用于痴呆症研究方法优化的人工智能

Magda Bucholc，博士[a*]、Charlotte James，博士[b*]、Ahmad Al Khleifat，博士[c]、AmanPreet Badhwar，博士[d, e, f]、Natasha Clarke，博士[d]、Amir Dehsarvi，博士[g]、Christopher R. Madan，博士[h]、Sarah J. Marzi，博士[i, j]、Cameron Shand，博士[k]、Brian M. Schilder，哲学硕士[i, j]、Stefano Tamburin，医学博士、哲学博士[l]、Hanz M. Tantiangco，理学硕士[m]、Ilianna Lourida，博士[n]、David J. Llewellyn，博士[n, o^]、Janice M. Ranson，博士[n^]

[a] 阿尔斯特大学计算、工程与智能系统学院认知分析研究实验室，英国德里[b] NIHR布里斯托尔生物医学研究中心、布里斯托尔和韦斯顿大学医院NHS基金会信托及布里斯托尔大学，英国布里斯托尔[c] 伦敦国王学院精神病学、心理学与神经科学研究所基础与临床神经科学系，英国伦敦[d] 蒙特利尔老年学大学研究所研究中心神经退行性疾病多组学研究（MIND）实验室，加拿大蒙特利尔[e] 蒙特利尔大学生物医学工程研究所，加拿大蒙特利尔[f] 蒙特利尔大学药理学与生理学系，加拿大蒙特利尔[g] 阿伯丁大学医学、医学科学与营养学院阿伯丁生物医学成像中心，英国阿伯丁[h] 诺丁汉大学心理学学院，英国诺丁汉[i] 伦敦帝国理工学院英国痴呆症研究所，英国伦敦[j] 伦敦帝国理工学院 [illegible] 学图像计算中心，英国伦敦[l] 维罗纳大学神经科学、生物医学与运动科学系，意大利维罗纳

y

[m] 信息学院，谢菲尔德大学，谢菲尔德，英国

n 埃克塞特大学医学院，埃克塞特，英国。艾伦·图灵研究所，伦敦，英国

通讯作者： Magda Bucholc；计算机、工程与智能系统学院，阿尔斯特大学，英国，BT48 7JL。
电子邮件：m.bucholc@ulster.ac.uk

* 共同第一作者^ 共同资深作者

摘要

引言： 机器学习（ML）在从高维数据集中识别关键特征以及以达到或超过人类专家水平的准确率执行复杂任务方面取得了极大成功。

方法： 我们总结并批判性评估了机器学习在痴呆症研究中的当前应用，并指出了未来研究的方向。

结果： 我们概述了痴呆症研究中最常用的机器学习算法，并重点介绍了机器学习在临床实践、实验医学和临床试验中的未来应用机遇。我们探讨了可重复性、可再现性和可解释性问题，以及这些问题如何影响痴呆症研究的临床适用性。最后，我们举例说明了如何应用迁移学习、多任务学

讨论： 基于机器学习的模型有望极大增进我们对痴呆症潜在病因和病理机制的理解。

关键词： 痴呆、人工智能、机器学习、深度学习、分类、回归、监督学习、无监督学习、半监督学习、方法优化、泛化能力、可解释性、可再现性、可迁移性、临床效用

1. 引言

痴呆是一种与衰老相关的 疾病，其全球患病率不断上升，全球年度成本估计约为1万亿美元[1]。及时检测痴呆对于实现有效的疾病管理并提供最佳医疗保健 [2]至关重要。然而，潜在病理的复杂

2. 痴呆症研究中使用的机器学习技术类型

应用于痴呆症研究的机器学习方法大致可分为三类：i）监督学习；ii）无监督学习；以及 iii）半监督学习

2.1 监督学习

大多数机器学习方法采用监督学习，这是一种机器学习子类别，利用标注数据来学习一个能够将输入变量最佳映射到输出的目标函数[7,8]。监督学习通常分为两类问题，即分类和回归。

2.1.1 分类方法

分类模型旨在根据一组带标签的示例，确定某个实例属于一组预定义群组/类别中的哪一类。在痴呆的语境下，分类方法已应用于疾病检测、预后和管理[3,5]，其中许多方法源自逻辑回归[9-11]

比较这些方法时，最合适的模型取决于问题，并且会受到数据类型、数据收集过程及其基础分布、分类任务本身（例如痴呆症状态多分类预测）、训练和优化过程（在大多数情况下涉及人工干预）

特征[11,18]。一项比较九种监督模型、利用认知储备代理指标预测从轻度认知障碍（MCI）进展为阿尔茨海默病（AD）的研究显示，梯度提升决策树的总体准确率最高，达到93%[19]。梯度提

痴呆分类模型的构建采用了数量和类型各不相同的输入数据[15,22]。单模态机器学习框架基于认知和功能评估（CFA）[22]、磁共振成像（MRI）[23]、正电子发射断层扫描（PET）[24]、脑脊液

证据表明，纳入多种数据模态的分类方法通常能够提升模型性能[15, 28]。然而，数据类型对多模态机器学习框架整体性能的单独贡献往往未得到评估，这引出了一个问题：所提出的解决方案在

2.1.2 回归方法

与分类类似，回归旨在学习连续因变量与若干自变量之间的关系。痴呆研究中已应用多种回归方法[29-31]。最简单且最常用的形式是线性回归，即以自变量对连续因变量进行回归。在此过程

尽管线性回归是痴呆症研究中应用最广泛的回归模型，但纳入大量预测变量的线性回归模型往往会导致较差的泛化性能。为缓解这一问题，研究者提出了不同的惩罚函数，每种函数施加的约束

线性回归模型是经验模型，只描述观测数据，而并未真正理解生成这些数据的底层机制。非线性回归方法所采用的函数更为复杂，用于刻画因变量与自变量之间的关系，通常基于生成数据的底

监督式机器学习方法是痴呆研究中最常用的方法之一[7,8]。然而，分类和回归通常都需要大量的标注数据（尤其是在关系复杂时），因此较难应用于例如罕见痴呆亚型的研究，因为这类研究难

2.2 无监督学习

与监督算法不同，无监督算法会在此前未标注的数据中搜索未知模式。因此，它们在痴呆研究中特别有用，因为这类研究中的标签（例如临床诊断）可能无法获取[46]或不确定[47]。

无监督学习包含多种方法，其中包括混合效应模型[52], 项目反应理论模型[53], 高斯过程[54], 核密度估计[55]以及混合模型[56,57]。总体而言，这些模型已被用于识别疾病轨迹或亚型

无监督学习方法中最大的一类是聚类分析，其目标是根据适当的相似性度量，在数据中寻找不同的群组。传统的聚类方法，例如层次聚类和基于密度的空间聚类含噪声应用（DBSCAN），已被

与监督算法一样，多种模态的生物学数据（影像、体液、临床数据、遗传学）日益丰富，为无监督学习提供了更多机会，有望揭示复杂关系并阐明其潜在的病理生理学[50,61]。此外，利用单一

2.3 半监督学习

半监督学习介于监督学习与无监督机器学习之间，同时利用有标签和无标签数据。它通常用于数据量较大，但只有少部分样本已标注的情形。

半监督算法利用无标签数据点中的信息来提升使用少量标注数据训练的模型的性能 [62]。因此，这些方法最适用于标注数据有限的应用场景，例如缺少随访研究或获取标签的成本较高（如基于

痴呆症研究中的许多实例表明，相较于基于数据较少的监督算法，半监督算法在诊断或预后方面更具优势。Batmanghelich 等人 [64] 提出了一种降维框架，结果显示，无论是在分类器准确率

鉴于痴呆症研究中可用数据日益增多，且标签（即临床诊断）存在固有的不确定性，半监督学习提供了将临床相关特征（例如，全球临床痴呆评定量表评分）的预测与无监督学习的优势相结合

