

数智健康国际动态

北京市卫生健康大数据与政策研究中心

2025. 6. 7

（五）数智医疗

在传统医疗中，医院临床诊疗基本依赖医生经验，“千人一方”是常态，而依据每个人的基因组学特征、环境暴露史及生活方式等因素定制预防、诊断与治疗方案的精准医疗愿景则被视为难以企及的理想。如今，人工智能（Artificial Intelligence, AI）的出现正在快速改变着这一切。AI 正凭借其卓越的数据解析能力、模式识别精度及预测效能系统性地破解传统医疗的困局，加速精准医疗愿景的临床转化，重塑医疗健康领域的范式。本期专题遴选两篇标志性研究以阐释 AI 驱动精准医疗的机制。

第一篇是“通过 AI 辅助诊断让肺癌识别更精准”的研究。晚期肺癌患者并发恶性胸腔积液（Malignant Pleural Effusion, MPE）的细胞病理学诊断具有重要临床意义，但反应性间皮细胞与肿瘤细胞的形态学相似性导致鉴别诊断困难。本研究致力于开发适用于临床场景的机器学习（Machine Learning, ML）辅助工具，旨在提高分析胸腔积液细胞样本的诊断准确性和效率，提供一种不需要深厚数据科学专业知识就能准确诊断的 AI 解决方案。研究者采用圣安德烈亚大学医院提供的 969 张病理图像数据集，包含经病理专家标注的间皮细胞（ $n=3130$ ）与腺癌细胞（ $n=3260$ ）样本。通过部署 YOLO v8 与 YOLO v11 实例分割模型进行细胞分类，获得稳健的模型性能（交并比 IoU=0.72），其中 YOLO v11 在多项指标上呈现显著优势。研究表明：ML 模型可有效提升细胞病理学鉴别诊断准确性。MPE 显著的形态学异质性要求高质量数据集支撑。ML 技术的整合将强化精准医疗时代的数据解读能力，为个体化治疗策略提供循证依据。

第二篇是“通过 AI 赋能遗传风险分析而驱动精准医疗”的研究。研究系统地论证了遗传风险因素（Genetic Risk Factors, GRF）与 AI 协同优化疾病预测、个体化治疗及健康管理的路径，推动精准医疗在疾病筛查、患者分层、药物研发与机制

解析领域的突破：在方法论上，构建了 GRF 整合前沿 AI 模型的技术路线，通过对 20 种主流算法进行系统评估，筛选出 15 种高效遗传风险评估工具，形成跨组学应用优化方案。在临床价值验证上，结合临床案例与可视化分析，证实 AI 增强的 GRF 分析可提升诊断精确度、优化靶向治疗并改善患者预后。在伦理与实施框架上，提出应对算法可解释性、技术可及性及伦理风险的综合解决方案，强调发展可解释人工智能（Explainable AI, XAI）与公平基因组技术的迫切性。最后研究指出，AI 与生物医学大数据的深度融合，正推动医疗模式向个体基因组导向的精准范式转型。

两项研究分别从诊断技术革新和风险因素解析的维度实证了 AI 是如何通过提升诊断敏感性、治疗靶向性与预防预见性，重塑精准医疗实践，为肿瘤患者提供定制化治疗方案，缩短罕见病诊断路径和实现慢性病动态管理。但研究也指出，技术转化并非一人一时之功，是需要所有参与者（研究者、医生、患者、政策制定者、技术开发者等）共同协作，以创建符合伦理规范的多源数据共享生态，发展可验证的临床决策算法和完善监管科学框架，才能得以真正实现。当跨学科协作突破数据壁垒与算法黑箱，精准医疗方能实现从理论范式向普惠实践的跨越，最终实现“一人一方”的个体化疾病治疗和健康管理的前景。

（徐健编辑）

译文一：

将 AI 带给临床医生：用友好模型简化胸腔积液细胞学诊断

Enrico Giarnieri, Elisabetta Carico, Stefania Scarpino, Alberto Ricci,
Pierdonato Bruno, Simone Scardapane, Daniele Giansanti, 徐健（译）

来源：Diagnostics (Basel).

时间：2025 年 5 月.

链接：<https://doi.org/10.3390/diagnostics15101240>.

1. 引言

1.1. 诊断性胸腔积液细胞病理学的挑战

恶性胸腔积液（MPE）是指以胸腔积液中存在肿瘤细胞为特征的副恶性疾病。MPE 通常与肺癌、乳腺癌和淋巴瘤等癌症的晚期有关。值得注意的是，壁层胸膜转移在肺癌中很常见，占有肺癌病例的 45–65%。这些晚期胸腔积液的发作通常与不良预后有关，中位生存时间（MST）为 4 个月或更短。因此，早期准确检测 MPE 对于确定最有效的治疗策略至关重要。

治疗策略和预后

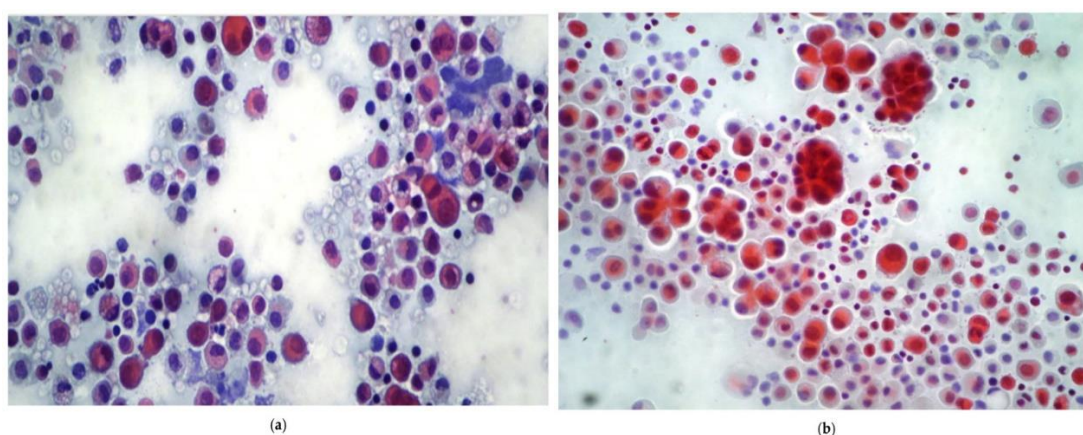
MPE 的主要治疗方法包括胸腔穿刺术，这是一种用于去除胸腔积液和缓解症状的手术。然而，胸腔穿刺术并不能提供长期解决方案，其有效性可能有限。虽然一些患者可能不会复发，但其他患者会迅速复发，有时在手术后一个月内。复发说明我们需要更积极的治疗选择。因此，快速准确地识别胸腔积液的来源对于定制个体化治疗计划至关重要，包括管理液体分泌和改善症状控制的干预措施。

诊断障碍和方法

MPE 的诊断通常从实验室检查开始，包括胸腔积液细胞学检查。尽管这种传统方法对于区分恶性细胞和反应性间皮细胞（常见于非肿瘤性积液）仍然至关重要，但当原发性肿瘤的来源不明确时，它可能具有挑战性。病理学家的专业知识在提高这

种诊断方法的可靠性方面起着关键作用。此外，与临床医生的合作为更准确地解读细胞学提供了必要的背景信息。

然而，胸腔积液的细胞病理学容易出现错误，包括假阳性和假阴性。这些不准确通常是由于积液中细胞的复杂形态造成的，这可能类似于恶性细胞。间皮细胞的反应性变化，如深染核、细胞质空泡化、双核化、多核化、吞噬活性和细胞内构型，可以模拟恶性细胞的外观，尤其是腺癌或间皮瘤中发现的恶性细胞。这些形态学上的相似性构成了假阳性诊断的风险（图 1）。这一挑战突出了补充诊断方法的必要性，以补充传统的形态学分析。



(a) 反应性胸腔积液显示间皮细胞表现出退行性和反应性特征（巴氏染色， $\times 10$ ）。

(b) 腺癌胸腔积液显示间皮细胞和肿瘤细胞共存，以单个和簇聚集体组织（巴氏染色， $\times 10$ ）。

肿瘤簇很容易识别，而单个肿瘤细胞可能类似于间皮细胞。

人工智能的作用越来越大

近年来，人工智能（AI）因其在提高细胞病理学诊断准确性方面的潜力而受到关注。机器学习（ML）和深度学习（DL）算法，尤其是卷积神经网络（CNN），在分析医学影像方面已显示出前景。

机器学习通过提高诊断准确性和治疗个性化来改变医疗保健。它支持分析庞大的医疗数据集，以发现超出人类能力的模式和见解。

通过支持临床决策，它有助于减少错误并改善患者预后。ML 驱动的工具可以更早地发现疾病并提出最有效的干预措施。预测模型有助于在症状出现之前识别高危个体。通过 ML 实现自动化还可以简化工作流程，节省时间和资源。

它有助于实时监测和适应性护理，尤其是在慢性病管理方面。最终，机器学习有助于建立更主动、更精确、更以患者为中心的医疗保健系统。

AI 模型的诊断准确性与人类专家相当，甚至超过人类专家。最初，这些 AI 模型被应用于妇科细胞学，例如巴氏涂片检查，产生了有希望的结果。然而，AI 的应用已经扩展到妇科细胞学之外，包括非妇科样本，例如来自泌尿道、乳腺、甲状腺、肺和胸膜转移性细胞病理学的样本。与此同时，文献正在探索生殖健康和生育问题的遗传基础。这些研究旨在识别和表征使个体易患综合征和非综合征型不孕症的基因，从而有助于更好地了解生殖障碍的分子机制。这些见解不仅对诊断目的很有价值，而且对生殖保健中有针对性的治疗策略和个性化医学方法的开发也很有价值。另外，尽管 AI 在细胞病理学的各个领域应用越来越多，但专门针对胸腔积液细胞学的研究仍然有限。

一些研究应用 AI 模型来区分胸腔积液中的正常细胞和恶性细胞。这些研究使用不同的深度学习架构，专注于来自各种肿瘤类型的细胞，包括肺癌、胃癌和乳腺癌。结果表明，深度学习模型可以提高 MPE 诊断的准确性和效率，为临床实践提供新的可能性。

文献和 AI 在胸腔积液诊断中的概述

表 1. 肿瘤性和非肿瘤性胸腔积液细胞学的深度学习文献概述

参考	胸腔积液数据集	细胞检测方法	指标
Xieetal.	肺腺癌	ResNet-18 系列	AUC0.67 - 0.83
Park 等人	乳腺癌	初始镜像 V2	灵敏度 95%、特异性 98.6%、准确度 81.1%
Wang 等人	恶性/非恶性细胞	ViT 视觉变压器、ResNet-50、Vgg-16、Fundus-DeepNet	精度 96.8、精度 87.3、精度 83.9、精度 88.7
任 etal.	肺腺癌、乳腺癌、胃癌	ResNet-2	AUC0.807、AUC0.737、AUC0.955
Chen 等人	肺腺癌	盗梦空间 v3YOLO v4	mAP20% 准确率 98%
Baykal 等人	胸腔积液、核素检测	使用 ResNet-101 实现更快的 R-CNN	准确率 99.34%、召回率 96.93%
Mavropoulos 等人	乳腺癌	初始 v3	AUC0.96

表 1 总结了有关基于深度学习的模型的文献，这些模型旨在区分细胞病理学中的肿瘤性和非肿瘤性胸腔积液。这些研究表明，使用 CNN 等 AI 技术来提高诊断准确性，特别是在区分恶性细胞和良性细胞方面。这些研究的成功凸显了人工智能帮助

病理学家更准确、更有效地诊断 MPE 的潜力。

1.2. 细胞病理学中的人工智能：通过细胞和细胞核分割提高诊断准确性

随着细胞病理学诊断方法的不断发展，特别是随着人工智能（AI）的整合，显微图像中的细胞和细胞核分割已成为分析过程的关键组成部分。这涉及注释、计数和跟踪细胞等基本任务，这些任务对于精确诊断至关重要，尤其是在胸腔积液的情况下。虽然前面的部分解决了识别恶性胸腔积液的挑战，但人工智能驱动图像分析的最新发展现在正在促进更准确和自动地分割胸腔积液样本中的细胞，进一步提高诊断可靠性。

许多研究引入了细胞分割的方法，一般可分为语义分割和实例分割两大类。语义分割涉及将图像像素分配给特定的对象或感兴趣的区域，例如细胞或细胞核，从而有效地将这些元素与周围的背景分离。一些用于语义分割的著名深度学习模型包括全卷积网络（FCN）和 U-Net，它们已被证明在分割显微图像中的细胞结构方面取得了成功。

另一方面，实例分割通过区分同一类中的单个对象，例如区分同一图像中的多个细胞或细胞核，从而进一步进行分割。此方法生成边界框和相应的掩码，从而在一个步骤中同时实现检测和分割。一个著名的实例分割模型是由 Facebook AI Research（FAIR）开发的 Mask-RCNN，它已成为在复杂图像中分割对象的领先算法之一。Mask-RCNN 的各种迭代已经被创建，以增强其在生物成像中细胞和细胞核分割的性能。

最近，一种名为 Segment Anything（SAM）的基于 transformer 的模型在图像分割任务中取得了巨大的成功，为包括细胞成像在内的众多应用树立了新的性能标准。SAM 的有效性导致了一种混合方法 Cells SAM 的开发，该方法将 SAM 与另一种基于 transformer 的对象检测模型 Cell Finder 相结合。这种混合方法为细胞图像分割产生了更好的结果，为自动分析具有不同细胞特征的显微图像提供了一种创新的解决方案。

