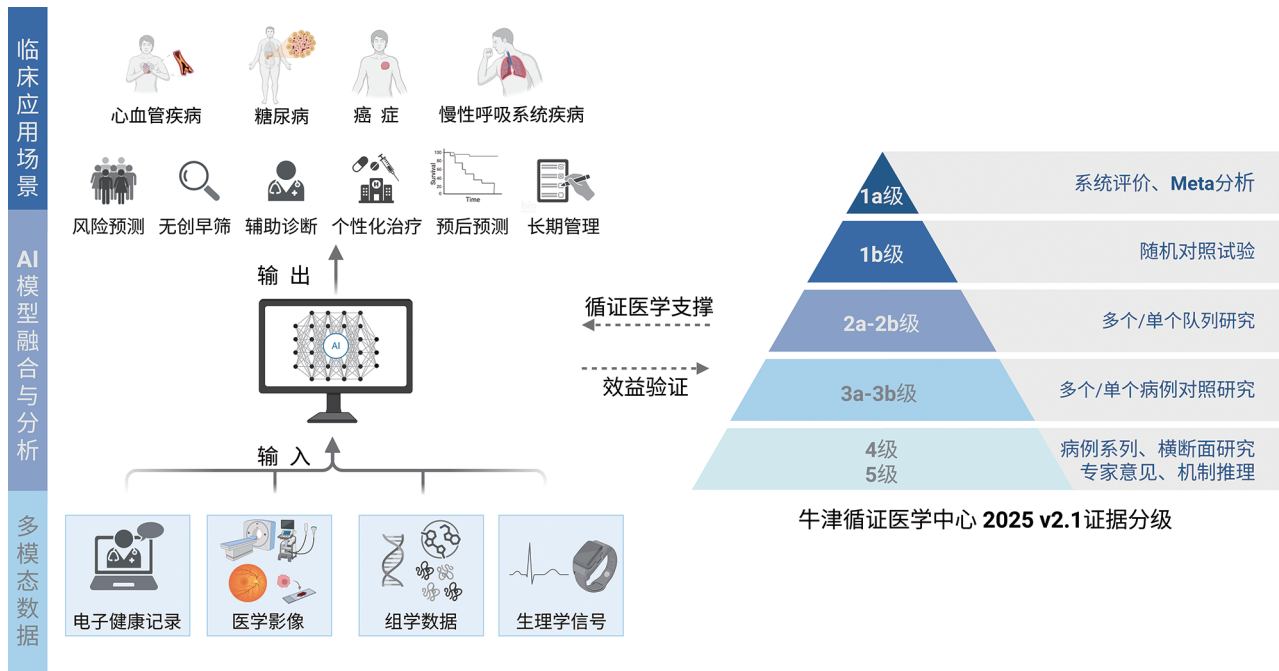


图1 多模态AI模型临床转化路径

2021 年全球 CVD 患病人数估计为 6.12 亿人，且负担在低收入和中等收入国家尤为突出^[14]。尽管防治水平不断提升，但由于复杂的病理生理学机制涉及遗传、生活方式和环境因素的交互作用，CVD 的患病人数仍居高不下。临床上，CVD 患者就医时大多数已进入中晚期，此时的治疗更复杂、预后较差。因此，早期精准预测与个性化干预至关重要，是改善疾病预后和生活质量、降低死亡率的关键所在^[15]。

CVD 的传统诊疗高度依赖影像学检查，如心脏彩超、冠状动脉计算机断层扫描 (computed tomography, CT)、冠脉造影、心脏磁共振 (magnetic resonance imaging, MRI) 等，不仅费用昂贵、耗时较长、部分方法具有侵入性，还对操作者的专业能力要求高，使得在医疗资源有限的地区难以开展大规模和长期的随访评估^[16]。为此，大规模人群研究明确了 CVD 的常见可控危险因素，包括高血压、不良饮食、缺乏运动、DM 及血糖管理不佳、吸烟饮酒以及睡眠不足等^[17]。基于这些临床信息，已开发出多种风险预测工具，例如 Framingham 模型 (美国)、SCORE 模型 (欧洲) 及 QRISK 模型 (英国) 等^[18-20]。这些评估方法显著提高了 CVD 高危人群的识别率，并指导实施包括戒烟、控制血压、服用他汀类药物等在内的干预措施。然而，这些模型在不同人群中的表现并不一致。Framingham 风险模