2.4 深度学习

深度学习（DL）是机器学习的一个子集，其灵感来自生物神经元的结构和信息处理方式。生物神经元层层堆叠，形成深度神经网络。与传统机器学习方法相比，深度学习的主要优势在于，数据

卷积神经网络通常用于医学影像数据，因为它们能够捕捉多尺度空间信息[74]。然而，近期研究通过整合多模态，扩展了卷积神经网络的应用数据[69]。例如，Spasov等人[69]构建了一个卷积

这些结果表明，整合多种数据模态的深度学习模型，或将成为充分利用痴呆研究中丰富信息的关键。

自编码器（AE）是一种神经网络，可用于将高维数据输入投影为潜在低维编码（编码器），同时忽略信息量较低的特征。随后，模型会使用潜在编码重建输入。例如，有研究使用自编码器，以

循环神经网络通常用于处理具有序列性或时间依赖性的数据，因为其隐藏状态组件（即“记忆单元”）允许先前的输入影响给定的输出[75]。例如，Alam 等人 [72] 采用基于长短期记忆（

深度学习方法转化为临床实践受到若干关键挑战的限制。深度学习算法需要大量数据：数据处理（例如整合不同的生物标志物）既耗时又成本高昂，而且也缺乏足够高质量的数据，尤其是在多

这是因为多个隐藏层产生了非线性特征变换，从而与其他机器学习方法相比，降低了其可解释性。缺乏可解释性可能导致人们对模型缺乏信任，进而成为实施障碍（更多详情请参见第4.2节）。

3. 采用机器学习方法开展痴呆症研究的目标研究

在痴呆症研究中，使用机器学习方法的目标是改善痴呆预防、诊断、治疗和照护。由于痴呆的复杂性以及可用数据集的规模和复杂性不断增加，传统统计方法往往已不足以应对。机器学习能够

3.1 临床实践中的机器学习

在临床实践中，机器学习已应用于痴呆风险评估、临床诊断、预后和照护[3]。在这些应用中，诊断研究主要侧重于区分认知障碍的不同阶段或痴呆亚型[15,77]。许多研究将神经影像数据作为输

策略，可用于早期检测痴呆状态或易感性，更稳健地识别疾病靶点，并确定不同生物标志物、症状与临床亚型之间的因果联系。

由于目前尚无针对阿尔茨海默病及其他神经退行性疾病的成熟疾病修饰疗法，大多数机器学习预后研究都着重于预测轻度认知障碍向痴呆的转化。所采用的数据包括磁共振成像[80]、脑电图

根据痴呆症研究中机器学习方法相关文献，可以合理地说，已有多种不同算法被证明能够以相对较高的预测准确率检测痴呆及其前驱期，但在考察其他性能指标时，其表现差异显著[15]。在某

除了用于比较性能的机器学习指标存在显著差异外，在诊断、预后和风险预测算法能够常规应用于记忆门诊等场景之前，还需要解决其他问题。这些问题包括可重复性、模型验证和数据泄漏：

参数迁移学习方法（先在规模更大、任务针对性较低的数据集上训练，再在规模较小、针对性更强的数据集上微调），以更新基线模型的决策边界，在所有性能指标上均取得更好表现，包括敏

最后，在新一代旨在发挥数据驱动决策潜力的临床决策支持系统（CDSSs）得到采用并在临床实践中常规使用之前，必须向临床终端用户清晰说明机器学习系统的局限性。尽管已有少数研究报

3.2 实验医学中的机器学习

理论上，独特小分子的数量比可观测宇宙中的原子数量高出一个数量级以上 [93]。因此，通过穷举暴力计算来优化合成分子的设计，以获得理想的治疗效果，在计算上是无法完成的任务，

而人工智能有望以高得多的效率搜索这一复杂空间，从而设计出最接近可行候选分子的药物[94]。尤其是深度学习，已证明能够胜任这些任务，因为它能够学习多模态数据（例如神经影像学、基因组学和临床记录）的新表征，并运用非线性映射 。因此，人工智能在药理学及其相关领域——化学信息学和结构生物学——的应用迅速增加 [96]。应用案例包括：材料维结构和功能 [97]；蛋白质-蛋白质或蛋白质-药物相互作用建模；临床结局或生物标志物预测[98,99]；以及个体化基因组-药物相互作用（即药物基因组学）。除了设计新药，人工智能还被用于重新利用原本针对其他治疗用途而开发的现有疗法。与痴呆尤其相关的是，Dias 等人近期的一项研究 [100] 使用 IBM Watson 药物发现平台在线工具，并结合自然语言处理算法，从数百万篇医学领域已发表的研究文章中提取基因-疾病和药物-疾病关系列表。随后，他们将这些关系与人脑样本的基因共表达数据相结合，构建了一个扩展的知识网络，揭示了不同精神疾病和神经系统疾病与数百种药物之间此前未知的关系。因此，他们确定了若干可重新定位为阿尔茨海默病（n=25）、帕金森病（n=1）和痴呆（n=1）疗法的候选药物（详情见原文的补充表4）[100]。此外，Rodriguez 等人 [101]利用基于机器学习的阿尔茨海默病药物再利用（DRIAD）框架，评估了阿尔茨海默病病理阶段与基因之间的潜在关联。在这项研究中，研究人员将 DRIAD 应用于一组基因，这些基因在神经元细胞暴露于由80种临床批准药物组成的测试组合后出现差异表达；由此生成了可能的再利用候选药物排名。经过在相关体外和体内阿尔茨海默病模型系统中的进一步验证后，这些候选药物可在临床试验中进行评估。

在药物发现领域之外，人工智能也已应用于疾病基因组学，例如预测蛋白质编码和非编码基因组序列中突变的功能后果 [102,103]。与明确的基于规则的模型相比，这类序列信息驱动的机器学习方法具有多项显著优势，包括能够进一步开展计算机模拟 诱变 ，以探究所有可能突变的影响[104]。这些是进行功能影响预测（例如基因表达、染色质修饰）的强大工具。然而，在根据基因组数据预测疾病状态方面，基于机器学习的方法尚未显示出显著优于更简单的加性方法的性能提升，例如用于自身免疫性疾病或阿尔茨海默病的多基因风险评分 [26]。这可能由多种因素造成，包括样本量不足、使用基因分型芯片而非全基因组测序、使用经过高度处理的输入数据（例如通过简单线性模型识别出的疾病相关变异），或针对这一问题采用了次优的机器学习架构。

另一项挑战是，序列预测模型通常并未考虑突变效应的组织和细胞类型特异性。单细胞转录组学、表观基因组学和蛋白质组学的最新进展，使研究人员得以在痴呆相关疾病的模式生物、在世患者的原代细胞（例如血液）[105],、尸检样本（例如脑组织）[106], 以及患者来源的诱导多能干细胞（iPSCs）[107]中积累单细胞图谱。已有一些专门的数据库用于整合和托管阿尔茨海默病相关的单细胞数据集[108]。随后，这些数据集可用于训练深度学习模型，以学习细胞类型特异性响应各种遗传和化学扰动的潜在表征；此前，这种方法已应用于癌细胞系[109], 以及天然基因组变异[110,111]。因此，大规模评估痴呆相关突变的细胞类型特异性效应正变得越来越可行。

3.3 临床试验中的机器学习

人工智能和机器学习在痴呆临床试验中的应用是一个备受研究关注的领域。这项研究的一个重点是开发具备识别并诊断患者疾病能力的算法，这有助于提高临床试验患者筛选的准确性，并协助

多项研究探讨了机器学习在临床试验设计中的作用[112,113]。Ezzati 和 Lipton[114]开发了一种采用 K近邻算法的机器学习框架，用于识别在随访期间更有可能出现认知衰退的个体，并将这一

人们日益认识到，痴呆症发病前会经历时长不一的无症状期或前驱期，在此期间，潜在疾病进程逐渐展开。这凸显了在疾病不同无症状阶段延缓疾病进展的机会；在这些阶段，病理变化更有可