YOLO：显微图像分割的实时解决方案

人工智能领域的一个重要发展也是由 You Only Look Once (YOLO) 推动的，这是一系列最初开发的算法，例如自然图像中的分割 (<https://medium.com/tag/yolo-object-detection>，于 2025 年 4 月 5 日访问; <https://pjreddie.com/darknet/yolo/>，于 2025 年 4 月 5 日访问)。由于其速度和准确性，YOLO 越来越多地用于医学图像分析 ([https://github.com/sevdaimany/YOLO v8-Medical-Imaging](https://github.com/sevdaimany/YOLO-v8-Medical-Imaging)，于 2025 年 4 月 5 日访问)。与 R-CNN 等传统模型相比，YOLO 的主要优势之一是它能够在一次通过中检测多个对象。这使得它特别适用于需要快速高效分析大型数据集的实时应用程序。与多阶段模型不同，YOLO 在一个统一的步骤中执行检测，使其对于实时决策场景特别有价值，例如外科手术或紧急诊断。

YOLO 对象检测能力的速度和准确性使其成为分析胸腔积液样本的有前途的应用，这是本研究的重点。通过快速分割细胞并区分恶性细胞和良性细胞，YOLO 算法可以显著改善诊断工作流程，为临床医生提供更快、更可靠的结果。此外，YOLO 处理各种图像类型和复杂性的能力确保了其对不同临床应用的适应性，包括在显微成像中分割细胞。

1.3. 使用 YOLO 推进医疗诊断：医疗保健中 AI 驱动的对象检测

YOLO (You Only Look Once) 是一种对象检测算法，它将输入图像划分为网格系统，以识别和分类图像中的对象。其结构包括三个主要部分：(i) 脊椎，(ii) 颈部和 (iii) 头部。

1. Backbone：主干负责特征提取。它处理输入图像并生成特征图，这些特征图是图像中重要图案和对象的表示。从本质上讲，主干充当算法的“眼睛”，捕获视觉数据中的关键元素。

2. 脖子：脖子充当脊椎和头部之间的中介。它聚合了主干网提取的特征，并通过特征金字塔网络 (FPN) 等网络对其进行优化。该组件对于提高目标检测的分辨率和准确性至关重要，尤其是在具有复杂目标结构的情况下。

3. 头部：头部负责生成最终预测，例如边界框（包含检测到的对象的矩形区域）的坐标、对象性分数（表示区域包含对象的可能性）和分类分数（识别检测到的对象类型）。

YOLO 于 2015 年随着 YOLO v1 的发布而首次推出，并因其执行实时对象检测的能力而迅速受到关注。多年来，该算法经历了多次更新，每个新版本都提供了更高的准确性、速度和多功能性。这些更新在 Ultralytics 于 2024 年发布的 YOLO v11 中达到高潮，它代表了 YOLO 算法的最新版本。

最重要的版本之一是 YOLO v8，它有五个缩放版本以满足不同的性能需求：YOLO v8n（纳米）、YOLO v8s（小）、YOLO v8m（中）、YOLO v8l（大）和 YOLO v8x（超大型）。这些变体支持一系列视觉任务，包括：

- 对象检测：识别和定位图像中的对象。
- 分割：将图像划分为多个段以进行更详细的分析。
- 姿态估计：确定物体的方向和位置，尤其是人物。
- 跟踪：监控视频数据中对象跨帧的移动。
- 分类：将对象分类为预定义的类别（例如，汽车、狗、人）。

YOLO 在医疗中的应用

表 2. YOLO 模型在细胞病理学中的应用总结

参考	YOLO 模型	细胞病理学应用
Chen 等人	YOLO v4	胸腔积液细胞学检查
Nambu 等人	YOLO v4	宫颈细胞学检查
Wu 等人	YOLO v5	支气管肺泡灌洗细胞
Shietal	YOLO v4	宫颈细胞学检查
Tarimo 等人	YOLO v5	血细胞
Wijaya 等人	YOLO v8	宫颈细胞学检查
Wang 等人	YOLO v7 版本	肺癌细胞学检查（活检）
Terasaki 等人	YOLO v5	子宫内膜细胞学
Rumpf 等人	YOLO v7 版本	支气管肺泡灌洗细胞
Awad 等人	YOLO v8-YOLO v11	血细胞（ALL）

YOLO 的主要优势之一是其多功能性，使其高度适用于图像分析至关重要的医疗保健等领域。在医学领域，YOLO 用于检测和定位诊断图像中的异常。例如，在细胞病理学中，检测异常细胞对于准确诊断至关重要，YOLO 的对象检测功能可以快速可靠地识别组织样本中的癌细胞。这种能力可以显著提高诊断的速度和准确性，尤其

是在需要快速分析大量医学图像的情况下。

表 2 总结了 YOLO 在细胞病理学中的应用，展示了使用不同版本 YOLO 的研究，包括最新的 YOLO v8 和 YOLO v11 模型。YOLO v11 在计算机视觉方面提供了进一步的进步。与 YOLO v8 一样，它支持对象检测、实例分割、分类、关键点检测和定向边界框（OBB）检测等任务。YOLO v11 在准确性和处理速度方面的改进使其在需要高精度和快速分析的应用中特别有用，例如医疗诊断。这些进步使 YOLO v11 成为需要快速、准确结果的实时应用的理想选择，例如检测医学图像中的异常或识别 X 光片中的肿瘤。

YOLO v11 和 AI 在医疗保健领域的未来

尽管相对较新，但 YOLO v11 已经在医学领域显示出有前途的应用，尽管专门针对细胞病理学的研究数量仍然有限。YOLO 设计的持续发展，尤其是 YOLO v11，正在为更高效、更准确的医疗保健 AI 驱动解决方案铺平道路。在细胞病理学中，图像分析的速度和精度至关重要，这些进步可以产生重大影响，确保更快、更准确的诊断，并最终改善患者的预后。

通过增强对象检测能力和实现快速处理，YOLO 及其后续版本（如 YOLO v11）正在改变 AI 在医疗诊断中的格局，使其在临床环境中更易于访问、更高效和有效。

1.4. 扩大人工智能的可及性：为专业人士提供可访问的 AI 工具以进行实际应用

在使用人工智能的每个部门，越来越需要让人们更广泛地参与其开发，而不是将参与局限于被动使用应用程序。这代表着一个真实而具体的进步机会。为这个方向做出贡献的平台是必不可少的。在应用 AI 的每个领域，都越来越需要促进积极参与其开发，而不是将专业人员局限于被动使用现有应用程序。

采用 AI 的主要障碍之一，尤其是在医疗保健和诊断领域，是训练和部署模型的技术复杂性。例如，Roboflow 提供了一个直观的平台来构建和应用计算机视觉模型，即使没有编程或数据科学专业知识的用户也更容易使用 AI。此功能在医疗保健等领域特别有价值，在这些领域中，医生、生物学家和实验室技术人员等可能缺乏技术背景的专业人员仍然可以在图像检测和分类等任务中受益于 AI 支持。

该平台允许用户在视觉数据集上上传、注释和训练模型，极大地简化了实施过程。这减少了开发时间和精力，并允许用户专注于临床解释而不是算法细节。

例如，在数字细胞学中，在此平台上训练的 YOLO (You Only Look Once) 等对象检测模型可以准确识别组织样本中的恶性细胞，从而支持更快、更准确的诊断。将此类工具集成到日常工作流程中可以提高效率并降低误差幅度。

1.5. 研究和目标的基本原理

本研究旨在探索将人工智能 (AI) 整合到胸腔积液的诊断过程中，胸腔积液是一种液体在胸腔中积聚的疾病。具体来说，它侧重于使用 YOLO 等高级对象检测算法，尤其是其最新版本 YOLO v8 和 YOLO v11，以提高分析胸腔积液细胞样本的诊断准确性和效率。目的是研究 AI 如何支持细胞学家检测和分类胸腔积液样本中的异常，提供一种不需要深厚数据科学专业知识的解决方案，从而使所有细胞病理学家都可以使用 AI。

面向细胞病理学家的 AI：使其易于访问

这项研究背后的主要动机是弥合复杂的 AI 技术与处理胸腔积液样本的细胞病理学家的需求之间的差距。细胞学家虽然是各自领域的专家，但往往缺乏有效使用高级 AI 模型所需的专业知识。YOLO 等 AI 应用程序旨在用户友好，使细胞学家无需深入的技术知识即可利用 AI。这使专业人员能够利用 AI 驱动的模式进行图像分析来增强诊断流程，从而提高准确性和工作流程效率。

将 AI 整合到胸腔积液诊断中特别有益，因为它可以帮助细胞学家快速可靠地区分不同的细胞类型，例如恶性和良性细胞。这种支持可以在减少人为错误、加快诊断和改善患者预后方面发挥关键作用。在胸腔积液分析中采用 AI 最终使先进技术在医疗诊断中的使用大众化，为所有细胞学家提供了有价值的应用。

YOLO：关键创新

YOLO 代表 “You Only Look Once”，是一种对象检测算法，旨在快速准确地处理图像。多年来，YOLO 不断更新，最新版本 YOLO v8 和 YOLO v11 在速度、精度和召回率方面都有了显著的改进，使其非常适合细胞学等医疗应用。

YOLO v11 引入了优化功能，可提高对象检测、分类和分割任务的准确性和速度。

这些功能在细胞病理学中特别有价值，因为精确度是正确识别胸腔积液样本中细胞类型和异常的关键。该模型可以帮助细胞学家自动化识别和分类恶性细胞的过程，从而提高诊断准确性。

在本研究中，我们重点评估 YOLO v8 和 YOLO v11 在检测和分类胸腔积液样本细胞方面的性能。结果表明，两种模型都表现良好，YOLO v11 在精度和平均精度均值（mAP）方面略优于 YOLO v8。这表明 YOLO v11 可以在胸腔积液样本分析中发挥强大的作用，有可能帮助细胞学家更有效地区分恶性和良性细胞。

本研究的目的

本研究的主要目的是评估基于 YOLO 的 AI 模型在分析胸腔积液样本方面的有效性。通过评估最新版本的 YOLO，特别是 YOLO v8 和 YOLO v11，本研究旨在确定这些模型如何增强细胞病理学家的诊断能力，提高胸腔积液分析的准确性和效率。该研究强调了如何将 AI（特别是 YOLO）用作应用程序，以支持细胞学家识别和分类细胞，而无需机器学习方面的专业知识。

这项研究通过展示如何将 AI 应用于胸腔积液分析的特定环境，为医学诊断中更广泛的 AI 领域做出了贡献。它探讨了使用 AI 模型检测恶性细胞和良性细胞之间的细微差异的好处，从而有可能减少诊断错误并有助于早期发现严重疾病。

总体而言，本研究背后的基本原理是探索 AI 在胸腔积液诊断中的整合，重点是基于 YOLO 的模型的使用。这些模型可以帮助细胞学家更精确地分析胸腔积液样本，使 AI 成为改进诊断工作流程的宝贵应用。该研究强调了 AI 支持细胞学家检测胸腔积液样本异常的潜力，最终提高诊断的准确性和效率，改善患者护理，并使所有细胞病理学家都可以使用 AI。

2. 材料和方法

这项研究利用由 YOLO 系列深度学习模型提供支持的先进计算机视觉技术来分割和分析胸腔积液图像中的细胞。主要目标是开发一种强大的方法来区分胸腔积液中的恶性细胞和良性细胞，特别关注肺腺癌和间皮细胞。

YOLO 架构以其高效的对象检测而闻名，特别适合这项任务，因为它可以执行实时分割，为检测微观图像中的细胞结构提供速度和精度。该方法可确保准确识别和分割细胞和细胞核，这对于在细胞病理学中诊断胸腔积液样本至关重要。下面，我

们描述了图像采集、数据集准备的过程，以及在模型训练和评估中采取的具体步骤。

2. 1. 图像收集、数据集准备和细胞形态学标准

这项初步研究回顾性分析了 2022 年 1 月至 2023 年 12 月期间在罗马圣安德烈医院收集的良性和肺腺癌胸腔积液病例。该数据集包括 78 例胸腔积液病例，分为 40 例肺腺癌胸腔积液和 38 例良性胸腔积液。载玻片由高级细胞病理学家（EG、EC）选择，诊断基于细胞学特征。免疫组织化学检测（包括钙调蛋白、TTF-1 和 NapsinA 抗体）用于确认或排除肺腺癌来源的恶性胸腔积液。排除临床随访信息，所有载玻片均匿名化评估。使用标准方法制备样品，并使用 Papanicolaou 染色。使用安装在 BX45 显微镜（奥林巴斯，东京，日本）上的数码相机 DP27（奥林巴斯）捕获每张载玻片多达 100 张图像。放大倍率设置为×200，手动调整焦距，并优化焦深以捕获每个细胞的整个区域。捕获后应用颜色归一化以保持一致性。由于制备缺陷、照片模糊、细胞重叠、假状结构、细胞质过多或细胞簇杂乱无章等问题，一些图像被排除在外。为研究选择的间皮细胞表现出核增大、高核质比、突出的核仁和双/多核等特征。腺癌细胞表现为孤立的非典型细胞，细胞核增大、圆核、双核或多核，细胞质致密，偶尔含有空泡化。

2. 2. 注释、数据增强、模型训练和指标评估

表 3 概述了数据集，包括每个类的图像参数和注释。根据收集到的图像，构建了一个计算机视觉数据集，由 969 张手动注释图像组成，显示间皮细胞（3130）和腺癌细胞（3260）（图 2）。注释后，图像的大小调整为 640×640 像素。

表 3. 数据集分析和每个类的注释数量概述

图片	注释	平均大小	图像中位数比率
969	6390	3. 92mp	2288×1712
所有分裂	训练集	验证集	测试集
adk 细胞 3260	adk2573	adk345	adk342
ms 细胞 3130	ms2468	ms309	ms353