型在英国及其他欧洲人群中常常高估心血管风险，而 SCORE 模型则在部分群体中低估风险^[21]。这意味着部分患者即便未接受干预也未发生 CVD，而另一些评分较低的患者仍发生 CVD。因此，从循证角度来看，当前的风险预测方法虽已改善人群水平的 CVD 防控，但其在个体精准预测方面仍有不足。未来仍需开发新型、无创、便捷且高效的风险预警与诊疗工具。

AI 多模态模型凭借对大规模、多样化及实时数据的整合与分析能力，正逐渐成为改善 CVD 风险预测与临床管理的重要工具，并已积累初步循证证据。多模态模型将眼底照片和临床危险因素相结合用于动脉粥样硬化风险分层，在韩国三星医疗中心与 UK Biobank 的外部验证中表现稳健，受试者工作特征曲线下面积 (area under the receiver operating characteristic curve, AUC) 达 0.781~0.872，为无创、可扩展的群体早筛提供了证据^[22]。最新提出的 MAARS 框架融合电子健康记录 (electronic health record, EHR)、超声心动图报告和心脏 MRI 影像，在预测致命室性心律失常时，内外部验证 AUC 分别为 0.89 和 0.81，显示了其在高危人群精准分层中的循证价值^[23]。在远程监测与无创筛查方面，SleepCVD-Net 融合心电图 (electrocardiography, ECG)、血氧饱和度和呼吸气流等多源生理信号，在心绞痛与心衰分类中的准确率高达 97.6%，展现了可穿戴

监测在持续健康管理中的应用潜力^[24]。进一步地, LVH-fusion 模型将 ECG 与超声心动图时间序列联合分析, 不仅提升了左室肥厚的检测精度, 还通过可解释机制揭示了跨模态特征对模型决策的贡献, 为复杂心脏疾病的智能诊断提供了新思路^[25]。

总的来说, 现有循证证据表明, 多模态 AI 模型已在 CVD 的风险预测、早期筛查和辅助诊断方面展现出优于传统方法的潜力, 并为构建更精准的无创风险评估奠定了基础。

2 糖尿病

DM 已成为全球性的公共卫生挑战。最新流行病学数据显示, 截至 2024 年, 全球 20~79 岁成年人中约有 5.89 亿人患有 DM, 预计到 2050 年将增加至 8.53 亿人^[26]。其中, 约四分之三的 DM 患者生活于中低收入国家, 凸显出 DM 地域覆盖范围广、分布不均衡的特征。在如此庞大且不均衡的流行态势下, 现有健康管理模式在实践中面临明显不足, 难以支撑循证医学所要求的有效防控^[27]。DM 是一种由遗传、环境及生活方式等多因素共同作用的复杂代谢性疾病。在疾病进展过程中, 并发症的综合管理始终是核心挑战。由于 DM 起病隐匿, 部分患者确诊时已伴有糖尿病视网膜病变 (diabetic retinopathy, DR)、糖尿病肾病 (diabetic kidney disease, DKD) 和周围神经病变等并发症, 延迟诊断和治疗不仅降低患者生活质量, 还显著增加失明、肾功能衰竭和过早死亡等不良预后风险^[28, 29]。并发症的发生往往与病程延长及血糖控制不良等因素密切相关, 因此, 提高知晓率和早期诊断率, 识别高危人群并进行前移干预, 是循证医学强调的核心目标。然而, 传统 DM 筛查主要依赖病史、人体测量和实验室指标。这些单一模态方法的预测效能有限, 难以覆盖 DM 的复杂病因, 也难以将疾病识别提前至代谢损伤之前^[30, 31]。由此可见, 迫切需要开发更高效、可推广且兼顾成本效益的手段, 而多模态 AI 模型的出现为 DM 预测、筛查、诊断和管理提供了新路径。

在众多多模态数据中, 眼底图像成为无创评估 DM 全身多系统健康状况与并发症风险的重要“窗口”, 也为多模态 AI 模型提供了循证医学意义上的理想数据源。前瞻性队列研究表明, 融合眼底图像和人口学信息的 DL 模型可预测未来 DM 风险 (AUC $\approx$ 0.82), 并实现有效风险分层^[32]。本团队构建的 DeepDR Plus 在超 20 万例多国、多种族患者的纵向队列中, 仅凭基线眼底图像即可预测 5 年内