尽管已经开发出大量用于痴呆症研究的机器学习模型，但在临床试验开始前用于患者与试验匹配及招募，或在试验期间用于患者监测的模型仍然不多。仍有一些挑战有待解决，包括确保患者数

4. 痴呆症研究中的可重复性、可再现性、可解释性和临床适用性问题

4.1 可重复性和可再现性

随着我们开发新的计算模型以理解痴呆，可重复性和可再现性问题日益重要。可重复性是指获得相同的

模型和数据所产生的结果，而可再现性则是将相同模型应用于独立数据集，并观察研究发现的泛化能力。

可重复性是科学研究，尤其是实验医学中长期存在的问题，因为结果会受到可控因素（例如遵循相同的方案）和不可控因素（例如系统随机性）的影响。因此，使用计算模型可显著提高可重复

与可重复性相比，可再现性更难实现：针对某个队列优化、例如用于预测新发痴呆症的算法，并不能保证在完全不同的队列中也能表现良好。对于神经网络等在训练过程中容易过拟合的算法而

迁移学习和多任务学习等最先进的机器学习方法，有望提高结果的可再现性和泛化能力。例如，迁移学习可用于在来自不止一项研究的数据上训练算法，使两者中部分变量

数据集相同，并通过在数据域之间寻找映射关系。此类方法可有助于提高结果的泛化能力，并减少对用于训练的队列的过拟合，从而增加在后续队列中实现可再现性的可能性。Danso 等人已展

迁移学习意味着表征按顺序共享，而多任务学习则涉及在不同任务间并行开发的共享表征。迄今为止，只有少数研究在痴呆症背景下采用了多任务机器学习方法[123]。其中一个例子是采用基于

4.2 机器学习模型的可解释性

除了可重复性和可再现性之外，机器学习算法的可解释性也是许多应用中的重要考量，尤其是在算法所做的决策会改变人生的情况下。痴呆症研究和医疗保健应用显然属于此类。在临床决策者

算法的可解释性可分为两类：全局可解释性和局部可解释性。全局可解释性指算法的各个组成部分（例如特征和权重）如何组合起来做出决策。相比之下，局部可解释性关注单个决策：算法如

。不过，可以在模块层面探查算法，以了解改变某个特征或权重会如何影响所做出的决策。

痴呆症研究中用于解释算法决策的最常见方法之一是（置换）特征重要性，它正是通过告知用户每个特征对所作决策的贡献程度来实现这一点[127]。特征重要性的一项优势是它与模型无关，因

痴呆症研究中常用的机器学习算法，例如决策树、支持向量机和贝叶斯网络，本身具有可解释性；它们会学习变量之间的一组规则或关系，其结构已知，用户可以对其进行探查。相比之下，“黑

人们可能会质疑，是否有必要采用方法来探查“黑箱”模型，以用于痴呆症研究。文献中有许多案例比较了不同机器学习算法的表现。在痴呆症诊断任务中，可解释算法的表现往往不逊于，甚至

[4,135]；可解释算法在预测能力方面可能表现更好[136]，而“黑箱”模型通常具有更高的判别准确率[135]。

准确率的提升可能会使“黑箱”模型显得更具吸引力，但这是以牺牲可解释性为代价的。近期综述发现，尽管“黑箱”模型正被用于痴呆症研究，研究人员仍然倾向于选择具有可解释性的模型[3,81]。

4.3 临床适用性问题

目前，成功应用于临床实践的痴呆症诊断、预后和管理机器学习模型案例仍然有限。与数据集收集相关的问题，例如数据集偏移（即训练集和测试集中的输入与输出联合分布不同的情形），遗

这些问题的影响在神经影像学领域有所体现：使用 ADNI 数据训练、用于预测区域性脑萎缩的卷积神经网络，其均方误差从 0.31 上升至0.41（在记忆门诊数据上测试时）[138]。当训练集纳入

基于认知数据训练模型的研究也面临类似问题。值得欣慰的是，一款用于评估整体认知的人工智能 iPad应用在检测轻度认知障碍和轻度阿尔茨海默病方面表现良好，使用两个语言和文化背景各

可能限制临床适用性的相关问题包括资源的成本和可用性，例如，依赖昂贵的影像数据。多组学方法可能会加剧这一问题。因此，采用互补且独特的方法来发现最具信息性的特征将是关键[144]。

以及可解释性较低的模型；采用透明、可解释的方法[146]，并报告算法对其决策的置信度 [147]，可以提高信任度。此外，性能指标应力求反映真实的临床适用性，并且易于临床医生理解。鉴于

5. 讨论

5.1 现有方法及其局限性概述

显然，人工智能和机器学习在痴呆症研究中的应用日益普遍。监督学习算法的应用最为频繁：分类通常用于预测痴呆症进展风险和进行痴呆鉴别诊断，而回归则用于检测认知变化、评估痴呆症严重程度、估计痴呆症发病时间、识别风险因素 [149] 和痴呆亚型 [38,150]。在无法获取或标注数据不足的情况下，研究人员正采用无监督和半监督方法，并取得了令人鼓舞的结果[58,66]。此外，近年来，深度学习模型在痴呆症早期检测和分类中的应用备受关注，这主要归因于其性能优于传统机器学习方法[69]。然而，目前在临床环境中成功应用机器学习模型仍是高风险之举，因为这些模型过度依赖单一数据来源；外部验证质量欠佳；对模型预测的机制基础及不同性能指标的临床效用缺乏或未充分理解；并且缺失数据或不恰当地使用缺失数据插补方法可能带来偏倚。其他偏倚来源，包括测量偏倚、评估偏倚、抽样偏倚，

以及算法偏差，也会削弱机器学习模型向新人群泛化的能力，并可能导致对特定群体或个人的歧视。例如，正如Gianattasio等人所指出的，用于训练的无代表性数据集以及模型校准不佳，可能导致模型应用中的种族偏差。[151]此外，大型数据集（例如ADNI、NACC及其他类似的、经过高度筛选的临床队列）所得研究发现缺乏泛化能力，难以适用于其他人群（如基于社区的样本），这可能严重限制其生成可轻易迁移至其他环境的机器学习模型的能力[152]。痴呆症机器学习算法开发和验证所需数据的可获得性，也是阻碍其在临床实践中应用和常规使用的一项重要因素。大多数机器学习模型需要大量数据进行训练和测试，以确保其在训练集之外也具备泛化能力。许多机器学习模型具有“黑箱”特性，尤其是深度学习模型，其输入数据可能经过多层复杂变换，因此缺乏明确的声明式知识表示，也就无法在提供预测时附上相应的解释。其他机器学习方法或许能够列出目标预测输出与输入特征之间的依赖关系，但这些关系过于复杂，难以理解或验证。最后，由于尚未针对某些模型结构和应用场景规定必须采用的性能指标，因此难以客观比较不同痴呆研究中的模型。机器学习应用面临的所有这些障碍，都需要在模型开发、验证和使用的各个阶段予以关注，以增强利益相关者（包括监管机构、研究人员、临床医生和行业合作伙伴）对这一过程及其结果的信任。

5.2 未来研究建议

虽然机器学习在痴呆症研究中的应用为未来带来了许多机遇，但要将算法应用于临床实践，仍需解决可再现性、泛化能力、可解释性、可比性以及决策可信度等相关问题。

可再现性

问题：大多数采用机器学习方法的痴呆研究都假设，只要模型在其所处理的特定领域或问题的数据上表现良好，就可以成功应用于新数据。然而，这一假设在许多情况下并不成立。由于模型是针对特定问题/数据集构建的，因此必须针对任何新问题/数据集重新训练和调优。这不仅会影响所提出解决方案的时间和成本效益，还可能导致模型性能发生变化，而这种变化通常是由数据分布或特征空间的不一致造成的。