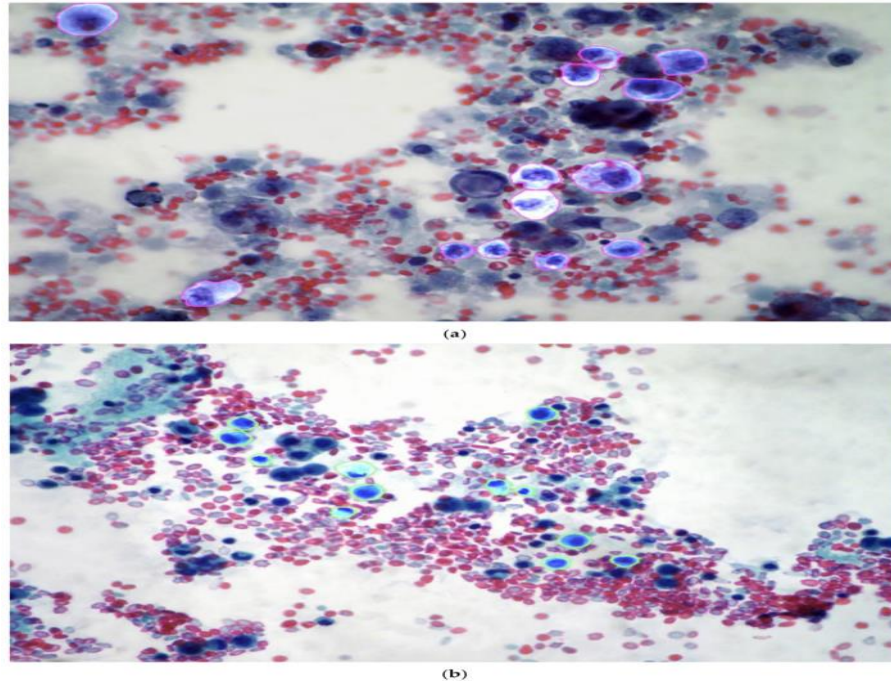


图 2 显示反应性/正常和肿瘤性胸腔积液细胞的不同病例分为两类并手动注释

(a) 具有紫色注释的单个腺癌细胞 (adk) (巴氏染色)

(b) 具有绿色注释 (巴氏染色) 的单个间皮细胞 (ms)

数据集分为 80%训练、10%验证和 10%测试子集，遵循生物医学图像分类任务中普遍采用的约定。这种拆分通过单独的验证和测试集平衡了对足够训练数据的需求与可靠的评估。80-10-10 的比率已在多项研究中被证明可产生卓越的性能。例如，Golchubian 等人证明，在用于照片质量分类的深度学习模型中，与 60-20-20 配置相比，这种分割实现了更高的准确率 (81.6%)。同样，在细胞学图像分析中，Boussel 等人采用相同的比率来确保巴氏涂片细胞图像分类中所有亚群的类别分布均衡。鉴于我们研究的试点性质和有限的数据集大小，80-10-10 的拆分被认为是支持稳健模型开发和评估的适当且方法合理的。

Roboflow 技术用于促进将数据集导出到 Google Colab，从而实现模型性能的自定义训练和评估。采用数据增强技术来提高模型性能并减少过拟合。图像在水平和垂直方向上翻转，以提高模型识别对称物体的能力。随机裁剪使模型能够识别部分可见的物体，并应用 90 度旋转以促进识别不同方向的细胞。训练后，计算平均精度 (mAP) 和精度-召回率 (PR) 曲线等性能指标以评估模型的疗效。mAP 用于衡量不同阈值下的模型性能，它同时考虑精度 (预测的正确性) 和召回率 (模型识别所有相关实例的能力)。在数据增强之后，使用 YOLO v8 和 YOLO v11 模型执行实例分割，

并在具有 NVIDIA 驱动程序版本 535.104.05 和 CUDA 12.2 的 Tesla T4 GPU 上进行训练。

3. 结果

在本节中，我们介绍了两个深度学习模型 YOLO v8 和 YOLO v11 的评估结果，这些模型应用于胸腔积液细胞的分类任务。本研究的目的是评估模型区分胸腔积液样本中恶性腺癌细胞（adk）和良性间皮细胞（ms）的能力。此处显示的结果包括基于一系列指标（包括精度、召回率、F1 分数和平均精度均值（mAP））对两种模型的性能进行详细比较。

我们还通过混淆矩阵检查模型的性能，该矩阵提供了对每种细胞类型的分类准确性的见解。此外，我们使用来自体腔液的恶性和良性细胞的公共数据集进行外部验证。这种额外的验证用于评估模型在应用于新的、看不见的数据时的泛化能力。

以下部分介绍了 YOLO v8 和 YOLO v11 的定量性能细分，以及模型在训练和外部数据集上的预测的直观示例。通过评估两种模型的优缺点，我们旨在确定哪种模型提供了最可靠的胸腔积液细胞分类，并探索这些模型在临床诊断中的潜在应用。

3.1. 模型训练结果：YOLO v8 vs. YOLO v11

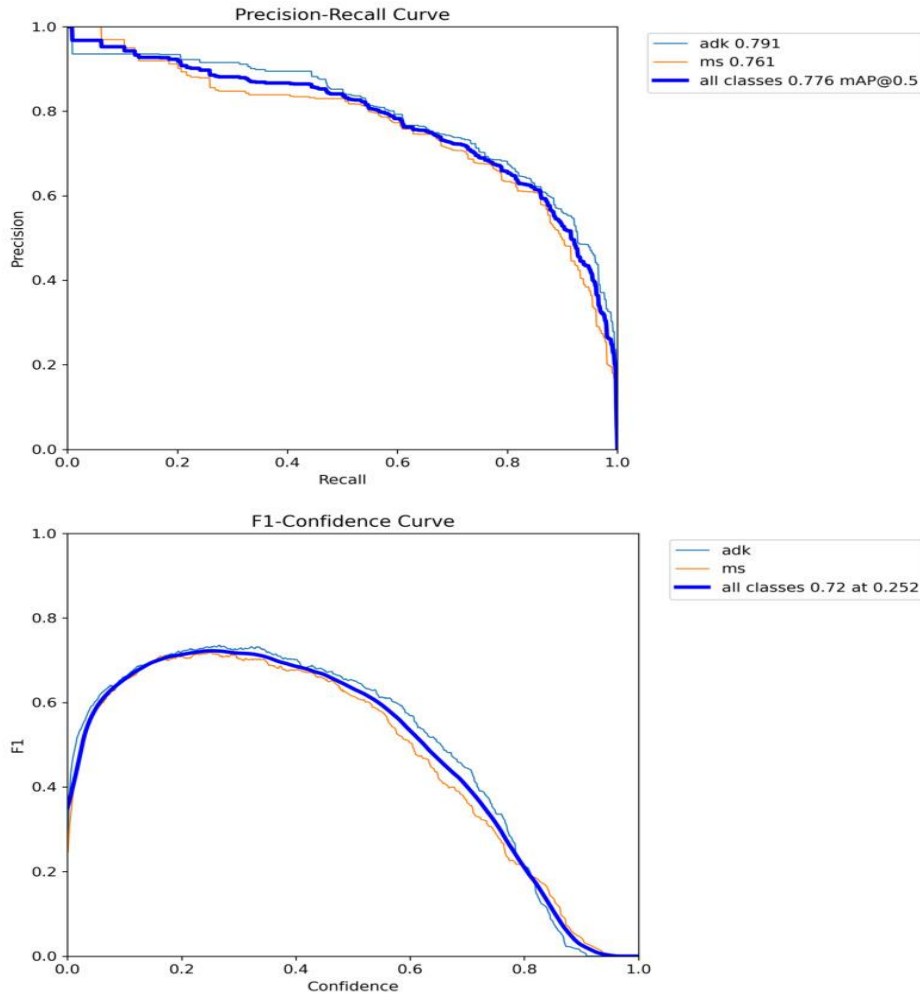
胸腔积液细胞评估数据集的训练结果揭示了 YOLO v8 和 YOLO v11 的关键性能指标，如表 4 所示。这些结果突出了两种模型在两种细胞类别中的精确率、召回率和整体分类性能：间皮细胞（ms）和腺癌细胞（adk）。

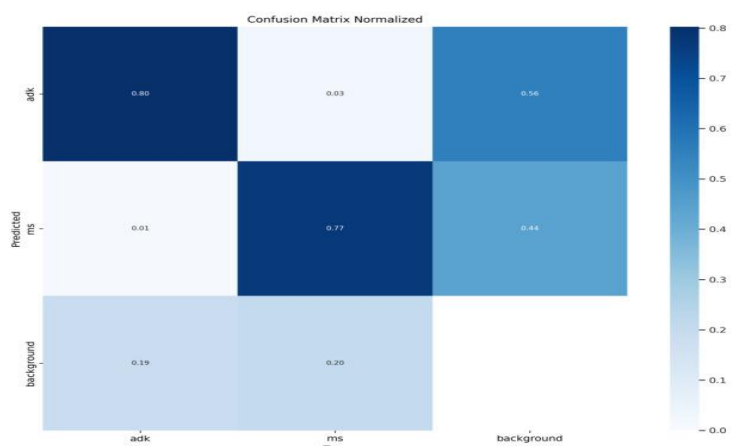
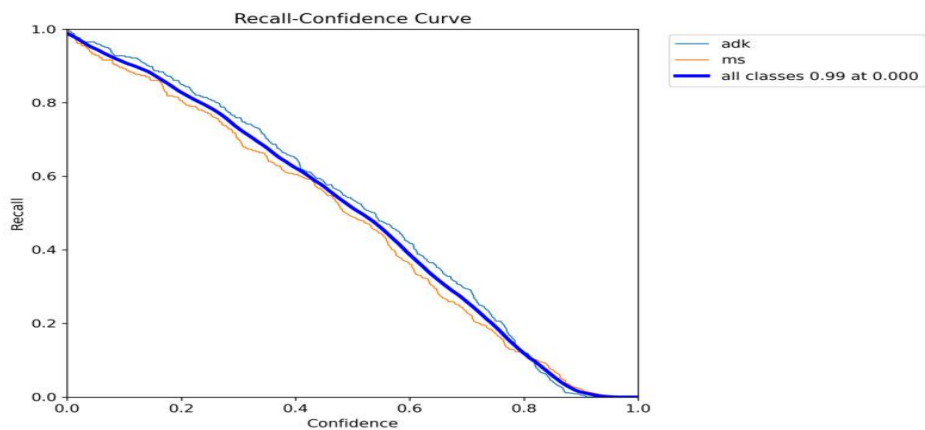
从胸腔积液细胞评估数据集获得的训练结果显示了按类别分类的值和指标（表 4）。YOLO v8 对所有细胞的准确度为 0.673，其中间皮细胞（ms）为 0.672，腺癌细胞（adk）为 0.674，而总体召回值为 0.784，ms 为 0.806，adk 为 0.761。通过平均精密均值（mAP）测量的总体性能为 0.777。根据归一化混淆矩阵，adk 的分类准确率达到 0.80，ms 的分类准确率达到 0.77，表明模型在区分两种细胞类型方面的可靠性能。

表 4. 所有类的 YOLO v8 和 YOLO v11 指标的比较

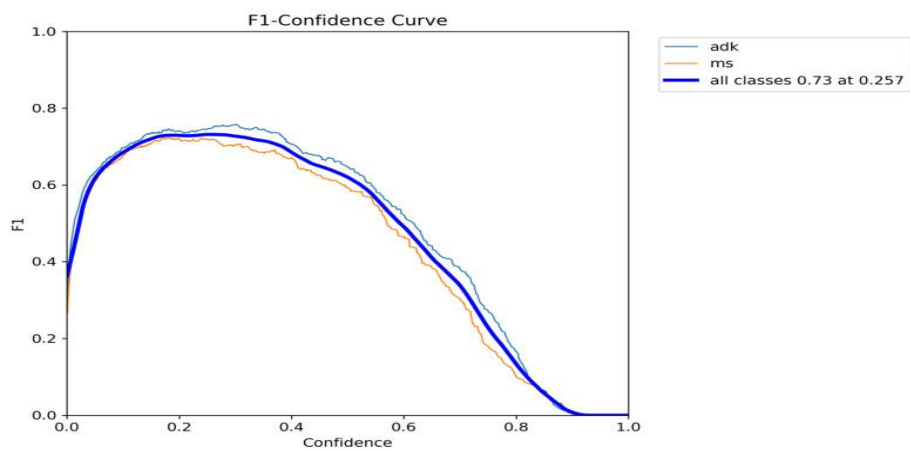
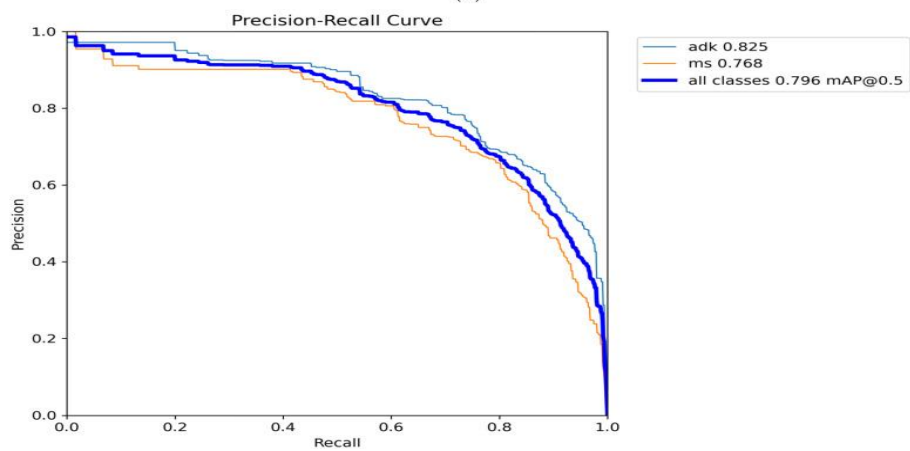
模型	F1 分数	精度	召回	mAP50 抗体	mAP50-95
YOLO v8	0.72	0.673	0.784	0.777	0.662
YOLO v11 版本	0.72	0.701	0.756	0.796	0.681