DR 进展, 一致性指数 (concordance index, C-index) 大于 0.77, 并支持将低危人群的筛查间隔由 1 年延长至 3 年而不增加漏诊风险^[33]。除筛查外, 眼底图像驱动的 AI 模型还可揭示多系统并发症的诊断价值。将 DL 对眼底图像的分析结果与肾活检病理学这一“金标准”直接进行对照, 发现 AI 模型能够较准确地反映 DKD 病理特征^[34]。另外, 本团队还拓展了跨系统并发症的研究, 基于眼底图像开发的 DL 模型能够在多中心人群中检测隐匿性脑梗死, 识别性能 AUC 达 0.92, 并能有效预测未来卒中事件 (C-index: 0.74~0.76), 显著优于传统风险评分^[35]。这些循证证据表明, 眼底影像不仅能用于 DM 本身的识别, 也能作为全身并发症诊断与预测的重要切入点。进一步地, 本团队研发的 DeepDR-LLM 系统结合视觉模型与大语言模型, 能够在全球多中心队列中提供 DR 辅助诊断及个体化 DM 管理建议。该系统在中国基层医疗的前瞻性真实世界验证中显著提高了 DR 筛查与管理决策水平, 为基层 DM 防控提供了高质量循证证据^[36]。

除了眼底图像, 多模态 AI 也拓展至其他数据类型。如结合 ECG 与传统临床危险因素的多模态神经网络模型——ECG-DiaNet^[37], 显著提升了 DM 风险预测的准确性 (AUC=0.845), 显示出在早期识别与精准分层方面的优势。整合基因组学与代谢组学信息的多模态机器学习模型在一项基于 10 年随访的大规模前瞻性队列研究中表现突出, 预测 DM 发生的 AUC 达到 0.88, 较单一基因组或代谢组模型均有显著提升, 显示出多源组学数据在疾病风险预测中的互补价值^[38]。另一方面, 影像学模态的融合也逐渐拓展至糖尿病足的严重程度评估与愈合预测, 例如结合足部热成像与临床图像训练的多模态 DL 模型, 在区分溃疡分级及预测愈合结局方面较单一图像模式显著提高, 提示了其在并发症管理与个体化干预中的潜在临床应用价值^[39, 40]。

现阶段证据显示, DM 领域的多模态 AI 研究正在由单模态向多模态与跨系统整合演进, 其中基于眼底图像的模型因其循证证据相对集中, 为并发症筛查、诊断及个体化管理的拓展提供了重要范例。

3 癌症

据统计, 在非传染性疾病死亡中, 约有 20% 的死因是癌症, 给当今社会带来沉重的经济压力和公共卫生挑战^[41]。癌症的早期筛查与风险预测、精准诊断与分型、个性化治疗、疗效监测与预后评估

一直是医学研究领域极为重要的目标。现代医学迭代更新, 数据积累呈现出飞速增长的态势。EHR、影像报告以及基因组数据等多模态信息全面涵盖了癌症患者的诊疗过程。然而, 这些数据分散存储于不同系统, 且具有异构性, 这给数据的整合与分析带来了巨大的挑战。AI 的机器学习技术、DL 技术和自然语言处理技术可获取这些自由文本的信息, 推动多模态模型助力癌症精准医学的进一步发展。

在癌症领域, 多模态 AI 模型逐渐展现出从疾病筛查、预后预测与风险分层到个体化治疗的循证价值。在高危人群识别方面, Mirai 模型利用钼靶影像结合患者风险因素预测未来 1~5 年乳腺癌风险 ($C\text{-index}=0.76$), 并在多国多中心外部验证中依旧保持稳健性能, 后续通过强化学习将风险分数转化为个体化筛查间隔, 实现了风险预测与筛查策略的闭环^[42-44]。另外, M3FM 模型整合 CT 影像与临床数据, 在肺癌风险预测中相较单模态提升 20%, 体现了多任务融合在高危人群识别中的潜力^[45]。在预后预测方面, 2022 年发表的 PORPOISE 多模态模型采用弱监督的 DL 框架, 将全景病理切片与多组学特征融合, 在涵盖 5 720 名患者、14 种癌症类型的研究中, 显著提高了生存风险分层的准确性, $C\text{-index}$ 提升幅度约为 0.05~0.08, 并通过注意力机制揭示了形态学与分子特征间的联系^[46]。在个体化治疗环节, 整合放射学、病理学与基因组数据的多模态 AI 模型可有效预测非小细胞肺癌患者对程序性死亡配体 1 (programmed death ligand-1, PD-L1) 抑制剂的疗效 ($AUC=0.80$)^[47]。纳入了来自四家医院共 998 名高级别浆液性卵巢癌患者的诊断研究显示, 融合 MRI、临床及病理信息的多模态模型在预测铂类化疗敏感性方面优于单一模态 (外部验证 AUC 达 0.807), 提示了其在个体化治疗中的潜力^[48]。