建议：迁移学习允许将预训练模型重复用于特定领域的任务，是一种节省时间的替代方案。在痴呆症研究中应用迁移学习框架已显示出一些令人鼓舞的成果，包括提高准确率，以及在处理相对较小的数据集时防止过拟合 [153]。

泛化能力

问题：在为临床研究生成的数据集中，某些疾病阶段和患者特征不太可能得到充分体现。临床队列对其他人群的泛化能力不足，限制了其生成能够适应其他数据集的机器学习模型的能力。此外，使用回顾性数据集开发机器学习模型，并不总能顺利应用于前瞻性应用。

痴呆领域中使用的机器学习模型绝大多数都是纯粹的关联模型，即侧重于根据预先定义的一组变量来预测结局。由于这些模型并非为理解输入数据与输出预测之间的因果关系而构建，因此当输入数据发生变化，或患者症状存在多种可能病因时（例如痴呆的鉴别诊断），它们无法很好地泛化。痴呆领域机器学习算法的关联性限制了其性能，并可能导致诊断不够理想。

建议：为确保机器学习模型能够泛化至训练数据集之外，开发者应使用多个数据集来测试模型性能的稳定性。此外，可利用因果推断，即理解如何得出诊断并明确界定预期输出，从而构建更稳健、泛化能力更强的机器学习模型 [154]。针对痴呆领域机器学习干预措施开展的临床试验，可以评估机器学习模型的临床效用，包括其在实际部署中可能遇到的问题。

可解释性

问题：算法决策的可解释性和可信度，是阻碍研究成果转化到临床环境的因素。

建议：可以使用 LIME 和 Shapley 值等方法来提升模型的可解释性和可信度。这些方法使临床医生能够探究机器学习模型的决策过程，从而提高透明度。

可比性

问题：根据机器学习模型的预期使用场景不同，不同指标会对模型性能产生不同的解读。

建议：可以针对特定模型类型和使用场景规定强制性性能指标，以确保预期用途相近的模型能够进行比较 [155]。应定期重新评估这些指标，以检测潜在的模型漂移。

结论

要提高模型的泛化能力和可重复性，构建这些模型所依据的痴呆症数据必须广泛提供给其他研究人员并便于其获取，同时还应能代表模型将在临床环境中应用于的人群。未来，先进的

方法，如迁移学习、因果建模、强化学习和多任务学习，或将得到应用。这些方法不仅可用于标注数据稀缺的情境，还可用于提高模型的泛化能力，增强其在不同人群中的有效性，从而提升临床适用性。如果痴呆症研究人员与临床医生合作，克服临床转化的障碍，机器学习有望在未来改变痴呆症研究与照护的格局。

致谢

感谢深度痴呆表型研究（DEMON）网络科学现状研讨会的合作者和参与者（按字母顺序排列）：Peter Bagshaw、Robin Borchert、Magda Bucholc、James Duce、Charlotte James、David Llewellyn、Donald Lyall、Sarah Marzi、Danielle Newby、Neil Oxtoby、Janice Ranson、Tim Rittman、Nathan Skene、Eugene Tang、Michele Veldsman、Laura Winchester、Zhi Yao。

资金

本文源自DEMON网络举办的科学现状研讨会“利用数据科学和人工智能开展痴呆症研究”，该研讨会由英国阿尔茨海默症研究会资助。JMR和DJL获得英国阿尔茨海默症研究会以及艾伦·图灵研究所/工程与自然科学研究委员会（EP/N510129/1）的支持。DJL还获得了国家健康研究院（NIHR）应用研究协作（ARC）西南半岛、国家健康与医学研究委员会（NHMRC）、JP Moulton基金会、国家老龄研究所/美国国立卫生研究院（RF1AG055654）以及艾伦·图灵研究所/工程与自然科学研究委员会（EP/N510129/1）的资助。本研究还获得英国阿尔茨海默症研究会（MB、CJ）、欧盟INTERREG VA计划（MB）、George Moore博士数据捐赠基金的支持

阿尔斯特大学科学（MB）、英国国家健康与护理研究院布里斯托尔生物医学研究中心（CJ）、魁北克健康研究基金 - 初级研究员1（AB）、加拿大衰老神经退行性疾病联盟和库尔图瓦基金会（AB、NC）、运动神经元疾病协会研究奖学金（Al Khleifat/Oct21/975-799）（AAK）、肌萎缩侧索硬化症协会米尔顿·萨芬诺维茨研究奖学金（资助编号22-PDF-609。
doi:10.52546/pc.gr.150909)（AAK）、英国国家健康与护理研究院莫兹利生物医学研究中心（AAK）、达比里默基金会（AAK）、英国研究与创新署未来领袖奖学金（MR/S03546X/1）（CS）、E-DADS项目（欧盟神经退行性疾病联合计划）（CS）、EuroPOND项目（欧盟“地平线2020”计划，编号666992）（CS）。SJM的经费来自埃德蒙和莉莉·萨夫拉青年研究人员奖学金计划以及英国痴呆症研究所。该研究所的经费由英国痴呆症研究所有限公司提供，而英国痴呆症研究所有限公司则由英国医学研究委员会、阿尔茨海默病协会和英国阿尔茨海默症研究会资助。

利益冲突

所有作者均声明不存在利益冲突。

参考文献

[1] Patterson C.《2018年世界阿尔茨海默病报告》。2020年。[2] Robinson L、Tang E、Taylor J。痴呆：及时诊断与早期干预。*BMJ*。2015；350。[3] Pellegrini E、Ballerini L、Hernandez、Maria del C Valdes 等。利用神经影像学进行认知障碍和痴呆辅助诊断的机器学习：系统综述。*Alzheimer's & Dementia: Diagnosis, Assessment & Disease Monitoring*。2018；10：519-535。[4] Ahmed MR、Zhang Y、Feng Z、Lo B、Inan OT、Liao H。神经影像学与机器学习在痴呆症诊断中的应用：最新进展与未来展望。《*IEEE*生物医学工程评论》。2018；12：19-33。[5] Enshaeifar S、Zoha A、Markides A 等。利用家庭传感器和机器学习技术进行痴呆患者日常生活活动的健康管理与模式分析。*PLOS ONE*。2018；13(5)：e0195605。[6] Mitelpunkt A、Galili T、Kozlovski T 等。采用数据与知识驱动策略识别出的新型阿尔茨海默病亚型。《科学报告》。2020年；10(1)：1-13。[7] Kumar S、Oh I、Schindler S、Lai AM、Payne PRO、Gupta A。利用临床数据对阿尔茨海默病性痴呆进展进行建模的机器学习：系统性文献综述。《*JAMIA Open*》。2021年；4(3)：ooab052。doi：10.1093/jamiaopen/ooab052。[8] Grueso S、Viejo-Sobera R。预测轻度认知障碍进展为阿尔茨海默病性痴呆的机器学习方法：系统综述。《阿尔茨海默病研究与治疗》。2021年；13(1)：1-162。doi：10.1186/s13195-021-00900-w。[9] Barnes DE、Covinsky KE、Whitmer RA、Kuller LH、Lopez OL、Yaffe K。痴呆风险指数：识别高痴呆风险个体的框架。*Alzheimer's & dementia: thejournal of the Alzheimer's Association*。2010年；6(2)：138。[10] Chary E、Amieva H、Pérès K、Orgogozo J、Dartigues J、Jacqmin-Gadda H。短期与长期-

教育水平较低和较高的受试者中痴呆的终末期预测。《阿尔茨海默病与痴呆》。2013年;9(5):562-571。[11] Maroco J, Silva D, Rodrigues A, Guerreiro M, Santana I, de Mendonça A. 用于痴呆预测的数据挖掘方法：基于真实数据，对线性判别分析、逻辑回归、神经网络、支持向量机、分类树和随机森林的准确率、敏感度及特异度进行比较。《*BMC*研究札记》。2011年;4(1):1-14。[12] Lebedev AV, Westman E, Van Westen G, 等。用于检测和预测阿尔茨海默病且具有良好队列间稳健性的随机森林集成。《神经影像：临床》。2014年;6:115-125。