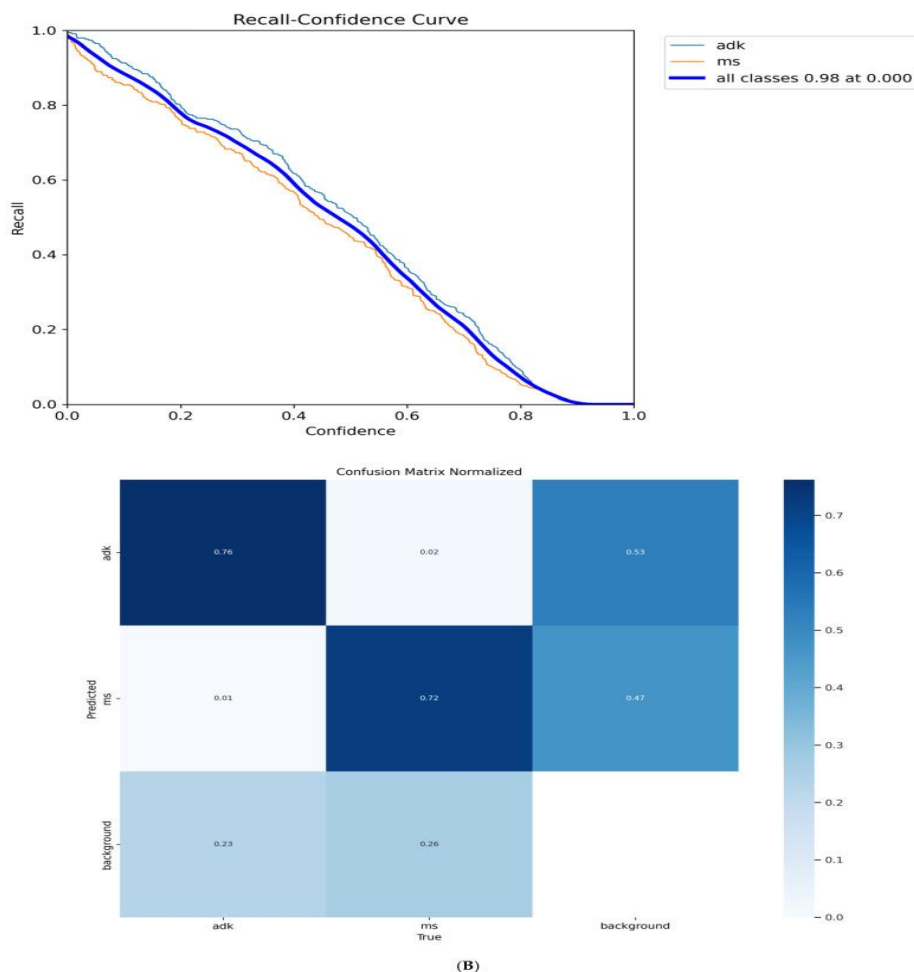
与 YOLO v8 相比, YOLO v11 在精度、mAP50 和 mAP50-95 方面取得了略好的结果, 反映了不同阈值下平均分类准确性的提高。然而, 它在特定类的准确性方面表现略差, 如混淆矩阵所示 (图 3A、B), 表明精度和整体类区分之间可能存在权衡。





(A)





(A) YOLO v8 性能指标。(B) YOLO v11 性能指标。

总体而言，考虑到 YOLO v11 在多个关键指标上的一致表现，YOLO v11 成为更有利的模型。虽然两个模型都获得了相同的 F1 分数，但 YOLO v11 表现出更高的总体精度(0.701 对 0.673)，表明它能够更好地减少跨类的误报。此外，它在 mAP50(0.796 对 0.777)和 mAP50-95 (0.681 对 0.662) 方面的表现都优于 YOLO v8，这不仅反映了它在标准阈值下的卓越检测精度，还反映了它在不同水平 IoU 中的稳健性。这些增强功能在临床环境中尤其重要，因为在临床环境中，精确性和可靠性对于支持诊断决策至关重要。尽管 YOLO v11 在混淆矩阵中的类特异性准确性略低，但其整体性能概况（尤其是在更好地捕获泛化和模型质量的指标方面）表明它可能更适合在 AI 辅助细胞学分析中实际实施。

图 3A, B 提供了模型性能指标的直观比较。精度-召回率 (PR) 曲线为评估模型性能提供了一个有价值的工具，尤其是在具有类不平衡的数据集中，因为它们侧重于精度和召回率之间的权衡。这在医疗应用中尤其重要，因为在医疗应用中，检测

真阳性（召回率）而不产生过多的假阳性（精确率）至关重要。

这些图显示了 YOLO v8 和 YOLO v11 模型在间皮（MS）和腺癌（ADK）细胞分类方面的性能，以及它们各自的性能曲线。深蓝色粗线表示所有类别的平均性能，而绿色和橙色线分别代表 ADK 和 MS 细胞。这两个模型都表现出可靠的区分能力，精度-召回率（PR）曲线往往接近图的右上角，这是精确率和召回率之间良好平衡的典型指标，尤其是在类分布和标签歧义会影响性能的数据集中。

F1 置信度曲线提供了对一系列置信度阈值中准确率和召回率之间相互作用的进一步见解。这些曲线有助于可视化模型的预测在分类阈值变化时的行为，尤其是在复杂情况下，例如涉及重叠细胞边界、细微形态特征或异质染色模式的场景。F1 分数在特定阈值处的下降可能反映了现实世界的挑战，例如类不平衡、模棱两可的细胞特征、不一致的注释或可能存在可能混淆分类的炎症或坏死细胞。细胞学样本中存在此类元素是很常见的，这强调了不仅在固定阈值下而且在整个置信范围内评估性能的重要性。

召回率-置信度曲线突出了每个模型对阳性情况的敏感性，尤其是在较低确定性水平下。在较低阈值下较高的召回率可以表明模型标记潜在相关检测的能力，这在通常优先考虑敏感性的临床筛查工作流程中可能很有用。在这种情况下，其中一个模型 YOLO v8 显示，与另一个模型相比，腺癌细胞的召回率略高，这表明在检测恶性特征方面具有潜在优势，尤其是在具有挑战性或模棱两可的情况下。

作为对这些性能曲线的补充，混淆矩阵提供了对模型分类行为的更精细视图。它包括三个类别-腺癌（ADK）、间皮（MS）和背景-并已标准化以反映不同类别的比例表现。这种归一化考虑了数据集中每个类的不同频率，从而可以更公平地评估模型的稳健性。该矩阵显示，两种模型都能够以令人满意的准确性区分目标细胞类型，即使在具有挑战性的条件下也能保持较低的误分类率。这些结果支持了模型在实际细胞学分析中的适用性，其中敏感性和特异性之间的平衡权衡至关重要。

3.2. 使用外部数据集进行评估和验证

为了进一步验证模型的泛化能力，使用包含来自体腔液的恶性和良性细胞图像的公共数据集进行了外部验证。该验证旨在测试模型在面对新的、看不见的数据时准确分类细胞的能力。

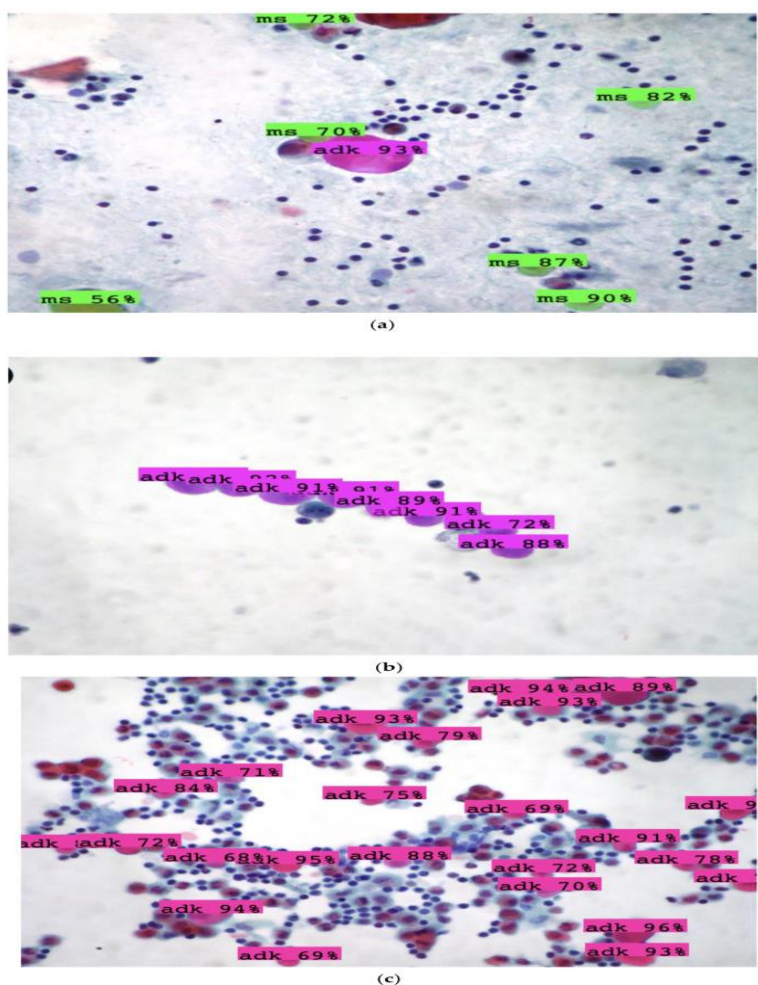


图 4 从测试集中随机选择图像以 50%置信度阈值验证两类细胞的预测性能

(a) 间皮细胞 (ms) 中的分离腺癌细胞 (adk)。(b) 腺癌细胞簇。(c) 在显着炎症的情况下分离的腺癌细胞。

- (a) 间皮 (ms) 细胞中的分离腺癌 (adk) 细胞。
- (b) 一簇腺癌细胞。
- (c) 明显炎症区域内的腺癌细胞。

利用从训练数据集和外部数据集中随机选择的图像来评估模型在区分恶性腺癌细胞 (adk) 和良性间皮细胞 (ms) 方面的性能。这些图像带来了额外的挑战，例如重叠的单元格和较差的图像质量，但这些模型仍然能够可靠地执行。图 4 展示了随机选择的测试图像，用于在 50%置信度阈值下评估模型的性能。这些图像专门评估区分腺癌 (adk) 和间皮 (ms) 细胞的能力，包括图 4。

这些图像突出了模型面临的挑战，例如在炎症条件下识别腺癌细胞。然而，它们也表明，即使在这些复杂的情况下，这两种模型也能够有效地识别和分类这些细胞。

图 5 显示了外部数据集的结果，其中包括来自体腔液的恶性和良性细胞。外部验证证实，YOLO v8 和 YOLO v11 都能够准确地对细胞进行分类，即使面对未包含在训练数据集中的新图像也是如此。验证结果包括图 5。

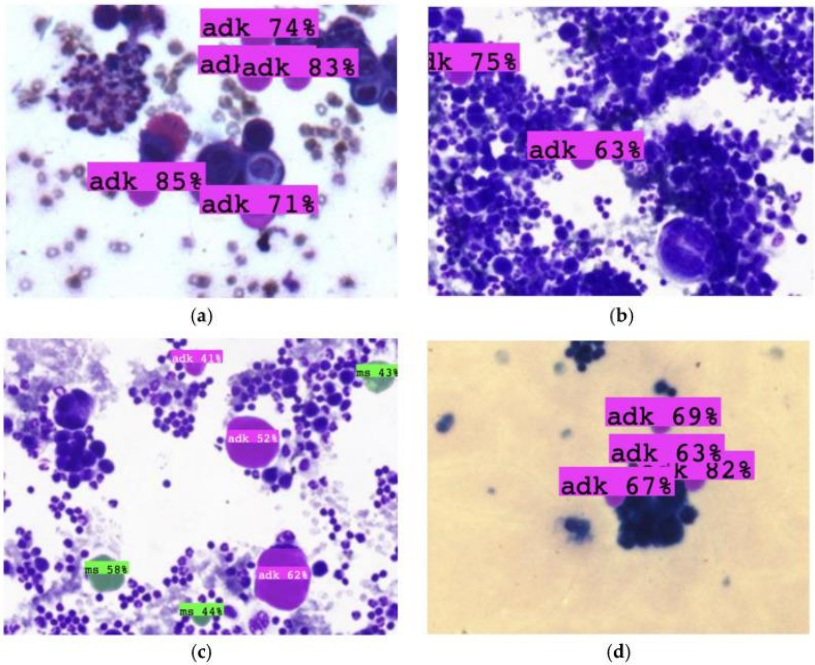


图 5 使用从包含来自体腔液的恶性和良性细胞图像的 Kaggle.com 获得的外部数据集进行了额外的验证

结果证明了与我们的数据集一致的预测，即使染色和放大倍率不同。该模型可用于初步观察和筛选，以确定是否应将样本提交进行免疫组织化学评估。

- (a) 到 (d)：来自外部数据集的图像，由两个模型正确分类，表明与训练数据集的一致性。

这些结果强调了模型泛化到新数据的能力，这对于它们在临床环境中的部署至关重要。即使细胞不清晰可见或出现在具有挑战性的条件下，细胞的准确分类表明，这些模型可能是在提交胸腔积液样本进行更详细的免疫组织化学评估之前初步筛选胸腔积液样本的有价值的应用。

3.3. 最终性能总结

综上所述，YOLO v8 和 YOLO v11 之间的比较表明，虽然 YOLO v11 在精确率、mAP 方面略优于 YOLO v8，但两种模型在胸腔积液细胞分类方面都表现出强大的性能。使用测试图像和外部数据集进行外部验证进一步支持模型泛化到新的、看不见