综上所述, 多模态 AI 不仅提升了恶性肿瘤早期筛查与精准诊断的性能, 还在预后评估与个体化治疗方面展现出独特价值, 正逐渐成为推动癌症精准诊疗的重要工具。

4 慢性阻塞性肺疾病

慢性阻塞性肺疾病 (chronic obstructive pulmonary disease, COPD) 是最常见的慢性呼吸系统疾病之一。全球患者超过 2.5 亿例, 预计到 2030 年将成为继 CVD 和癌症之后的第三大死亡原因, 对公共卫生和医疗系统构成沉重负担^[49, 50]。COPD 每年造成的直接医疗成本和生产力损失高达数十亿美元^[51]。这

种流行病学与经济学数据凸显出早期识别、精准分期和有效管理的迫切性。当前, COPD 筛诊依赖于需专业人员操作及判读的肺功能和影像学检查, 限制了早期识别的普及性。这一困境在中国的一项大型流行病学调查中也得到了印证^[52]。约 60.2% 的 COPD 患者缺乏典型自述症状, 且仅 12% 的患者在调查前接受过肺功能检查。这一结果凸显了 COPD 早期诊断的难度及公众认知的缺乏。

在此背景下, AI 技术被视为改善 COPD 早期识别的重要契机。与传统方法相比, AI 能够直接利用大规模临床与影像数据实现自动化分析, 从而提升筛查效率和可及性。在此基础上, 多模态 AI 通过进一步融合临床症状、影像学特征、生活环境和个人行为数据等多源信息, 相较于单一模态模型能更全面地刻画 COPD 的复杂性, 并在疾病诊治中展现出更高的循证潜力。

近期, 面向临床电子病历的多模态 DL 框架 COPD-MMDDxNet 在缺乏肺功能检测的环境下展现出良好的诊断性能^[53]。该模型整合了影像学报告、病史文本及血气分析和血液学指标等结构化数据, 有效弥补了基层医院和欠发达地区因肺功能检测缺乏而导致的诊断不足。Kumar 等^[54]提出的融合胸部 CT 图像与呼吸音或咳嗽声音样本的多模态 AI 模型在 COPD 早期诊断和分型中表现优异, 在准确率、灵敏度和特异性方面均显著优于单一模态, 提示影像与生理信号的互补信息能够提升早期识别的循证价值。在急性加重期预测领域, Atzeni 等^[55]建立了融合个人空气质量监测、EHR 和生活方式数据的多模态框架, 能够实现短期加重风险的有效预测, 并通过可解释性分析提升模型临床可用性。这为 COPD 的远程管理和个体化干预提供了循证支持。

综合来看, 多模态模型在 COPD 的早期识别、辅助诊断及急性加重监测中展现出优势, 并凸显了其在真实世界环境下提升筛查效率和疾病管理质量的潜力。

5 不足与挑战

尽管多模态 AI 模型在慢性病筛查、诊断与治疗中展现了广阔前景, 但从循证医学角度评估, 其应用仍面临诸多局限。

第一, 证据强度不足。当前大部分多模态 AI 模型虽已实现多中心和外部验证 (表 1), 但整体仍集中在 OCEBM 2b-3b 级^[56], 缺乏以 RCT 或长期