[13] Wei W, Visweswaran S, Cooper GF. 将朴素贝叶斯模型平均应用于基于全基因组数据预测阿尔茨海默病。《美国医学信息学会杂志》。2011年;18(4):370-375。

[14] Kruthika KR、Maheshappa HD、阿尔茨海默病神经影像学倡议。基于多阶段分类器的方法用于阿尔茨海默病预测与检索。《医学信息学开放》。2019年；14：34-42。

[15] Bucholc M、Ding X、Wang H 等。一种实用的计算机化决策支持系统，用于预测个体阿尔茨海默病的严重程度。《专家系统与应用》。2019年；130：157-171。

[16] Sharma A、Kaur S、Memon N、Fathima AJ、Ray S、Bhatt MW。使用支持向量机（SVM）并通过定量分析检测阿尔茨海默病患者。《神经科学信息学》。2021年11月1日；1（3）：100012。

[17] Moore PJ、Lyons TJ、Gallacher J、阿尔茨海默病神经影像学倡议。利用时间序列数据中的成对选择，通过随机森林预测阿尔茨海默病。*PLOS ONE*。2019年；14（2）：e0211558。

[18] Lehmann C, Koenig T, Jelic V, 等。用于识别阿尔茨海默病的分类算法的应用与比较

脑电活动（脑电图）。*J Neurosci Methods*。2007;161(2):342-350。[19] Facal D, Valladares-Rodriguez S, Lojo-Seoane C, Pereiro AX, Anido-Rifon L, Juncos- Rabadán O。运用机器学习方法研究认知储备在轻度认知障碍转化为痴呆过程中的作用。*Int J Geriatr Psychiatry*。2019;34(7):941-949。[20] Shigemizu D, Akiyama S, Asanomi Y, Boroevich KA, Sharma A, Tsunoda T, Sakurai T, Ozaki K, Ochiya T, Niida S。利用miRNA表达数据比较路易体痴呆的机器学习分类器。*BMC medical genomics*。2019;12(1):1-0。[21] James C, Ranson JM, Everson R, Llewellyn DJ。机器学习算法预测记忆门诊患者痴呆进展的表现。*JAMA network open*。2021;4(12):e2136553-。[22] Cui Y, Liu B, Luo S, 等。利用多变量预测因子识别轻度认知障碍向阿尔茨海默病的转化。*PloS one*。2011;6(7):e21896。[23] Moradi E, Pepe A, Gaser C, Huttunen H, Tohka J, 阿尔茨海默病神经影像学倡议。基于机器学习框架，利用MRI早期预测轻度认知障碍受试者向阿尔茨海默病的转化。*Neuroimage*。2015;104:398-412。[24] Grimmer T, Wutz C, Alexopoulos P, 等。视觉分析与全自动分析18F-FDG和淀粉样蛋白正电子发射断层扫描在预测轻度认知障碍患者因阿尔茨海默病导致痴呆方面的比较。*Journal of Nuclear Medicine*。2016;57(2):204-207。

[25] Forlenza OV, Radanovic M, Talib LL 等。阿尔茨海默病的脑脊液生物标志物：诊断准确性及痴呆预测。*Alzheimer's & Dementia: Diagnosis,Assessment & Disease Monitoring*. 2015;1(4):455-463。[26] De Velasco Oriol J, Vallejo EE, Estrada K, Tamez Pena JG。基于基因组数据预测晚发型阿尔茨海默病的机器学习模型基准测试。*BMC Bioinformatics*. 2019;20。

[27] Korolev IO, Symonds LL, Bozoki AC, 阿尔茨海默病神经影像学倡议。利用临床变量、磁共振成像和血浆生物标志物，通过概率模式分类预测轻度认知障碍向阿尔茨海默型痴呆的进展。*PLOS ONE*. 2016;11(2):e0138866。[28] Prakash M, Abdelaziz M, Zhang L, Strange BA, Tohka J, 阿尔茨海默病神经影像学倡议。通过多模态预测学习对阿尔茨海默病进行定量纵向预测。《阿尔茨海默病杂志》. 2020(预印本):1-14。[29] Duchesne S, Caroli A, Geroldi C, Collins DL, Frisoni GB。轻度认知障碍患者一年的认知变化与基线磁共振成像特征的关联。《神经影像》. 2009;47(4):1363-70。[30] Duchesne S, Caroli A, Geroldi C, Frisoni GB, Collins DL。根据磁共振成像特征预测临床变量：在轻度认知障碍患者简易精神状态检查中的应用。载于：2005年10月26日国际医学图像计算与计算机辅助干预会议（第392-399页）。施普林格，柏林，海德堡。[31] Youssofzadeh V, McGuinness B, Maguire LP, Wong-Lin K。采用 DARTEL 的多核学习改进基于澳大利亚影像、生物标志物与生活方式研究数据的阿尔茨海默病磁共振成像-正电子发射断层成像联合分类：群体与个体分析。《人类神经科学前沿》. 2017;11:380。[32] Mofrad SA, Lundervold A, Lundervold AS, 阿尔茨海默病神经影像学倡议。基于脑容量轨迹、可实现阿尔茨海默病早期检测的预测框架。《计算机化医学影像与图形学》. 2021;90:101910。[33] Growdon ME, Schultz AP, Dagley AS 等。临床认知正常老年人的气味识别与阿尔茨海默病生物标志物。《神经病学》. 2015;84(21):2153-2160。[34] Wolters FJ, Zonneveld HI, Hofman A, van der Lugt A, Koudstaal PJ, Vernooij MW。心脑关联协作研究组。脑灌注与痴呆风险。《循环》. 2017;136:719-728。[35] Mattsson N, Zetterberg H, Janelidze S 等。阿尔茨海默病中的血浆 tau 蛋白。《神经病学》. 2016;87(17):1827-1835。[36] Lee S, Viqar F, Zimmerman ME 等。脑白质高信号是阿尔茨海默病的核心特征：来自常染色体显性遗传阿尔茨海默病网络的证据。《神经病学年鉴》。

2016;79(6):929-939.[37] Irwin DJ, Grossman M, Weintraub D, 等. 突触核蛋白病患者生存情况和痴呆发病的神经病理学及遗传学相关因素：一项回顾性分析。《柳叶刀·神经病学》. 2017;16(1):55-65.[38] Asanomi Y, Shigemizu D, Akiyama S, Sakurai T, Ozaki K, Ochiya T, Niida S. 使用血清微RNA表达数据，通过惩罚回归方法构建用于多分类的痴呆亚型预测模型。《科学报告》 2021;11(1):20947.[39] Agbavor F, Liang H. 使用大语言模型根据自发性言语预测痴呆。《*PLOS*数字健康》. 2022;1(12):e0000168.[40] Yang T, Wang J, Sun Q, 等. 通过套索筛选从全基因组测序数据中检测阿尔茨海默病遗传风险因素。 . 2015:985-989.[41] Dayon L, Guiraud SP, Corthésy J, 等. 一碳代谢、认知障碍与阿尔茨海默病病理的脑脊液指标：同型半胱氨酸及其他因素。《阿尔茨海默病研究与治疗》. 2017;9(1):1-11.[42]. Silver M, Janousova E, Hua X, Thompson PM, Montana G, 阿尔茨海默病神经影像学倡议. 利用纵向影像表型和稀疏回归识别与阿尔茨海默病相关的基因通路。《神经影像》. 2012;63(3):1681-1694.[43] Teipel SJ, Grothe MJ, Metzger CD, 等. 使用弹性网络正则化回归稳健检测阿尔茨海默病患者受损的静息态功能连接网络。《衰老神经科学前沿》. 2017;8:318.[44] Kaut O, Schmitt I, Tost J, 等. 对散发性帕金森病患病状态不一致的同胞及同卵双胞胎进行全表观基因组DNA甲基化分析，发现外周血单个核细胞中存在不同的表观遗传模式。《神经遗传学》. 2017;18(1):7-22.[45] Bouts MJ, Möller C, Hafkemeijer A, 等. 使用解剖、弥散张量和静息态功能磁共振成像进行阿尔茨海默病与行为变异型额颞叶痴呆的单受试者分类。《阿尔茨海默病杂志》. 2018;62(4):1827-1839.