的数据的能力。图 3A、B、图 4 和图 5 提供了模型在各种实际场景中稳健性能的视觉证据，表明它们在临床实践中可用于胸腔积液样本的早期检测和筛查。虽然 YOLO v8 和 YOLO v11 都获得了相同的 F1 分数（0.72），但对各个指标的仔细分析会发现它们的性能存在细微的差异。YOLO v11 在 50% 的标准 IoU 阈值（mAP50：0.796 对 0.777）和从 50% 到 95% 的整个 IoU 范围内（mAP50-95：0.681 对 0.662）都表现出更高的精度（0.701 对 0.673）和更高的平均精度均值。这些结果表明，YOLO v11 总体上提供了更准确和一致的预测。相反，YOLO v8 的召回率略高（0.784 对 0.756），这表明它在检测相关实例时可能更敏感，尽管代价是假阳性更多。总体而言，YOLO v11 似乎在检测精度和定位精度之间取得了更有利的平衡，使其成为减少假阳性和确保一致分类是关键的临床应用的令人信服的选择。

4. 讨论

4.1. 总结和附加值

本研究调查了 YOLO v8 和 YOLO v11 对象检测模型在胸腔积液细胞学中的应用，重点关注区分反应性间皮细胞和腺癌细胞的挑战性任务。YOLO v11 在精度和 mAP 指标方面的性能略好，这可能是由于架构改进。探讨了在用户友好环境（例如基于 Web 的应用程序）中部署此类模型的可行性。这项研究为 AI 辅助细胞病理学的发展前景做出了多项贡献。

首先，它对应用于诊断要求苛刻的上下文的两种高级深度学习架构进行了重点比较评估。胸腔积液中反应性间皮细胞和腺癌细胞之间的细微形态差异通常给有经验的专业人士带来困难；本研究评估了最先进的对象检测模型如何通过基于指标的评估和可视化评估来应对这一挑战。

其次，该研究强调了细胞学中数据集创建和富集的实际问题。通过讨论基于 YOLO 的对象检测模型的应用并建议实验室间合作，它强调了加强医疗保健中 AI 系统开发和实施的重要策略。

第三个贡献在于围绕扩大 AI 开发机会的讨论。该研究强调了面向用户的应用程序如何支持数据注释、模型训练和集成过程，尤其是在细胞学等领域，这些领域通常缺乏专用的 AI 开发基础设施。

这为没有特定计算背景的医疗保健专业人员更广泛地参与 AI 模型开发打开了大门。

最后，通过将诊断模型构建为低细胞率环境（如恶性胸腔积液）中的潜在决策支持应用，该研究有助于围绕精准肿瘤学的对话。

4. 2. 讨论该研究对推进胸腔积液诊断技术的贡献

4. 2. 1. AI 用于胸腔积液诊断的最新进展

近年来，胸腔积液诊断方法的重点已经转移，不仅在实验室内，而且在需要细胞形态学诊断之外更多信息的临床医生中也是如此。在晚期肿瘤疾病患者中，尤其是肺腺癌患者，积液中的细胞是确定最合适的治疗策略的宝贵资源。肿瘤细胞可能在初步诊断时被发现，并提供对其分子特征的见解。在某些情况下，这些积液往往会在治疗过程中或之后重新积聚。在这种情况下，肿瘤细胞变得至关重要，尤其是对于其遗传特征的变化。即使对于经验丰富的细胞病理学家来说，准确鉴定和表征这些细胞也是首要挑战。

尽管人工智能（AI）主要用于组织病理学，但深度学习的进步极大地加速了计算细胞学的发展。近年来，将深度学习应用于不同细胞区室的筛选和分类任务、定位、识别和分割的趋势正在蓬勃发展。大多数基于深度学习的细胞学最初集中在宫颈，后来扩展到非妇科病变。

1995 年，Truong 等人基于积液细胞的密度测量和形态测量数据建立了人工神经网络（ANN）模型，分别报告了 95.3% 和 85.7% 的敏感性和特异性。2012 年，Barwad 等人使用监督学习模型进行网络训练，将 ANN 应用于积液细胞学，识别良性和恶性病例。最近，Xie 等人使用监督深度学习方法通过深度卷积神经网络（DCNN）对良性和恶性胸腔积液细胞进行分类，准确率、灵敏度和特异性分别为 91.67%、87.50% 和 94.44%。他们观察到，深度学习方法可以帮助病理学家诊断胸腔积液中的癌细胞，无论他们的经验如何。Park 等人应用深度学习来支持胸腔积液中转移性乳腺癌的诊断，证明深度卷积神经网络（DCNN）模型表现出更高的准确性、敏感性和特异性，优于病理学家。

4. 2. 2. 细胞病理学中的 YOLO 算法

对象检测算法是计算机视觉中的一项重要任务。传统上，感兴趣区域（ROI）的手动注释和分割是耗时且容易出错的过程。YOLO（You Only Look Once）算法在计算机视觉社区中引起了极大的关注。它是一个单级检测器，通过卷积神经网络（CNN）在一次正向传递中识别图像中的所有目标，预测图像中目标的边界框和类概率。YOLO在医学领域的应用引起了人们的兴趣，因为它能够实现各种病症的早期诊断和检测，包括视网膜疾病、脑萎缩、异常蛋白质沉积和心血管疾病。

在数字病理学方面，Rong 等人将 YOLO（HD-Yolo）应用于基于肺癌、肝癌和乳腺癌组织学的检测，加速细胞核分割和肿瘤微环境量化。Elazab 等人开发了一种混合模型，将 ResNet 50 的提取和分类功能与 YOLO v5 相结合，用于神经胶质瘤肿瘤诊断，显示出更高的准确性。Nambu 等人将 YOLO v4 与 ResNet 联合用于检测宫颈的非典型细胞，平均准确率和 F 测量值分别为 90.5% 和 70%。最近，Ultralytics 推出了 YOLO v8，这是对象检测和语义分割的重大进步，在各种应用中取得了可喜的成果。Topuz 等人证明了 YOLO v7 和 YOLO v8 检测有丝分裂细胞的能力，这是癌症侵袭性的关键指标。

Ikeda 等人将 YOLO v8 应用于 11 种具有不同染色方法的人类癌细胞系，观察其在临床环境中筛选和细胞分类的功效。YOLO v8 和 YOLO v11 已被证明在检测恶性和良性细胞方面非常有效，最近的研究报告了出色的性能。例如，Awad 等人使用 YOLO v8 和 YOLO v11 来区分恶性和良性白细胞，准确率达到 98.8%。

4.2.3. 研究贡献和与文献的比较

在我们拟议的研究中，我们比较了 YOLO v8 和 YOLO v11 在反应性间皮细胞和腺癌细胞之间强烈的类间相似性所带来的诊断挑战中的性能，特别是当这些细胞单独发生时。这些模型使用训练数据集和外部数据集进行了验证，这证明了它们能够泛化到看不见的数据。我们还证明了将这些模型集成到细胞检测过程的 Web 应用程序中的可行性。YOLO v8 和 YOLO v11 的指标结果都显示出良好的预测正确类别的能力，其中 YOLO v11 在精度、平均精度均值（mAP50）和 mAP50-95* 方面略优于 YOLO v8。

（注意*mAP50 以 50% 的最小重叠阈值测量模型的平均精度，而 mAP50-95 在 50% 到 95% 的重叠阈值范围内评估平均精度，从而提供更全面、更严格的模型性能评估。重要的是要记住，平均精度（mAP50）和 mAP50-95 都为模型检测和分类对象的能力提供了有价值的见解，其中 mAP50 侧重于特定阈值，而 mAP50-95 提供跨多个阈值的

更广泛评估)。

YOLO v11 性能的提高可归因于其增强的架构，其中包括更好的骨架设计和更精细的特征提取等修改，使其更有效地区分细胞形态的细微差异。

此外，来自公共数据集的单个图像的预测可视化(图 5)显示出类似的结果。YOLO v8 和 YOLO v11 都实现了准确的细胞检测，其中 YOLO v11 在稳健性方面略有优势，特别是在区分结构重叠的细胞方面，这是胸腔积液样本中的常见问题。尽管存在诊断挑战和有限的数据集，但获得的结果还是有希望的。但是，必须考虑几个因素。YOLO 模型的缺点包括需要一个大型数据集来训练强大的 AI 模型，以及难以检测较小的物体，这可能导致潜在的假阳性或假阴性。此外，虽然 YOLO v11 提供了更准确和可靠的预测，但由于细胞学中使用的细胞形态和染色方法的复杂性，它可能仍难以处理某些不太常见或更难分类的病理变异。

我们的结果与越来越多的研究结果一致，这些研究支持将深度学习整合到细胞病理学工作流程中。值得注意的是，以前的研究表明，基于 YOLO 的模型可以有效管理数字细胞学中形态异质性和重叠结构的挑战，尤其是在胸腔积液等基于液体的制备中。例如，Nambu 等人展示了 YOLO v4 与 ResNet 相结合如何增强对非典型宫颈细胞的检测，与传统的基于 CNN 的方法相比，准确性显著提高。同样，Topuz 等人强调了 YOLO v7 识别有丝分裂象的能力，这是肿瘤侵袭性细胞学评估的一个关键特征。这些研究强调了 YOLO 架构适用于需要精细区分细胞亚型的任务。此外，Ikeda 等人将 YOLO v8 应用于多种人类癌细胞系和染色条件，支持其跨细胞学域的稳健性和可转移性。基于这些见解，我们的研究证实，最近的 YOLO 迭代，例如 YOLO v11，在应用于间皮细胞与腺癌细胞区分等诊断具有挑战性的环境时，在泛化和精确性方面提供了具体优势。

我们的研究结果表明，YOLO v11 在与 AI 模型在可变和复杂临床数据集中的应用特别相关的指标上表现出比 YOLO v8 显著的性能改进。这些增强功能表明，YOLO v11 可能为将 AI 辅助细胞学整合到常规诊断工作流程中提供更强大的基础，其中一致性、适应性和可靠的检测至关重要。

这与 Ultralytics (<https://docs.ultralytics.com>, 2025 年 4 月 5 日访问) 报告的比较见解一致，支持较新的 YOLO 架构越来越完善以部署在现实世界的医疗环境中的观点。

这些发现为未来将 AI 辅助细胞学整合到临床环境中提供了一条潜在的途径，其中实时诊断支持可能成为一种宝贵的资源，尤其是在资源有限或患者数量多的环境中。

然而，虽然 YOLO v11 和 YOLO v8 等较大的模型确实需要强大的计算能力，而许多涉及胸腔积液样本诊断的医疗中心可能不容易获得这种能力，但重要的是要注意医院的计算能力正在迅速发展。近年来，硬件有了重大改进，包括 GPU 和专用处理单元（例如 TPU）的可用性增加，以及基于云的计算解决方案的使用越来越多。这些发展正在逐渐解决与深度学习模型的高计算需求相关的挑战。

此外，模型优化方面的创新（例如量化和修剪）使得在不牺牲性能的情况下减小 YOLO v11 和 YOLO v8 等模型的大小和计算要求成为可能。这些优化可以进一步促进这些模型在临床环境中的部署，即使在计算基础设施更有限的医院中也是如此。

因此，这些模型的实际采用变得越来越可行，不应低估它们增强诊断工作流程的潜力。快速准确地处理大量医学影像的能力可以显著提高诊断效率，使医疗保健专业人员能够及时做出更明智的决策。此外，医疗机构内部不断增长的计算能力可能会加速将高级模型整合到常规临床实践中，有助于简化诊断过程并提高胸腔积液分析的准确性。

总体而言，尽管计算限制仍然是一个挑战，但硬件和软件的不断进步正在逐步缓解这些障碍。因此，将先进的 AI 模型集成到临床实践中变得越来越容易。特别是，此类系统有可能通过以下方法增强诊断工作流程：

- 协助细胞病理学家对细胞学样本进行初步评估，在筛选阶段有效地充当虚拟第二意见；
- 减少需要额外调查（例如免疫组织化学分析）的病例数量；
- 降低总体诊断成本并最大限度地缩短周转时间。

4.3. 挑战和未来方向

用户友好的 AI 工具的日益普及使更多的实验室和临床中心能够探索基于深度学习的细胞学。

这种方法不仅可以加速 AI 原型设计，还可以通过开放存储库促进透明度和数据共享。这些发展促进了基于社区的模型和协作框架的出现，这些模型和框架可以在

标准化诊断工作流程中发挥核心作用，尤其是在资源匮乏或分散的卫生网络中。

虽然可能会出现协作挑战，例如数据隐私问题、地理障碍和基础设施差异，但这些挑战可以通过战略合作伙伴关系和创新解决方案来解决。例如，可以建立实验室间协议和联盟来共享匿名数据并共同开发强大的数据集。利用云平台和联合学习模型还可以促进跨中心安全高效的数据交换，从而在不损害患者隐私的情况下进行协作。

展望未来，生成对抗网络（GAN）提供了一个令人兴奋的机会来补充现有方法，特别是在解决数据稀缺和阶级失衡等问题方面。GAN 可以生成与真实细胞结构非常相似的合成图像，从而有可能扩展训练数据集并提高模型性能。例如，GAN 已成功用于医学应用，例如生成合成组织学图像。在一项研究中，Teramoto 等人使用渐进式生长 GAN（PGGAN）对真实和合成的肺细胞学图像进行分类，发现生成的图像与真实图像非常相似。当与 YOLO 等 AI 模型集成时，GAN 可以增强对稀有或代表性不足的细胞类型的识别，进一步提高诊断准确性。

此外，AI 与分子生物学数据的集成为推进精准医疗提供了巨大的潜力。通过分析肿瘤细胞环境和传统细胞学图像，机器学习模型可以实现更加个性化的癌症治疗策略，从而改善诊断和预后，尤其是在细胞数量有限的恶性胸腔积液（MPE）等复杂病例中。