表1 循证视角下多模态AI模型在四大慢性病筛防治中的代表性研究及证据等级

疾病	模型	数据模态	应用	研究类型	证据等级
CVD	眼底图像结合危险因素模型 ^[22]	眼底图像和危险因素	动脉粥样硬化风险预测	队列研究 (含外部验证)	2b-3b
	MAARS ^[23]	EHR、超声心动图和心脏MRI	室性心律失常预测	单队列研究 (外部验证)	2b
	LVH-fusion ^[25]	ECG和超声心动图	左室肥厚诊断	单队列研究	2b-3b
DM	DeepDR Plus ^[33]	眼底图像和临床数据	DR进展预测	多国前瞻性队列研究(临床结局)	2b
	DeepDR-LLM ^[36]	眼底图像和大语言模型	DR筛查与DM管理	多中心前瞻性研究(流程结局)	2c
	ECG-DiaNet ^[37]	心电图和临床危险因素	DM风险预测	单队列研究	2b-3b
癌症	多组学风险模型 ^[38]	基因组学和代谢组学	DM发生预测	前瞻性队列研究	2a-2b
	Mirai ^[42-44]	钼靶影像和危险因素	乳腺癌风险预测	多中心前瞻性队列研究(临床结局)	2b
	M3FM ^[45]	胸部CT和临床数据	肺癌风险预测	单队列研究	2b-3b
	PORPOISE ^[46]	全景病理切片和多组学	多癌种预后分层	多队列回顾性研究 (外部验证)	2b-3b
	多模态免疫疗效预测 ^[47]	影像、病理和基因组数据	非小细胞肺癌PD-L1抑制剂疗效预测	单队列研究	2b-3b
	COPD-MMDDxNet ^[53]	EHR、影像报告和检验指标	基层辅助诊断	多中心队列研究	2b-3b
COPD	影像和声音模型 ^[54]	胸部CT和呼吸音/咳嗽声	早期诊断与分型	单队列研究	2b-3b
	环境-生活方式框架 ^[55]	空气质量监测、EHR和生活方式数据	急性加重预测	前瞻性队列研究 (临床结局)	2b

注：证据等级依据OCEBM 2025 v2.1分级判定。以患者临床结局(如乳腺癌发病、COPD 急性加重)作为终点的前瞻性研究一般判定为2b级；以筛查率、诊断一致性等流程结局作为主要终点的前瞻性研究一般判定为2c级。CVD：心血管疾病；DM：糖尿病；COPD：慢性阻塞性肺疾病。

随访验证临床结局的证据^[7, 57-59]。其中，对于CVD和DM，已有较大规模前瞻性队列，但仍缺RCT；对于癌症，已具备跨人群外部验证；对于COPD，多停留在探索性队列，证据薄弱。这些差异凸显了多模态AI临床转化的瓶颈。

第二，方法学与数据局限。多模态AI的有效性依赖于大规模、高质量、多源异构数据的融合^[60]。然而，不同医疗中心在影像采集、实验室检测、电子病历书写上的标准差异，导致模型外部推广性受限^[61]；部分研究过度依赖单一数据集，存在过拟合与选择偏倚风险。同时，多模态数据的缺失值处理、模态间权重分配及算法可解释性问题尚未得到统一解决^[57]。

第三，临床转化潜力不足。许多研究聚焦于算法性能指标(如AUC、C-index)，而缺乏对临床流程优化、患者预后改善、成本-效益分析的系统评估。因此，即便模型在实验环境中表现优异，其临床实际应用价值仍不明确。此外，AI工具的集成需要与

临床 workflow 高度兼容，而当前缺少有效的临床试验和指南化路径。

第四，伦理与可解释性挑战。多模态AI的“黑箱”特征在循证医学框架下仍难以满足临床可解释性要求，限制了医生和患者对模型结果的信任。隐私保护、数据共享和公平性问题(如不同性别、种族、地区人群间的性能差异)同样对循证证据的普适性构成挑战^[62, 63]。

6 总结与展望

总体而言，多模态AI模型通过整合临床、影像、多组学、生理监测及生活方式等多源数据，为慢性病的筛查、风险预测、辅助诊断和个体化防治提供了新的可能。在CVD、DM、癌症和慢性呼吸系统疾病等领域，已有多项研究显示其在模型性能上优于传统方法。因此，为促进多模态AI模型在慢性病中的发展与临床应用，未来研究需从以下几方面加以完善。(1) 强化循证支撑。应在OCEBM框架

下推进多中心前瞻性研究和嵌入式真实世界试验,以验证模型对临床决策和结局的真实影响。(2) 规范方法学与提升模型泛化性。推动数据采集、标注与质量控制标准化,优化模态融合算法及缺失值处理策略,增强模型的可解释性、泛化性和跨中心可移植性,确保模型在各级医疗机构和人群中的稳定表现。(3) 促进临床转化与流程融合。在模型验证阶段引入成本-效益与流程优化指标,开展基于真实临床环境的实施研究,确保 AI 工具与临床 workflow 相兼容,提高其在慢性病诊疗中的实用性和可及性。(4) 完善伦理与公平性框架。加强隐私保护和透明度建设,系统评估不同性别、种族及地区人群间的模型表现差异,防止潜在偏倚扩大,确保 AI 模型在慢性病防治中的公平与社会信任。

只有在循证医学的系统验证与规范研究设计支持下,多模态 AI 模型才能真正从实验室走向临床实践,并最终纳入指南和决策支持体系。随着证据的不断积累与技术的不断成熟,其有望成为慢性病防治的重要工具,为精准医学与公共健康提供核心支撑。