[46] Cleret de Langavant L, Bayen E, Bachoud-Lévi A, Yaffe K. 采用无监督机器学习方法估算全球衰老人群调查中的痴呆患病率。《阿尔茨海默病与痴呆：转化研究与临床干预》. 2020;6(1):e12074.[47] Ranson JM, Kuźma E, Hamilton W, Muniz-Terrera G, Langa KM, Llewellyn DJ. 使用简短认知评估时痴呆误分类的预测因素。《神经病学：临床实践》. 2019;9(2):109-117.[48] Whitwell JL, Przybelski SA, Weigand SD, 等. 行为变异型额颞叶痴呆的不同解剖亚型：一项聚类分析研究。《大脑》. 2009;132(11):2932-2946.[49] Toschi N, Lista S, Baldacci F, 等. 生物标志物引导的阿尔茨海默病临床综合征聚类。《衰老神经生物学》. 2019;83:42-53.[50] Oxtoby NP, Alexander DC. 影像加X：神经退行性疾病的多模态模型。*Curr Opin Neurol*. 2017;30(4):371.[51] Arbelaitz O, Gurrutxaga I, Muguerza J, Pérez JM, Perona I. 聚类有效性指标的广泛比较研究。《模式识别》. 2013;46(1):243-256.[52] Raket LL. 阿尔茨海默病的疾病进展统计建模。《大数据前沿》. 2020;3.[53] Leoutsakos J, Gross AL, Jones RN, Albert MS, Breitner J. “阿尔茨海默病进展评分”：为阿尔茨海默病预防试验开发生物标志物汇总结局指标。《阿尔茨海默病预防杂志》. 2016;3(4):229.[54] Lorenzi M, Filippone M, Frisoni GB, Alexander DC, Ourselin S, 阿尔茨海默病神经影像学倡议. 概率性疾病进展建模，用于描述诊断不确定性：在阿尔茨海默病分期与预测中的应用。《神经影像》. 2019;190:56-68.[55] Firth NC, Primativo S, Brotherhood E, 等. 典型认知衰退的序列

采用一种新型疾病进展事件模型估算阿尔茨海默病和后部皮质萎缩。《阿尔茨海默病与痴呆》。2020年;16(7):965-973。[56] Fonteijn HM, Modat M, Clarkson MJ, 等。用于疾病进展的基于事件的模型及其在家族性阿尔茨海默病和亨廷顿病中的应用。《神经影像》。2012年;60(3):1880-1889。[57] Young AL, Oxtoby NP, Daga P, 等。散发性阿尔茨海默病中生物标志物变化的数据驱动模型。《大脑》。2014年;137(9):2564-2577。[58] Eshaghi A, Young AL, Wijeratne PA, 等。利用无监督机器学习和磁共振成像数据识别多发性硬化亚型。《自然·通讯》。2021年;12(1):1-12。[59] Vogel JW, Young AL, Oxtoby NP, 等。在阿尔茨海默病中识别出四种不同的tau蛋白沉积轨迹。《自然·医学》。2021年;27(5):871-881。[60] Young AL, Marinescu RV, Oxtoby NP, 等。利用亚型与阶段推断揭示神经退行性疾病的异质性和时间复杂性。《自然·通讯》。2018年;9(1):1-16。[61] Khanna S, Domingo-Fernández D, Ayyappan A, Emon MA, Hofmann-Apitius M, Fröhlich H。利用多尺度遗传、神经影像学和临床数据预测阿尔茨海默病，并重建相关生物学机制。《科学报告》。2018年;8(1):1-13。[62] Van Engelen JE, Hoos HH。半监督学习综述。《机器学习》。2020年;109(2):373-440。[63] An L, Adeli E, Liu M, Zhang J, Shen D。用于阿尔茨海默病诊断的半监督分层多模态特征与样本选择。. 2016:79-87。[64]. Batmanghelich KN, Dong HY, Pohl KM, Taskar B, Davatzikos C。通过半监督降维进行疾病分类与预测。. 2011:1086-1090。[65] Filipovych R, Davatzikos C, 阿尔茨海默病神经影像学倡议。医学图像的半监督模式分类：轻度认知障碍（MCI）的应用。《神经影像》。2011年;55(3):1109-1119。

[66] Teramoto R。利用标签传播的半监督距离度量学习预测阿尔茨海默病诊断。*Comput Biol Chem*。2008年;32(6):438-41。[67] Bengio Y, LeCun Y。将学习算法扩展至人工智能领域。大规模核机器。2007;34(5):1-41。[68] Morabito FC, Campolo M, Ieracitano C, 等。基于头皮脑电图记录，使用深度卷积神经网络对轻度认知障碍患者和阿尔茨海默病患者进行分类。. 2016:1-6。[69] Spasov S, Passamonti L, Duggento A, Lio P, Toschi N, 阿尔茨海默病神经影像学倡议。一种参数高效的深度学习方法，用于预测轻度认知障碍向阿尔茨海默病的转化。《神经影像》。2019年;189:276-287。[70] Ithapu VK, Singh V, Okonkwo OC, 等。利用深度学习算法制定基于影像的富集标准，以提高轻度认知障碍临床试验的效率。《阿尔茨海默病与痴呆》。2015年;11(12):1489-1499。[71] Vu T, Ho N, Yang H, Kim J, Song H。非白质组织提取与深度卷积神经网络用于阿尔茨海默病检测。*Soft Computing*。2018年;22(20):6825-6833。[72] Alam R, Anderson M, Bankole A, Lach J。利用智能手表和序列行为模型推断痴呆患者的身体躁动。. 2018:170-173。[73] Li H, Fan Y。基于基线海马磁共振成像和1年随访认知测量指标，利用深度循环神经网络早期预测阿尔茨海默病性痴呆。发表于：2019年IEEE第16届生物医学成像国际研讨会（ISBI 2019），2019年4月8日（第368-371页）。IEEE。[74] Sarvamangala DR, Kulkarni RV。医学图像理解中的卷积神经网络：综述。*Evolutionary intelligence*。2021年：1-22。[75] Navamani TM。健康信息学中的高效深度学习方法。载于：生物工程系统的深度学习与并行计算环境。Elsevier；2019：123-137。

[76] Ebrahimighahnavieh MA, Luo S, Chiong R. 利用深度学习从神经影像中检测阿尔茨海默病：系统性文献综述。*Comput Methods Programs Biomed*.2020;187:105242.[77] Klöppel S, Stonnington CM, Chu C, et al. 阿尔茨海默病磁共振扫描的自动分类。《大脑》. 2008;131(3):681-689.[78] Burgos N, Colliot O. 用于脑部疾病分类和预测的机器学习：最新进展与未来挑战。*Curr Opin Neurol*. 2020;33(4):439-450.[79] Gupta Y, Lama RK, Kwon G, et al. 基于载脂蛋白E基因型、脑脊液、磁共振和氟脱氧葡萄糖正电子发射断层成像生物标志物的组合特征预测和分类阿尔茨海默病。*Frontiers in Computational Neuroscience*. 2019;13:72.[80] Bron EE, Klein S, Reinke A, Papma JM, Maier-Hein L, Alexander DC, Oxtoby NP. 痴呆领域图像分析与机器学习竞赛的十年发展。《神经影像》.2022;253:119083.[81] Yang S, Bornot JMS, Wong-Lin K, Prasad G. 用于预测轻度认知障碍和阿尔茨海默病的基于脑磁图/脑电图的生物标志物：机器学习视角下的综述。*IEEE Transactions on Biomedical Engineering.*2019;66(10):2924-2935.[82] Bucholc M, Titarenko S, Ding X, Canavan C, Chen T. 一种用于预测轻度认知障碍向痴呆转化的混合机器学习方法。*Expert Systems withApplications.* 2023:119541.[83] Lin RH, Wang CC, Tung CW. 一种利用基因生物标志物预测轻度认知障碍病情稳定患者的机器学习分类器。 *International Journal of Environmental Research and Public Health.*2022;19(8):4839.[84] Lee G, Nho K, Kang B, Sohn KA, Kim D. 利用多模态深度学习方法预测阿尔茨海默病进展。《科学报告》. 2019;9(1):1-2.[85] Ansart M, Epelbaum S, Bassignana G, et al. 利用机器学习预测轻度认知障碍进展：系统性、定量且批判性的综述。*Med Image*