随着该领域的发展，人工智能有望在支持诊断和治疗决策方面发挥越来越重要的作用，提供更准确、高效和个性化的癌症护理方法。通过技术解决方案和中心间战略合作伙伴关系克服协作障碍将是确保 AI 技术得到有效实施并在不同的医疗保健系统中实现其优势的关键。

4. 4. 研究的局限性

本研究提出了一些在解释结果时应考虑的局限性。该数据集是从圣安德烈亚大医院的细胞病理学部门收集的；然而，它包括相对大量的注释细胞用于细胞学研究，支持训练过程的稳健性。尽管如此，整合多中心数据可以增强形态学多样性并进一步提高模型的泛化性。

分析的重点是裁剪和注释的细胞图像，而不是全玻片图像（WSI）。虽然这种方法允许进行准确的对象检测训练，但未来的研究可以探索这些模型在真实世界诊断

环境中的表现，包括 WSI 和临床工作流程，其中载玻片质量和成像可变性是重要因素。此外，尽管 YOLO v11 在关键指标上优于 YOLO v8，但与其他 AI 架构（例如基于 transformer 或多模态模型）的基准测试可以为它们的比较诊断性能提供有价值的见解。

最后，该研究不包括可解释的 AI (XAI) 方法。将 XAI 集成到未来的实施中可以提高临床信任度，并促进 AI 辅助应用程序在细胞病理学实践中的采用。

5. 结论

本研究旨在探索机器学习的潜力，特别是对对象检测模型，在增强细胞病理学的诊断过程方面，特别是用于区分恶性胸腔积液（MPE）中的间皮细胞和腺癌细胞。主要目标之一是探索 AI 如何帮助使细胞病理学更容易获得，使细胞病理学家在各种医疗保健环境中（包括资源有限的机构）能够使用先进的诊断应用程序。通过提供可扩展且用户友好的 AI 解决方案，这种方法使临床医生（即使是那些在数据科学方面没有广泛专业知识的临床医生）能够采用基于 AI 的诊断工具并为开发做出贡献。这可以显着减少技术障碍，并使世界各地的医疗保健专业人员能够提高诊断效率和准确性，无论他们的技术专长如何。

结果表明，YOLO v8 和 YOLO v11 模型在检测间皮细胞和腺癌细胞方面均有效，其中 YOLO v11 在精度和 mAP 指标方面的表现略好。这些发现突出了机器学习模型的潜力，通过帮助病理学家做出更准确和有效的诊断来补充人类专业知识。虽然最终决策始终需要人类专业知识，但机器学习的集成具有显著的优势，例如通过揭示可能难以手动检测的细微形态差异来缩短诊断时间并可能提高准确性。

此外，这项研究强调了扩展数据集的重要性，尤其是在细胞病理学等领域，在这些领域，数据可用性可能是一个限制因素。通过生成对抗网络（GAN）生成合成图像以及促进实验室之间的协作以共享数据集等策略对于克服这一挑战和提高模型稳健性至关重要。

此外，这项研究表明，作为未来的步骤，将分子生物学数据与机器学习模型相结合可以进一步提高诊断准确性。通过结合生物学见解，AI 模型可以增强对细胞特征的解释，为精准医疗时代更有效和个性化的治疗策略铺平道路。

总体而言，虽然细胞病理学中的 AI 仍处于早期阶段，但其支持诊断决策、实现

高级应用程序访问民主化以及与分子数据集成的潜力为改善患者预后提供了有希望的机会。AI 在医疗保健领域的未来取决于克服当前的挑战，例如数据集限制和模型可解释性，但迄今为止取得的进展标志着临床诊断和治疗计划的变革性转变。展望未来，AI 模型在细胞病理学中的持续开发和集成为增强诊断工作流程提供了重大前景。借助使临床医生能够在没有深厚技术专业知识的情况下创建和部署机器学习模型的平台，AI 有可能使高级诊断应用程序更易于访问，即使在资源有限的环境中也是如此。随着该领域的发展，克服数据限制和模型可解释性等挑战对于充分发挥 AI 在改变诊断实践和改善患者护理方面的潜力至关重要。

***注：**原文和译文版权分属作者和译者所有，若转载、引用或发表，请标明出处。

译文二：

AI 驱动精准医疗：通过优化遗传风险因素彻底改变医疗保健

Sakhaa Alsaedi, Michihiro Ogawara, Mohammed Alarawi,

Xin Gao, Takashi Gojobori, 徐健 (译)

来源: NAR Genom Bioinform.

时间: 2025 年 5 月.

链接: <https://doi.org/10.1093/nargab/lqaf038>.

介绍

人工智能 (AI) 正在改变组学数据的分析, 包括基因组学、转录组学和蛋白质组学, 从而增强精准医学中的生物信息学。精准医疗是一种根据个人特定需求制定医疗保健的医疗方法, 包括健康个体和患者, 分为早期筛查、精准诊断、精准临床治疗和人工智能增强的个性化健康管理四个阶段, 如图 1 所示。在这种方法中, 经常采用诊断测试, 根据患者的基因图谱或其他生化或细胞分析来选择合适和最佳的药物。目前, 大多数精准医学干预都涉及基因分析, 其中包括识别预测性生物标志物和遗传风险因素 (GRF)。这种方法可以识别有某些疾病风险或严重疾病表现的患者, 从而允许采取预防性干预措施, 从而降低疾病复杂性并提高医疗保健质量。人工智能和精准医学的融合对于理解疾病的 GRF 尤为重要。例如, 使用 AI 整合不同模式的组学数据可以增强生物医学知识平台, 并提供疾病致病性的深入表观遗传学分析。

该图说明了精准医疗的框架, 分为四个关键阶段: 早期筛查、精准诊断、精准临床治疗和 AI 增强的个性化健康管理。这种方法整合了生物多模式和临床数据, 通过 AI 工具和可穿戴技术实现疾病风险的早期检测、准确疾病分类、量身定制治疗方案和持续健康管理。

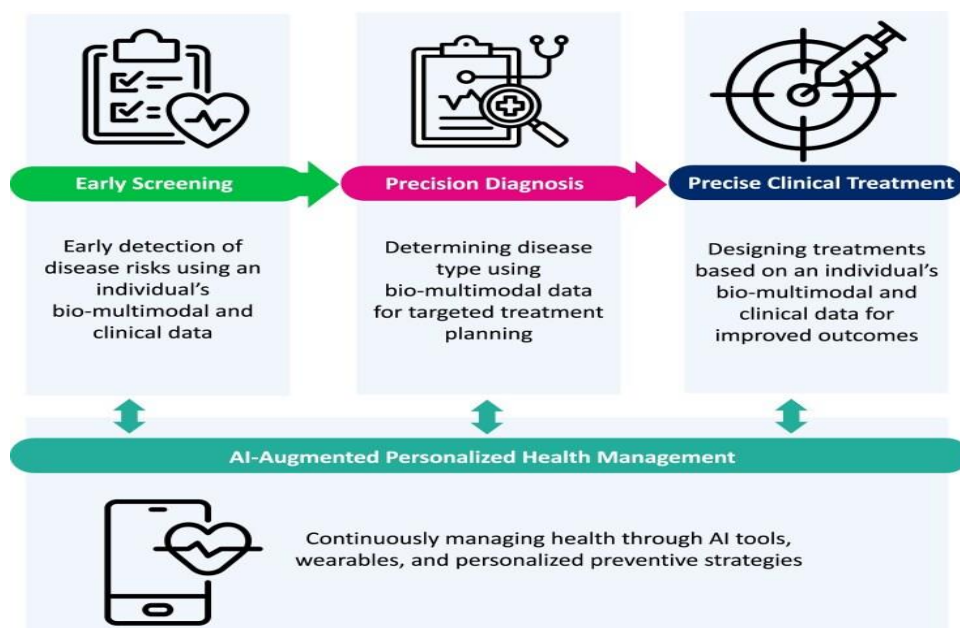


图 1 定制医疗保健的精准医疗工作流程概述

尽管 GRF 代表了医疗保健领域的重大进步，但它们目前在精准医疗中的应用受到一些挑战的限制。例如，由于 GRF 的表达变量、复杂的相互作用、持续的发现以及对人群的不同影响，目前尚无明确的定义。此外，GRF 可能并不同样适用于所有疾病，尤其是受多个基因和环境因素影响的复杂疾病。破译 GRF 与环境因素之间复杂的相互作用仍然很困难，这会导致潜在的误解和不准确的风险预测。在基因检测中，数据隐私和安全等伦理问题至关重要。确保知情同意并解决基于遗传信息的潜在歧视至关重要。此外，基因检测和个体化治疗的成本和可及性可能令人望而却步，限制了一些患者的可及性。

为了解决这些问题，本文展示了利用高级 AI 模型应对与 GRF 相关的挑战的适用性，并提供了解决这些问题的潜在解决方案。它采用人工智能模型来解码和解释复杂的遗传信息，从而加深我们对个体遗传风险概况的理解。人工智能在精准医疗中的兴起代表了一种范式转变，使医疗保健能够根据每个人的基因组进行定制。通过各种先进的 AI 模型、案例研究和示例，我们说明了如何将 AI 增强的遗传洞察力整合到医疗实践中，从而实现更准确的诊断、有针对性的治疗和更好的患者护理，如图 1 所示。这篇综述强调了人工智能利用的 GRF 在推动精准医疗的未来和彻底改变医疗保健方面的作用。

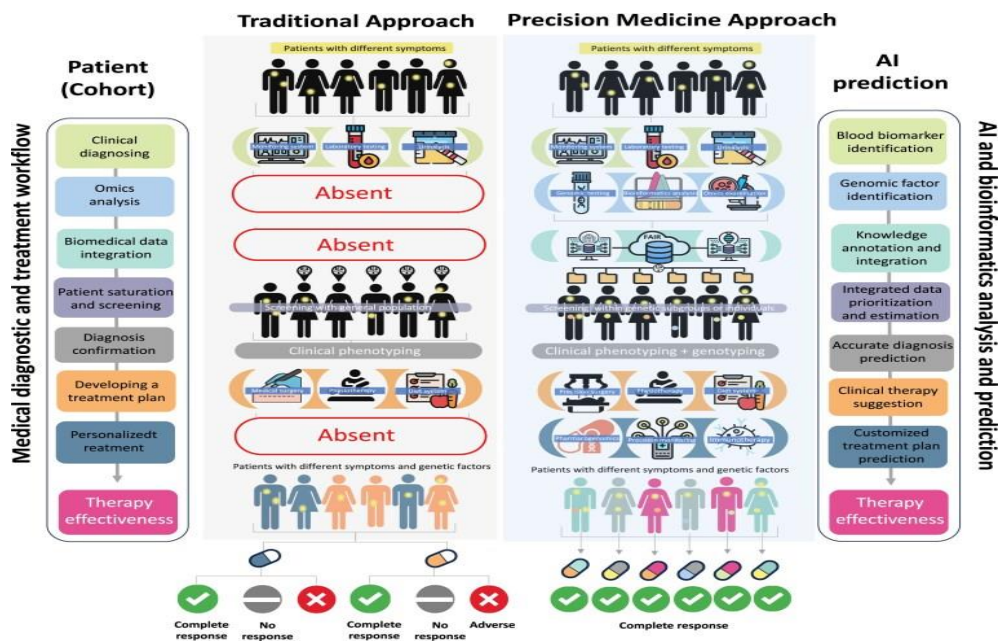


图 2 医学诊断和治疗中的精准医学方法概述

该图比较了诊断和治疗中的传统医学和精准医学方法。传统方法包括一般临床诊断和治疗效果评估。然而，精准医疗使用基因组检测、生物信息学分析、组学检查和个性化治疗计划。顶层的点表示来自临床问卷的症状，而中间层的点表示通过遗传分析确定的遗传风险因素（GRF），这些因素不包括在传统方法中。精准医疗将患者数据与公共生物医学资源相结合，以实现准确诊断和个性化护理，利用 AI 提供更好的治疗建议并提高疗效。该图说明了从通用医疗保健解决方案到个性化医疗保健解决方案的转变。

精准医疗中的 GRF

GRF 是精准医学的核心，它允许将对疾病机制的遗传理解与分析遗传数据的复杂生物信息学方法相结合。本综述在单核苷酸多态性（SNP）水平上考虑了 GRF，这些 SNP 有助于疾病发展或增加疾病易感性和复杂性。值得注意的是，与这些 SNP 相关的宿主基因和蛋白质也被认为是风险因素。我们将 GRF 分为三类：稀有、常见和模糊 GRF。如图 3 所示，分类基于次要等位基因频率（MAF）、遗传贡献、分子网络中的功能参与、变异数量、有效大小、环境影响、疾病突变的外显率以及对疾病风险及其病因的贡献。然而，这种分类简化了遗传变异的复杂相互作用和连续体，这可能导致这些类别无法完全捕捉到细微的风险因素。