[参 考 文 献]

- [1] GBD 2019 Diseases and Injuries Collaborators. Global burden of 369 diseases and injuries in 204 countries and territories, 1990-2019: a systematic analysis for the Global Burden of Disease Study 2019. *Lancet*, 2020, 396: 1204-22
- [2] Roth GA, Mensah GA, Johnson CO, et al. Global burden of cardiovascular diseases and risk factors, 1990-2019: update from the GBD 2019 Study. *J Am Coll Cardiol*, 2020, 76: 2982-3021
- [3] Organization WH. Noncommunicable diseases progress monitor 2020 [EB/OL]. (2020-06-12). <https://www.who.int/publications/i/item/ncd-progress-monitor-2020>
- [4] Beaglehole R, Bonita R, Horton R, et al. Priority actions for the non-communicable disease crisis. *Lancet*, 2011, 377: 1438-47
- [5] Daar AS, Singer PA, Persad DL, et al. Grand challenges in chronic non-communicable diseases. *Nature*, 2007, 450: 494-6
- [6] Guan Z, Li H, Liu R, et al. Artificial intelligence in diabetes management: advancements, opportunities, and challenges. *Cell Rep Med*, 2023, 4: 101213
- [7] Topol EJ. High-performance medicine: the convergence of human and artificial intelligence. *Nat Med*, 2019, 25: 44-56
- [8] Sackett DL, Rosenberg WM, Gray JA, et al. Evidence based medicine: what it is and what it isn't. *BMJ*, 1996, 312: 71-2
- [9] Murad MH, Asi N, Alsawas M, et al. New evidence pyramid. *Evid Based Med*, 2016, 21: 125-7
- [10] OCEBM Levels of Evidence Working Group. OCEBM levels of evidence (v2.1)[EB/OL]. <https://www.cebm.ox.ac.uk/resources/levels-of-evidence/ocebm-levels-of-evidence#:~:text=TheLevelsofEvidence,version2.1:%C2%A0English%C2%A0|%C2%A0French%C2%A0%E2%80%9393%C2%A0NancyDurieux,FrancoisePasleau|Japanese%C2%A0%E5%92%8C%E8%A8%B3%C2%A0%C2%A0%E2%80%9393%C2%A0MikikoShigemori>
- [11] Guyatt GH, Oxman AD, Vist GE, et al. GRADE: an emerging consensus on rating quality of evidence and strength of recommendations. *BMJ*, 2008, 336: 924-6
- [12] Bothwell LE, Greene JA, Podolsky SH, et al. Assessing the gold standard--lessons from the history of RCTs. *N Engl J Med*, 2016, 374: 2175-81
- [13] Collins GS, Moons KGM. Reporting of artificial intelligence prediction models. *Lancet*, 2019, 393: 1577-9
- [14] Tan SCW, Zheng BB, Tang ML, et al. Global burden of cardiovascular diseases and its risk factors, 1990-2021: a systematic analysis for the Global Burden Of Disease Study 2021. *QJM*, 2025, 118: 411-22
- [15] Joseph P, Leong D, McKee M, et al. Reducing the global burden of cardiovascular disease, Part 1: the epidemiology and risk factors. *Circ Res*, 2017, 121: 677-94
- [16] Marcu CB, Beek AM, van Rossum AC. Clinical applications of cardiovascular magnetic resonance imaging. *CMAJ*, 2006, 175: 911-7
- [17] Yusuf S, Joseph P, Rangarajan S, et al. Modifiable risk factors, cardiovascular disease, and mortality in 155 722 individuals from 21 high-income, middle-income, and low-income countries (PURE): a prospective cohort study. *Lancet*, 2020, 395: 795-808
- [18] D'Agostino RB Sr, Vasan RS, Pencina MJ, et al. General cardiovascular risk profile for use in primary care: the Framingham Heart Study. *Circulation*, 2008, 117: 743-53
- [19] Conroy RM, Pyörälä K, Fitzgerald AP, et al. Estimation of ten-year risk of fatal cardiovascular disease in Europe: the SCORE project. *Eur Heart J*, 2003, 24: 987-1003
- [20] Hippisley-Cox J, Coupland C, Brindle P. Development and validation of QRISK3 risk prediction algorithms to estimate future risk of cardiovascular disease: prospective cohort study. *BMJ*, 2017, 357: j2099
- [21] Brindle PM, McConnachie A, Upton MN, et al. The accuracy of the Framingham risk-score in different socioeconomic groups: a prospective study. *Br J Gen Pract*, 2005, 55: 838-45
- [22] Lee YC, Cha J, Shim I, et al. Multimodal deep learning of fundus abnormalities and traditional risk factors for cardiovascular risk prediction. *NPJ Digit Med*, 2023, 6: 14
- [23] Lai C, Yin M, Kholmovski EG, et al. Multimodal AI to forecast arrhythmic death in hypertrophic cardiomyopathy. *Nat Cardiovasc Res*, 2025, 4: 891-903
- [24] Kim Y, Jang TG, Park SY, et al. Multimodal AI-approach for the automatic screening of cardiovascular diseases based on nocturnal physiological signals. *NPJ Cardiovasc Health*, 2025, 2: 15
- [25] Milosevic M, Jin Q, Singh A, et al. Applications of AI in