Anal. 2020:101848.[86] Samper-Gonzalez J, Burgos N, Bottani S, Habert MO, Evgeniou T, Epelbaum S, Colliot O. 对预测临床和神经影像数据中阿尔茨海默病进展方法的可重复评估。载于《医学影像2019：图像处理》，2019年3月15日（第10949卷，第221-233页）。SPIE.[87] Kauppi K, Fan CC, McEvoy LK, Holland D, Tan CH, Chen CH, Andreassen OA, Desikan RS, Dale AM, 阿尔茨海默病神经影像学倡议。结合多基因风险评分、体积磁共振成像和认知测量指标，可改善对轻度认知障碍向阿尔茨海默病进展的预测。《神经科学前沿》。 2018;12:260.[88] Vonk JM, Greving JP, Gudnason V, Launer LJ, Geerlings MI. 一般人群的痴呆风险：在AGES-雷克雅未克研究中对预测模型进行大规模外部验证。《欧洲流行病学杂志》。2021;36(10):1025-41.[89] Wen J, Thibeau-Sutre E, Diaz-Melo M, et al. 用于阿尔茨海默病分类的卷积神经网络：概述及可重复评估。《医学影像分析》. 2020;63:101694.[90] Danso SO, Zeng Z, Muniz-Terrera G, Ritchie CW. 开发基于可解释机器学习的个性化痴呆风险预测模型：采用集成学习算法的迁移学习方法。《大数据前沿》. 2021;4:21.[91] Antila K, Lötjönen J, Thurfjell L, et al. PredictAD项目：开发用于阿尔茨海默病早期诊断的新型生物标志物和分析软件。《界面聚焦》. 2013;3(2):20120072.[92] Bruun M, Frederiksen KS, Rhodius-Meester HF, Baroni M, Gjerum L, Koikkalainen J, Urhemaa T, Tolonen A, Van Gils M, Rueckert D, Dyremose N. 临床决策支持工具对早期痴呆进展预测的影响：一项前瞻性验证研究。《阿尔茨海默病研究与治疗》。2019;11(1):1-1.[93] Gorse AD. 药物化学空间的多样性。《药物化学当前主题》。2006;6(1):3-18.

[94] David L, Arús-Pous J, Karlsson J, 等. 深度学习在工业制药研究中利用大规模异质化合物数据的应用。《药理学前沿》. 2019年;10:1303.[95] Lipinski CF, Maltarollo VG, Oliveira PR, da Silva AB, Honorio KM. 深度学习在药物设计与发现中的应用进展与展望。《机器人与人工智能前沿》. 2019年;6:108.[96] Rajkomar A, Dean J, Kohane I. 医学中的机器学习。《新英格兰医学杂志》. 2019年;380(14):1347-1358.[97] Gómez-Bombarelli R, Wei JN, Duvenaud D, 等. 使用基于数据的分子连续表征进行自动化学设计。《*ACS*中心科学》. 2018;4(2):268-276.[98] Vatansever S, Schlessinger A, Wacker D, 等. 人工智能和机器学习助力中枢神经系统疾病药物发现：现状与未来方向。《医学研究评论》. 2021;41(3):1427-1473.[99] Leung MK, Delong A, Alipanahi B, Frey BJ. 基因组医学中的机器学习：计算问题与数据集综述。《*IEEE*会刊》. 2015;104(1):176-197.[100] Dias TL, Schuch V, Beltrão-Braga PCB, 等. 通过网络医学方法为精神疾病和神经系统疾病重新定位药物。《转化精神病学》. 2020年;10(1):1-10.[101] Rodriguez S, Hug C, Todorov P, Moret N, Boswell SA, Evans K, Zhou G, Johnson NT, Hyman BT, Sorger PK, Albers MW. 机器学习发现阿尔茨海默病药物再利用的候选药物。《自然·通讯》. 2021;12(1):1-3.[102] Sundaram L, Gao H, Padigepati SR, 等. 利用深度神经网络预测人类突变的临床影响。《自然·遗传学》. 2018;50(8):1161-1170.[103] Dey KK, Van de Geijn B, Kim SS, Hormozdiari F, Kelley DR, Price AL. 评估深度学习注释对人类复杂疾病的信息价值。《自然·通讯》. 2020;11(1):1-9.[104] Breiman L. 统计建模：两种文化。《质量控制与应用统计》。

2003;48(1):81-82.[105] Navarro E, Udine E, de Paiva Lopes K, 等. 帕金森病人类髓系细胞中线粒体基因的转录特征存在差异。*bioRxiv*. 2020年.[106] Grubman A, Chew G, Ouyang JF, 等. 阿尔茨海默病患者内嗅皮层的单细胞图谱揭示了细胞类型特异性基因表达调控。*NatNeurosci*. 2019年;22(12):2087-2097.[107] Anderson NC, Chen P, Meganathan K, 等. 平衡偶然发现与可重复性：多能干细胞作为智力与发育障碍的实验系统。《干细胞报告》. 2021年.[108] Jiang J, Wang C, Qi R, Fu H, Ma Q. scREAD：阿尔茨海默病单细胞RNA测序数据库。*iScience*. 2020年;23(11):101769.[109] Lotfollahi M, Wolf FA, Theis FJ. scGen可预测单细胞扰动响应。《自然—方法》. 2019年;16(8):715-721.[110] Li H, Guan Y. 在单核苷酸分辨率下快速解析细胞类型特异性转录因子结合图谱。*Genome Res*. 2021;31(4):721-731.[111] Quang D, Xie X. FactorNet：一种用于根据核苷酸分辨率序列数据预测细胞类型特异性转录因子结合的深度学习框架。*Methods*. 2019;166:40-47.[112] Health TL. 指导更好地设计和报告人工智能干预试验。《柳叶刀·数字健康》.2020;2(10):e493.[113] Harrer S, Shah P, Antony B, Hu J. 人工智能助力临床试验设计。*TrendsPharmacol Sci*. 2019;40(8):577-591.[114] Ezzati A, Lipton RB, Alzheimer's Disease Neuroimaging Initiative. 机器学习预测模型可提高阿尔茨海默病临床试验的疗效。《阿尔茨海默病杂志》. 2020;74(1):55-63.[115] Reith FH, Mormino EC, Zaharchuk G. 利用机器学习预测痴呆患者未来的淀粉样蛋白生物标志物，以改进临床试验患者筛选。《阿尔茨海默病与痴呆》*:*