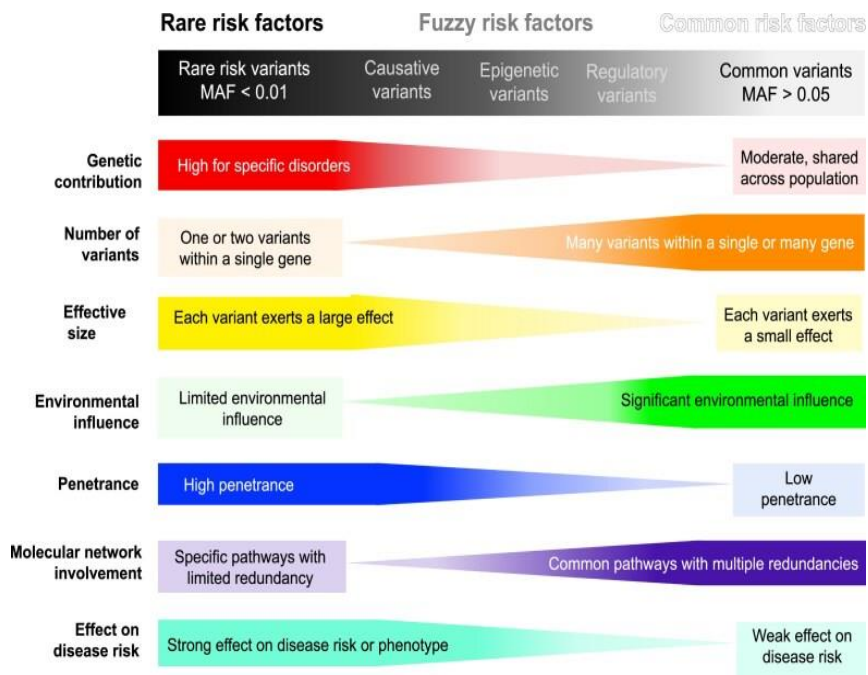


图 3 GRF 的特征

该图说明了导致疾病风险的遗传变异谱，较深的颜色表示该图说明了导致疾病风险的遗传变异谱，按其影响的强度进行组织。GRF 分为罕见的变体，通常具有较大的效应、高外显率和最小的环境影响。这些变异通常涉及特定的生物途径。常见变异影响小、外显率低，并且受环境因素的强烈影响。它们通常影响多个基因和通路。模糊风险因素，包括致病、表观遗传和调节变异，根据其相对 MAF 占据罕见和常见变异之间的中间空间。它们在基因调控和疾病易感性中表现出不同的作用。

稀有 GRF

罕见遗传因素是指在一般人群中不常发生的特异性遗传变异或突变 (MAF < 0.01)。这些变异通常具有较高的外显率，这意味着携带这些罕见变异的个体患特定疾病的可能性很高。罕见的遗传变异通常表现出最小的环境影响，但对疾病风险和表型有显著影响，通常直接导致特定的疾病或病症。尽管罕见变异的影响很大，但由于它们的患病率低，鉴定罕见变异带来了重大挑战，需要进行大规模测序研究，并且通常依赖于基于家族的设计。此外，数据解释、变体注释和计算限制的复杂性也是相当大的障碍。尽管与常见变异相比，罕见变异在人群中并不经常出现，但它们对疾病的影响很高，对于受影响家庭的准确遗传咨询和精准医疗至关重要。未来的研究应侧重于开发更有效的识别方法和计算工具，以更好地了解这些变异。

常见 GRF

常见的遗传因素是一般人群中经常发生的特异性遗传变异或突变，通常为（ $MAF > 0.05$ ）。这些变异具有低外显率和适度的效应量，这意味着它们单独对疾病发展的影响很小。然而，它们共同可以通过多基因效应显著影响疾病易感性和表型。与罕见变异不同，常见变异通常表现出大量的环境相互作用，其中生活方式和环境因素会调节它们对疾病风险的影响。这些环境影响很复杂，并且在不同的人群和背景下可能会有很大差异。常见变异主要与复杂的多因素疾病有关，如糖尿病、心脏病和多种形式的癌症。这些变异通常通过全基因组关联研究（GWAS）来识别。尽管它们的个体影响较小，但多个常见变异的累积效应可能很大，这使得它们对于理解常见疾病的遗传结构和开发用于个性化医疗的多基因风险评分（PRS）至关重要。识别和研究常见变异为复杂疾病的遗传基础提供了有价值的见解，并为预防、诊断和治疗策略提供了信息。它们在人群中的高频率使其成为识别人群范围危险因素和了解疾病病因的有用标志物。环境数据与遗传数据的全面整合对于更深入地了解基因-环境相互作用至关重要。

模糊 GRF

模糊 GRF 代表一类遗传变异，其中 MAF 介于稀有变异和常见变异之间（ $0.01 - 0.05$ ）。这些变异通常具有中等频率，并且对疾病风险或表型的影响各不相同。模糊遗传因素包括致病变异、调节变异和表观遗传变异，它们在遗传和表型多样性中发挥着不同但相互关联的作用。这些模糊变异的鉴定和研究需要基因组测序和 GWAS 之外的先进技术，例如染色质免疫沉淀测序、RNA 测序（RNA-seq）和亚硫酸氢盐测序，以捕捉基因调控和表观遗传修饰的动态性质。识别和研究模糊遗传因素对于理解多因素疾病的细微遗传结构至关重要。这些变异之间错综复杂的相互作用及其调节基因表达和表型的能力凸显了它们在精准医学和个性化治疗方法中的重要性。然而，应强调其影响的概率性质以及研究它们与环境因素和其他遗传变异相互作用的挑战。

致病变异

致病变异是通过改变基因功能或表达直接导致疾病或特定表型的遗传改变。这

些变异通常是导致单基因或孟德尔疾病的原因，其中单个基因突变会导致特定疾病。值得注意的是，一些致病变异也可能很常见；例如，血红蛋白亚基 β （HBB）基因的某些变异赋予对疟疾的抵抗力，同时导致纯合子个体出现镰状细胞贫血。了解致病变异的全部范围，包括它们的不同频率和影响，对于全面的遗传咨询和精准医学策略至关重要。

监管变体

调节变异通过修饰启动子、增强子和沉默子等调节元件来影响基因表达，从而影响基因活性的时间、位置和水平。这些变异在基因表达的微调调节和塑造复杂的性状结构中起着关键作用。例如，与脂肪量和肥胖相关的脂肪量和肥胖相关（FTO）基因的增强子区域的变异通过改变参与脂肪代谢的基因的表达而与肥胖有关。受遗传和环境因素影响的调节变异的动态性和上下文依赖性强调了需要先进的分析方法来全面研究其影响。

表观遗传变异

表观遗传变异在不改变 DNA 序列的情况下改变基因功能，通过 DNA 甲基化、组蛋白修饰和非编码 RNA 相互作用等机制修饰表观基因组。这些变异与环境因素的相互作用往往使它们对疾病风险和表型的影响复杂而多变。因此，表观遗传变异在遗传易感性和环境影响之间提供了重要的联系。对这些变异的研究需要一种综合方法，结合遗传、表观遗传和环境数据，以了解它们在健康和疾病中的多方面作用。

在精准医疗中利用 GRF 的 AI 方法

人工智能增强了实验、计算和遗传技术，以揭示疾病的遗传基础，从而实现更有效的治疗。这些 AI 模型利用来自文本、图像、表格数据和图表等各种模式的生物学大数据来训练其模型以执行不同的下游任务。人工智能方法的最新进展，包括统计模型、联邦学习、机器学习（ML）和深度学习（DL）方法，如生成模型、表示模型和语言模型，显著扩展了个性化医疗的潜力。无花果。说明了 AI 如何集成来自多种生物学模式的数据以识别风险因素并将其应用于精准医疗。

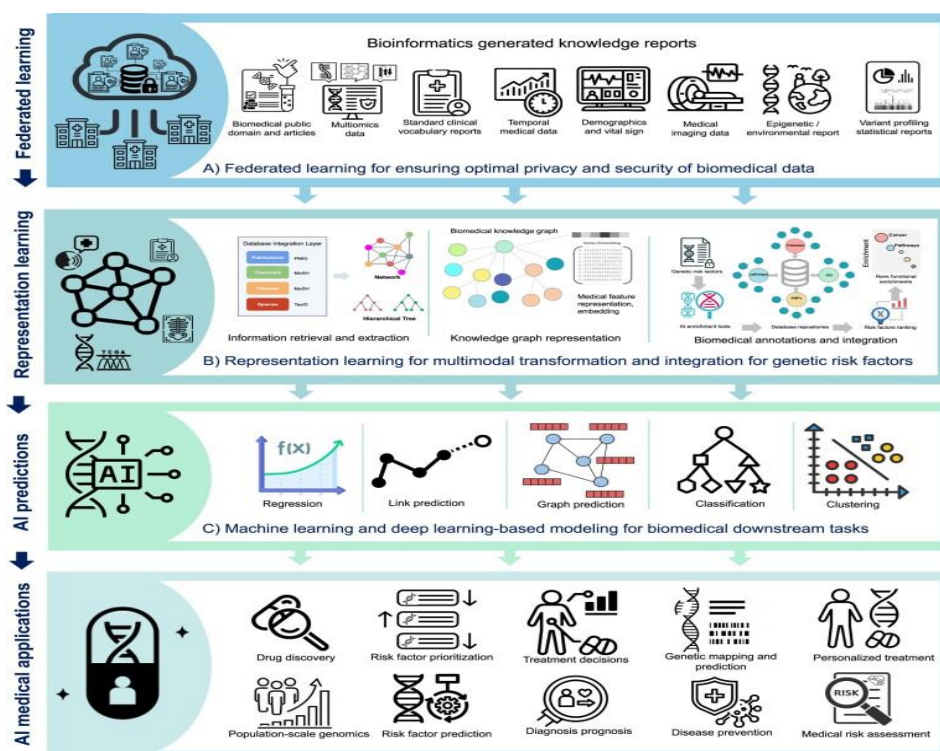


图 4. 基于 AI 的生物医学应用程序用于医疗保健中的 GRF

该图说明了如何利用先进的 AI 模型来利用医疗数据资源，为不同的医疗任务提供准确的应用程序。它包括：（A）用于确保生物医学数据的最佳隐私和安全的联邦学习，（B）用于 GRF 的多模态转换和集成的表示学习，以及（C）用于生物学下游任务的 ML 和 DL 建模。该图强调了人工智能的使用，通过利用基因图谱、风险因素预测和其他生物医学数据来增强治疗决策、医疗风险评估、药物发现和个性化治疗。

生物医学数据的数据隐私和成本效益

生物医学数据资源放置在图 1 的顶层。突出了它们在推进 AI 应用方面的关键作用。蛋白质数据库（PDB）、英国生物样本库和 TCGA 等高质量数据集通过为 AI 训练提供多样化的大规模数据，在蛋白质预测、基因组学和精准医学方面取得了重大进步。共享这些数据集可以促进协作，提高模型性能，并加快诊断和治疗进展。然而，共享敏感的生物医学和遗传数据会引发严重的隐私问题。必须保护患者的医疗和遗传信息不被滥用，同时仍然能够共享数据用于研究。

为了满足这些要求，开发了联合学习算法来加密敏感数据。例如，PPML-0mics 采用分散式差分隐私技术来安全地分析多组学数据，同时保持患者的机密性。遵循可查找、可访问、可互作和可重用原则（FAIR），联邦学习工作流程提供了鼓励患者共

享其遗传数据的安全渠道。EU Gene 是一个 FAIR 工具包，简化了基因组序列的 DL 分析，有助于基因功能预测和疾病变异分析等任务。这种方法具有成本效益，可以使用公共加密数据来更好地识别风险、早期疾病检测和复杂疾病的定制治疗。联邦学习使医院和诊所能够安全地共享患者的生物医学和遗传记录，通过将遗传因素、临床记录 and 生活方式信息整合到精准医疗中来改善医疗保健结果。Decagon 和 Path GNN 等工具使用此类数据集进一步增强了药物发现和疾病建模。此外，AI 通过为处理敏感数据集提供安全的解决方案，在推进生物医学研究的同时增强了数据隐私。Watsonfor Genomics 等平台以及 Human Cell Atlas 和 GenBank 等计划为训练 AI 模型提供了高质量的精选数据集，从而推动了各种生物医学应用的创新。根据 FAIR 原则共享私有数据集对于实现精准医疗、促进协作以及推进诊断和治疗至关重要。

生物医学数据集成和知识表示

集成和构建生物医学数据对于丰富功能、增强 AI 模型预测、确保可重复性和获得见解至关重要。AI 方法利用先进的数据挖掘和检索技术从 PubMed 和 Med Plus 等平台中提取重要信息。例如，文本挖掘和检索方法（如 SNP Map 和 Lit Var）通过可视化和检索与疾病相关的遗传因素，在自动化与变异相关的文献搜索中发挥着关键作用。这些工具通过利用 Pub Med 和 Pub Tator 等资源将遗传变异与疾病和化学物质联系起来。通过利用先进的文本挖掘技术，它们提供全面、最新和高效的语义搜索，显著提高了生物医学知识在研究和临床应用中的可访问性和可用性。

此外，将多组学数据与风险因素相结合有助于识别有风险的分子途径，并突出复杂疾病的潜在药物靶点。这种方法可以探索导致疾病复杂性的基因的复杂分子途径。Deep His CoM 分析多组学数据、GWAS 数据和本体，以预测复杂疾病中的通路效应。ExPecto 利用 DL 来预测遗传变异对基因表达和疾病风险的影响。AI 通过将多组学信息与带注释的资源（如 Open Targets Platform 和 Casual DB 以及更多带注释的数据库）集成，释放了复杂生物医学数据的力量。

尽管维护和集成数据丰富了 AI 模型，但稀疏性和不一致等挑战仍然存在。将原始数据转换为结构化格式有助于揭示变异、疾病和分子网络之间的关系，从而解决多因素疾病的复杂性。图 3 中的第三阶段。知识表示学习，构建数据以增强 AI 驱动的洞察力。例如，疾病知识图谱整合了不同的数据集，并系统地揭示了影响疾病进展的遗传、分子和环境因素的相互作用。使用 AI 进行表征学习从原始数据中提取模

式并将其组织成知识图谱，将基因、蛋白质、药物和表型等实体联系起来。这些图表为探索分子相互作用、致病机制和治疗途径提供了一个统一的框架。

知识表示学习对于推进生物医学研究至关重要，因为它将复杂的数据集构建成可作的框架。特别是知识图谱，能够系统地表示相互关联的实体，如基因、蛋白质和疾病，使研究人员能够发现隐藏的关系和模式。个性化医疗中的生物医学知识表示增强了复杂数据模式的处理，通过处理复杂的数据模式提高了风险预测的准确性。这些图表将不同的数据集（包括基因组变异、分子网络和疾病表型）集成到统一的系统中，从而提供多因素疾病的全面视图。通过利用图形卷积网络（GCN）、图形注意力网络（GAT）等高级模型和 Node2Vec 等嵌入技术，知识表示框架支持预测分析，从而促进风险评估、生物标志物发现和途径探索等任务。