- multi-modal imaging for cardiovascular disease. *Front Radiol*, 2023, 3: 1294068
- [26] Federation ID. IDF Diabetes Atlas (11th edition) [EB/OL]. 2025. <https://diabetesatlas.org/resources/idf-diabetes-atlas-2025/#:~:text=The%20IDF%20Diabetes%20Atlas%2011th%20edition%20provides%20detailed,region%2C%20country%20and%20territory%2C%20for%202024%20and%202050>
- [27] Bommer C, Sagalova V, Heesemann E, et al. Global economic burden of diabetes in adults: projections from 2015 to 2030. *Diabetes Care*, 2018, 41: 963-70
- [28] Zheng Y, Ley SH, Hu FB. Global aetiology and epidemiology of type 2 diabetes mellitus and its complications. *Nat Rev Endocrinol*, 2018, 14: 88-98
- [29] Harding JL, Pavkov ME, Magliano DJ, et al. Global trends in diabetes complications: a review of current evidence. *Diabetologia*, 2019, 62: 3-16
- [30] ElSayed NA, Aleppo G, Aroda VR, et al. 2. Classification and diagnosis of diabetes: Standards Of Care in Diabetes-2023. *Diabetes Care*, 2023, 46: S19-40
- [31] Colagiuri S, Lee CM, Wong TY, et al. Glycemic thresholds for diabetes-specific retinopathy: implications for diagnostic criteria for diabetes. *Diabetes Care*, 2011, 34: 145-50
- [32] Zhang K, Liu X, Xu J, et al. Deep-learning models for the detection and incidence prediction of chronic kidney disease and type 2 diabetes from retinal fundus images. *Nat Biomed Eng*, 2021, 5: 533-45
- [33] Dai L, Sheng B, Chen T, et al. A deep learning system for predicting time to progression of diabetic retinopathy. *Nat Med*, 2024, 30: 584-94
- [34] Meng Z, Guan Z, Yu S, et al. Non-invasive biopsy diagnosis of diabetic kidney disease via deep learning applied to retinal images: a population-based study. *Lancet Digit Health*, 2025, 7: 100868
- [35] Jiang N, Ji H, Guan Z, et al. A deep learning system for detecting silent brain infarction and predicting stroke risk. *Nat Biomed Eng*, 2025, 9: 1907-19
- [36] Li J, Guan Z, Wang J, et al. Integrated image-based deep learning and language models for primary diabetes care. *Nat Med*, 2024, 30: 2886-96
- [37] Mohsen F, Safa A, Shah Z. ECG features improve multimodal deep learning prediction of incident T2DM in a Middle Eastern cohort. *Sci Rep*, 2025, 15: 27164
- [38] Hahn SJ, Kim S, Choi YS, et al. Prediction of type 2 diabetes using genome-wide polygenic risk score and metabolic profiles: a machine learning analysis of population-based 10-year prospective cohort study. *EBioMedicine*, 2022, 86: 104383
- [39] Khandakar A, Chowdhury MEH, Reaz MBI, et al. A novel machine learning approach for severity classification of diabetic foot complications using thermogram images. *Sensors (Basel)*, 2022, 22: 4249
- [40] Kim RB, Gryak J, Mishra A, et al. Utilization of smartphone and tablet camera photographs to predict healing of diabetes-related foot ulcers. *Comput Biol Med*, 2020, 126: 104042
- [41] Bray F, Laversanne M, Sung H, et al. Global cancer statistics 2022: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. *CA Cancer J Clin*, 2024, 74: 229-63
- [42] Yala A, Mikhael PG, Strand F, et al. Toward robust mammography-based models for breast cancer risk. *Sci Transl Med*, 2021, 13: eaba4373
- [43] Avendano D, Marino MA, Bosques-Palomo BA, et al. Validation of the Mirai model for predicting breast cancer risk in Mexican women. *Insights Imaging*, 2024, 15: 244
- [44] Yala A, Mikhael PG, Lehman C, et al. Optimizing risk-based breast cancer screening policies with reinforcement learning. *Nat Med*, 2022, 28: 136-43
- [45] Niu C, Lyu Q, Carothers CD, et al. Medical multimodal multitask foundation model for lung cancer screening. *Nat Commun*, 2025, 16: 1523
- [46] Chen RJ, Lu MY, Williamson DFK, et al. Pan-cancer integrative histology-genomic analysis via multimodal deep learning. *Cancer Cell*, 2022, 40: 865-78.e6
- [47] Vanguri RS, Luo J, Aukerman AT, et al. Multimodal integration of radiology, pathology and genomics for prediction of response to PD-(L)1 blockade in patients with non-small cell lung cancer. *Nat Cancer*, 2022, 3: 1151-64
- [48] Bi Q, Ai C, Meng Q, et al. A multi-modal model integrating MRI habitat and clinicopathology to predict platinum sensitivity in patients with high-grade serous ovarian cancer: a diagnostic study. *Int J Surg*, 2025, 111: 4222-33
- [49] Adeyoye D, Chua S, Lee C, et al. Global and regional estimates of COPD prevalence: systematic review and meta-analysis. *J Glob Health*, 2015, 5: 020415
- [50] Fang L, Gao P, Bao H, et al. Chronic obstructive pulmonary disease in China: a nationwide prevalence study. *Lancet Respir Med*, 2018, 6: 421-30
- [51] Global burden of chronic respiratory diseases and risk factors, 1990-2019: an update from the Global Burden of Disease Study 2019. *EClinicalMedicine*, 2023, 59: 101936
- [52] Wang C, Xu J, Yang L, et al. Prevalence and risk factors of chronic obstructive pulmonary disease in China (the China Pulmonary Health [CPH] study): a national cross-sectional study. *Lancet*, 2018, 391: 1706-17
- [53] Yi Y, Shi L, Liu H, et al. COPD-MMDDxNet: a multimodal deep learning framework for accurate COPD diagnosis using electronic medical records. *Front Med (Lausanne)*, 2025, 12: 1601736
- [54] Kumar S, Bhagat V, Sahu P, et al. A novel multimodal framework for early diagnosis and classification of COPD based on CT scan images and multivariate pulmonary respiratory diseases. *Comput Methods Programs Biomed*, 2024, 243: 107911
- [55] Atzeni M, Cappon G, Quint JK, et al. A machine learning framework for short-term prediction of chronic obstructive pulmonary disease exacerbations using personal air quality monitors and lifestyle data. *Sci Rep*, 2025, 15: 2385
- [56] Sheng B, Pushpanathan K, Guan Z, et al. Artificial