《转化研究与临床干预》。2021;7(1):e12212.[116] Hane CA, Nori VS, Crown WH, Sanghavi DM, Bleicher P. 使用临床记录和机器学习预测痴呆症发病：病例对照研究。《JMIR 医学信息学》。2020;8(6):e17819.[117] Kumar S, Oh I, Schindler S, Lai AM, Payne PRO, Gupta A. 使用临床数据对阿尔茨海默病痴呆的进展进行建模的机器学习方法：系统性文献综述。《JAMIA Open》。2021;4(3):ooab052.[118] Wilkinson MD, Dumontier M, Aalbersberg IJ, Appleton G, Axton M, Baak A, Blomberg N, Boiten JW, da Silva Santos LB, Bourne PE, Bouw
Mueller SG. Weiner MW. Thal LJ. 等. 阿尔茨海默病神经影像学倡议。《北美神经影像学临床》。2005;15(4):869.[120] Aroklasally P, Bloom D, Lee J, Feeney K, Ozolins M. 印度纵向衰老研究：于：《亚洲衰老：来自新兴数据倡议的发现》。美国国家科学院出版社；2012.[121] Beekly DL, Ramos EM, Lee WW, Deitrich WD, Jacka ME, Wu J,
Hubbard JL, Koepsell TD, Morris JC, Kukull WA. 美国国家阿尔茨海默病协调中心（NACC）
M, Sudlow C. 英国生物样本库。《大脑》。2015;138(12):3463-5.[123] Khoei TT, Ahajjam MA, Hu WC, Kaabouch N. 一种用于纵向研究中阿尔茨海默病检测的深度学习多任务方法。载于《2022年IEEE国际电子信息技术会议（eIT）》2022（第315-319页）。IEEE.[124] Benri moh D, Israel S, Fratila R, 等. 医疗保健中机器学习与人工智能的安全性、有效性和可解释性。《大数据前沿》。2021;4.[125] Bruckert S, Finzel B, Schmid U. 新一代医疗决策支持：迈向透明专家助手的路线图。《人工智能前沿》。2020;3:75.

[126] Molnar C. 可解释机器学习。Lulu.com；2020.[127] Gómez-Ramírez J, Ávila-Villanueva M, Fernández-Blázquez MÁ. 使用随机森林和基于置换的方法，选出用于预测转为轻度认知障碍的最重要自我评估特征。《科学报告》。2020;10(1):1-15.[128] Cole J, Wood D, Booth T. 将视觉注意作为痴呆神经影像分类的可解释模型：Doctor AI：让计算机解释其决策。《阿尔茨海默病与痴呆》。2020;16:e037351.[129] Dyrba M, Pallath AH, Marzban EN. 比较卷积神经网络可视化方法，以帮助解释用于检测阿尔茨海默病的模型。载于：《医学图像处理*2020*》。施普林格；2020:307-312.[130] Qiu S, Joshi PS, Miller MI, 等. 开发并验证一种用于阿尔茨海默病分类的可解释深度学习框架。《大脑》。2020;143(6):1920-1933.[131] Ribeiro MT, Singh S, Guestrin C. "为什么要相信你？"：解释任何分类器的预测。2016:1135-1144.[132] Shapley LS. *17. n*人博弈中的一个值。普林斯顿大学出版社；2016.[133] Spooner A, Chen E, Sowmya A, 等. 用于痴呆症预测的高维临床数据生存分析机器学习方法比较。《科学报告》。2020;10(1):1-10.[134] Bansal D, Chhikara R, Khanna K, Gupta P. 用于检测痴呆的各种机器学习算法的比较分析。《计算机科学进展》。2018;132:1497-1502.[135] Williams JA, Weakley A, Cook DJ, Schmitter-Edgecombe M. 用于轻度认知障碍与痴呆诊断鉴别的机器学习技术。2013.[136] Chen R, Herskovits EH. 用于构建极轻度痴呆诊断模型的机器学习技术。《神经影像》。2010;52(1):234-244.[137] Park SH, Han K. 评估人工智能技术在医学诊断与预测中的临床性能及影响的方法学指南。《放射学》。

2018;286(3) 800 809
: - .[138] Mårtensson G、Ferreira D、Granberg T 等。临床分布外 MRI 数据中深度学习模型的可靠性：一项多队列研究。《医学影像分析》。2020年；66：101714。[139] Klöppel S、Peter J、Ludl A 等。将自动化 MR 诊断方法应用于记忆门诊：一项前瞻性研究。《阿尔茨海默病杂志》。2015年；47(4)：939-954。[140] Jun GR、Chung J、Mez J 等。跨族群全基因组扫描发现新的阿尔茨海默病位点。《阿尔茨海默病与痴呆》。2017年；13(7)：727-738。[141] Kalafatis C、Modarres MH、Apostolou P 等。基于人工智能的 5 分钟计算机化认知评估在轻度认知障碍和阿尔茨海默型痴呆中的有效性与文化泛化能力。《精神病学前沿》。2021年；12：706695。[142] Clarke N、Foltz P、Garrard P。如何利用（数千个）词语开展工作：阿尔茨海默病话语分析的计算方法。《大脑皮层》。2020年；129：446-463。[143] Fraser KC、Fors KL、Kokkinakis D。用于评估轻度认知障碍患者叙事性言语的多语言词嵌入。计算机语音与语言。2019年；53：121-139。[144] Badhwar A、McFall GP、Sapkota S 等。阿尔茨海默病异质性的多组学方法：重点综述与路线图。《大脑》。2020年；143(5)：1315-1331。[145] Graham SA、Lee EE、Jeste DV 等。预测和检测老年人认知衰退的人工智能方法：概念性综述。精神病学研究。2020年；284：112732。[146] Chandler C、Foltz PW、Cohen AS 等。用于门诊应用的神经心理学测验的机器学习。《智能医学》。2020年；1：100006。[147] El-Sappagh S、Alonso JM、Islam SR、Sultan AM、Kwak KS。基于可解释人工智能的阿尔茨海默病多层多模态检测与预测模型。《科学报告》。2021年；11(1)：1-26。[148] Plana D、Shung DL、Grimshaw AA、Saraf A、Sung JJ、Kann BH。医疗保健领域机器学习干预的随机临床试验：系统综述。《*JAMA* 网络开放》。

2022年；5(9)：e2233946-。

[149] Nation DA, Ho JK, Dutt S, Han SD, Lai MHC；阿尔茨海默病神经影像学倡议。神经心理功能衰退在阿尔茨海默病生物标志物和轻度认知障碍诊断之外，进一步提升了痴呆预测能力。《阿尔茨海默病杂志》。2019年；69(4):1171-1182。[150] Bottino CMC, de Pádua AC, Smid J, Areza-Fegyveres R, Novaretti T, Bahia VS；巴西神经病学学会阿尔茨海默病与血管性痴呆工作组。痴呆与精神障碍的鉴别诊断：诊断标准和辅助检查。巴西神经病学学会认知神经病学与衰老科学部建议。《痴呆与神经心理学》。2011年；5(4):288-296。[151] Gianattasio KZ, Ciarleglio A, Power MC。为美国健康与退休研究中的种族/族裔差异研究制定算法性痴呆判定。《流行病学》。 2020年；31(1):126-133。[152] Gianattasio KZ, Bennett EE, Wei J, Mehrotra ML, Mosley T, Gottesman RF, Wong DF, Stuart EA, Griswold ME, Couper D, Glymour MM, Power MC；阿尔茨海默病神经影像学倡议。临床样本研究结果对社区样本的泛化能力：ADNI与社区动脉粥样硬化风险研究的比较。《阿尔茨海默病与痴呆》。2021年；17(8):1265-1276。[153] Mehmood A, Yang S, Feng Z, 等。利用迁移学习方法基于磁共振成像实现阿尔茨海默病早期诊断。《神经科学》。2021年；460:43-52。[154] Richens JG, Lee C M, Johri S。利用因果机器学习提高医学诊断的准确率。《自然·通讯》。2020年；11(1):1-9。[155] Luo W, Phung D, Tran T, Gupta S, Rana S, Karmakar C, Shilton A, Yearwood J, Dimitrova N, Ho TB, Venkatesh S, Berk M。生物医学研究中机器学习预测模型的开发与报告指南：多学科视角。《医学互联网研究杂志》。2016年；18(12):e323。