- intelligence for diabetes care: current and future prospects. *Lancet Diabetes Endocrinol*, 2024, 12: 569-95
- [57] Kelly CJ, Karthikesalingam A, Suleyman M, et al. Key challenges for delivering clinical impact with artificial intelligence. *BMC Med*, 2019, 17: 195
- [58] CONSORT-AI and SPIRIT-AI Steering Group. Reporting guidelines for clinical trials evaluating artificial intelligence interventions are needed. *Nat Med*, 2019, 25: 1467-8
- [59] Fountzilias E, Pearce T, Baysal MA, et al. Convergence of evolving artificial intelligence and machine learning techniques in precision oncology. *NPJ Digit Med*, 2025, 8: 75
- [60] Qian B, Sheng B, Chen H, et al. A competition for the diagnosis of myopic maculopathy by artificial intelligence algorithms. *JAMA Ophthalmol*, 2024, 142: 1006-15
- [61] Justice AC, Covinsky KE, Berlin JA. Assessing the generalizability of prognostic information. *Ann Intern Med*, 1999, 130: 515-24
- [62] Price WN, 2nd, Gerke S, Cohen IG. Potential liability for physicians using artificial intelligence. *JAMA*, 2019, 322: 1765-6
- [63] Rajpurkar P, Chen E, Banerjee O, et al. AI in health and medicine. *Nat Med*, 2022, 28: 31-8