AI 驱动的知识图谱擅长集成多模态数据集，使研究人员能够从系统的角度研究疾病。疾病受多种遗传、环境和分子因素的影响，这使得传统方法无法完全捕捉其复杂性。STRING、DisGeNET 和 OpenTargets 等知识图谱揭示了遗传变异、分子网络和疾病机制之间的复杂关系。例如，STRING 绘制蛋白质-蛋白质相互作用图，而 DisGeNET 将遗传变异与疾病表型联系起来，有助于识别关键分子通路。这种能力通过桥接转录组学和蛋白质组学等数据模式来支持跨学科研究，揭示对疾病机制的关键见解，并能够发现治疗靶点。

在功能基因组学和疾病研究中，AI 驱动的知识图谱通过对调控网络及其对基因表达的影响进行建模来增强基因相互作用分析。通过整合来自 ENCODE 和 GTEx 等资源的组织特异性数据，这些图表揭示了非编码变异如何影响基因调控。例如，应用于 STRING 数据的 GCN 可以预测新的基因-疾病关联或突出转录途径中的关键调控元件。

PASNet 等应用程序使用多组学数据进行预后预测，而 Decagon 等工具则合成多模态信息来识别药物靶点并重新调整治疗用途。通过系统地捕获复杂的关系，知识图谱可以推动生物标志物的发现，确定治疗靶点的优先级，并促进个性化医疗。例如，SNPMap 通过从各种来源提取和可视化 SNP 数据来阐明与疾病相关的遗传因素。MecCog 是一个图形化知识框架，它整合了遗传变异和疾病机制，用于发现致病生物标志物和治疗。PathGNN 将通路特征捕获为生物医学知识图谱以预测癌症生存率，而 PASNet 使用高通量数据预测体细胞变异的预后。Decagon 整合了遗传风险、致病因

素和多组学数据以进行低维表示，有助于医疗决策。DL 平台，如 Gene Ontology、Pathology Workbench 和 Watsonfor Genomics，可以转化复杂的遗传风险，为医疗保健专业人员和患者提供支持。AI 简化了丰富的信息，以便于访问和理解生物医学知识表示。它们能够将不同的数据集转化为可作的见解，这表明它们在推进精准医学和加速疾病理解突破方面的潜力。

生物医学应用中 AI 驱动的计算下游任务的比较概述

图 3 的第 3 阶段和第 4 阶段。概述了将生物医学数据处理为 GRF 分析和预测的见解的主要计算任务。主要任务包括回归、链接预测、图形表示、聚类 and 分类，每个任务都由其计算方法和应用程序定义。

回归可识别输入特征和输出之间的连续关系，例如预测遗传风险评分或疾病进展。PRSice、DeepWAS 和 PolyPred 等工具利用统计模型和 DL 获取高维基因组数据，有效地计算 PRS。链接预测评估结构化数据集中的关联，例如基因-疾病或蛋白质-蛋白质相互作用。SNPMap 和 GRAM 等方法利用基于图的算法来推断有意义的关系，而 MiRLog 和 dbmiR 则专注于优先考虑涉及 microRNA 变体的相互作用。图表示将复杂的生物医学关系编码为结构化网络。GRAM、MecCog 和 SNP Map 等方法将数据整合到图表中，用于通路分析和变异可视化等任务。Variant Spark 等工具分析大规模多组学数据，通过基于图的嵌入揭示疾病相关变异。聚类通过根据相似性指标或潜在空间对数据进行分组来揭示固有模式。Deep His CoM 和 MUSSEL 等工具揭示了与疾病相关的亚群或调控元件。Reg Base 等方法应用聚类来高精度注释和预测非编码调节变异。分类将实体分配给预定义的类别，例如致病性或良性变体。CADD、EVE 和 Xrare 等工具在变异分类方面表现出色，在识别致病性或罕见病相关变异方面具有很高的准确性。同样，DeepSEA 和 RegVar 等工具优先考虑编码和非编码基因组区域的功能影响预测。

在图 4 的第 4 阶段中。通过利用 GRF 分析中的计算任务，包括风险评估、早期预防、疾病诊断、治疗预测和药物发现，增强了多个应用程序。例如，预测遗传变异的功能影响对于理解基因调控和疾病机制至关重要。经典的计算机预测方法，如 FATHMM，利用隐马尔可夫模型来评估蛋白质错义变异的功能影响和致病性。同样，PolyPhen-2 采用基于序列和结构的特征与朴素贝叶斯分类器相结合来预测蛋白质编码遗传变异的影响。此外，ASpedia FI 识别差异可变剪接事件和共调控基因，以使

用交互网络和随机游走算法预测风险变异的功能影响。基于 DL 的方法（如 Deep SEA 和 Deep CAGE）利用 CNN 来预测 GWAS 统计报告中的染色质特征、转录因子结合和变异效应，从而深入了解基因调控和潜在的疾病机制。ExPecto 使用 CNN 预测突变的组织特异性转录效应。RegVar 预测非编码变异的组织特异性调节影响。

在解码功能影响时，AI 算法集成了各种 ML 和统计方法，以识别变体的致病性。OncoBase 解码了癌症中的调节性体细胞突变，X-CNV、StrVCTVR 和 ClassifyCNV 已经根据该基因确定了变异的拷贝数。此外，parSMURF 使用具有过采样和欠采样技术的 ML 处理致病性变异检测中的不平衡数据。公共基因组学资源与多个带注释数据集的集成相结合，提供了另一种方法。FAVOR、Funseq2 和 Deeppost-GWAS 等 ML 模型利用这种策略，整合各种组学数据来查明与复杂疾病相关的风险基因和变异。iMEGES 是一种 DL 模型，它分析全基因组测序数据以查明精神障碍的潜在风险基因，有助于群体研究和个性化医疗。

对于编码变体，ESM1b 使用蛋白质语言模型来预测编码变体的效果。它还将变体注释为仅在特定蛋白质亚型中具有破坏性，并推广到复杂的编码变体。这些工具通过增强对不同组织中基因调控的理解来推进精准医学。

此外，优先考虑致病突变和致病性变异对于理解疾病机制至关重要。精细映射方法，如 FINEMAP 和 PICS2，通过从 GWAS 数据中识别致病因素来优化资源分配。人工智能驱动的工具，如 MUFFIN，整合了数据以对遗传变异进行功能注释，支持识别疾病相关突变，以精确定位疾病中的致病突变并预测变异的致病性和致病效应。

在风险评估和早期预防的背景下，AI 模型通过自动分析大量基因组数据集来增强 GWAS。Eigen GWAS 等方法采用特征向量分解来识别遗传关联。Deep WAS 使用 CNN 进行准确的 SNP 鉴定。这些进步通过支持个性化的疾病风险预测和预防，为精准医疗做出了重大贡献。这种 AI 驱动的方法增强了对疾病风险的预测，解决了精准医疗中遗传相互作用和环境影响的复杂性。

GRF 分析中的计算任务迎合了遗传研究的不同方面，其工具在特定领域表现出色。PRSice 和 Deep WAS 等回归工具对多基因风险评分非常有效，可同时利用统计模型和 DL 处理高维数据。和 SNP Map 等链接预测工具在发现关联方面表现出色，而 Variant Spark 和 MecCog 等图形表示方法则提供对复杂关系的结构化洞察。包括 CADD 和 Xrare 在内的分类工具对于识别致病性变异非常准确，而 Deep His CoM 等

聚类方法揭示了调节元件中的隐藏模式。虽然 FATHMM 和 PolyPhen-2 等传统工具轻量级且可解释，但 Deep SEA 和 ExPecto 等 DL 模型为功能影响预测提供了无与伦比的准确性，但计算成本更高。总体而言，工具的选择取决于计算需求和任务特异性之间的权衡，先进的 AI 模型为复杂分析提供了更深入的见解。

用于风险预测和估计的 AI 模型的比较评估

为了简化管理，我们专注于构成 GRF 分析综合工作流程的四个核心下游任务：风险变体注释、优先级排序、估计和解释。对这些任务的 20 种方法中的 15 种进行了比较评估，包括统计模型等非 AI 方法和 ML 和 DL 模型等 AI 方法，突出了它们基于性能、复杂性、鲁棒性和可扩展性的优势和局限性。此外，我们评估了这些方法，并根据五个重要标准对它们进行了排名（1-5）：实验验证、计算资源优化、模型性能和新颖性、灵活性和与其他工具的集成，以及对其他生物学任务的可扩展性。例如，这包括将不同 DNA 区域的分析应用于其他数据集或应用程序。表 2 提供了对这 15 种方法的全面分析和评估。

表 2. 用于基因组分析的 AI 工具比较，包括计算资源、性能指标和排名

工具	AI 模型	后阶任务	目的	优势	局限性	计算资源	性能	等级
Deep Sea	下载: CNN	功能注释	注释变体的功能影响	整合监管基因组数据	组织特异性有限	中等 (16GBRAM)	>90%的染色质特征预测 AUROC 准确率	5
LINSIGHT	ML: 线性模型	功能注释	对非编码变体影响进行评分	集成进化和功能数据	有限的上下文特异性	低 (标准笔记本电脑资源足够)	91%AUC 用于非编码有害性预测	4
MiRLog	ML: 逻辑回归	功能注释	注释 miRNA 序列变体	对 miRNA 相关研究有效	仅限于 miRNA 数据集	低 (标准笔记本电脑资源足够)	>85%的 miRNA 区域功能注释的平均 AUC	4
DeepVariant	下载: CNN	功能注释	检测和注释遗传变异	精度高;广泛验证	计算密集型	高 (需要>32GBRAM, 建议使用 GPU)	>99%的 SNP 准确率	3
DeepSNP	DL: 基于注意力的神经网络	功能注释	检测结构变异	结构变化精度高	缺乏实验验证	中等 (16GBRAM)	结构变异检测中 90%的宏 F1 评分	3
RegVar	DL: DNN	遗传风险优先级排序	优先考虑组织特异性非编码变异	强大的组织特异性优先级	需要高质量的组织数据	高 (需要>32GBRAM, 建议使用 GPU)	预测调节变异的 AUC 为 96%;>88%AUC 用于区分致病性变异	4

工具	AI 模型	后阶任务	目的	优势	局限性	计算资源	性能	等级
FINEMAP	SL: 贝叶斯变量选择	遗传风险优先排序	识别致病遗传变异	高效处理大型数据集	假设线性遗传结构	中等 (16GBRAM)	精细映射和因果变异预测的准确率达到 99%	4
DeepWAS	下载: CNN	遗传风险优先排序	精细映射 GWAS 关联	强大的多变量关联	计算成本高	高 (需要>32GBRAM, 建议使用 GPU)	GWAS 信号恢复准确率>90%	4
CADD	SL: 集成学习	遗传风险优先排序	识别有害变体	经过广泛验证;集成多个注释	仅限于预定义的注释	中等 (16GBRAM)	有害性和致病性变异预测准确率达到 98%	5
GREEN-DB	各种 ML 模型	遗传风险优先排序	使用多组学注释调节变体	强大的注释框架	需要全面的多组学数据	中等 (16GBRAM)	90%的注释可靠性	4
WEVar	统计学习	风险评估	预测变体的监管影响	包含多种评分方法	基准测试有限的较新工具	低 (标准笔记本电脑资源足够)	>91%AUC 用于非编码功能预测	4
GRAM	ML: 随机森林	风险评估	预测非编码变体的分子效应	对细胞类型分析具有高度特异性	计算成本很高	中等 (16GBRAM)	94%的 AUC, 在细胞类型预测中具有高特异性	4
Xrare	ML: 梯度提升决策树	风险评估	优先考虑罕见病相关变异	对临床应用有效	仅限于罕见病	低 (标准笔记本电脑资源足够)	95%稀有变体的平均准确率	4
EVE	DL: 生成模型	风险评估	预测罕见变异的致病性	对罕见变异有效	仅限于蛋白质编码区域	中等 (16GBRAM)	罕见致病性变异的平均 AUC 为 91%	4
PolyPred	SL: 线性混合模型+多基因模型	风险评估	预测不同群体的遗传风险	高跨种群准确率	需要插补数据集	中等 (16GBRAM)	>92%不同人群 PRS 的平均 AUC	4
PRSice	统计线性模型	风险评估	计算 PRS	灵活;处理大型数据集	计算密集型, 适用于大规模队列	中等 (16GBRAM)	PRS 预测精度高	2
MecCog	图形模型	可视化和解释	了解疾病机制	整合生物医学文献和真实世界数据	需要专家策划	低 (标准笔记本电脑资源足够)	90%的生物标志物识别准确率	4
DeepHisCom	下载: CNN	可视化和解释	使用多组学数据分析通路	整合多组学数据	需要大量的训练数据	高 (需要>32GBRAM, 建议使用 GPU)	94%的通路分析精度	4
SNPMap	SL: 自动化管道	可视化和解释	可视化和解释 SNP 效应	用户友好;集成来自多个来源的数据	仅限于预先存在的注释	低 (标准笔记本电脑资源足够)	SNP 解析准确率达到 85%	4
VariantS	ML: 随机	可视	识别复杂	处理复杂的	不太容易	高 (需	复杂性状的	2

工具	AI 模型	后阶任务	目的	优势	局限性	计算资源	性能	等级
park	森林系综	化和解释	的特征关联	数据集和 SNP 交互	解释;需要大量资源	要>32GBRAM)	预测准确率 达到 85%	

任务 1: GRF 注释和识别

非编码区影响基因表达，并在疾病风险方面超过编码区，突出了它们在识别风险变异方面的重要性。我们比较了靶向和注释监管要素的方法。

DeepSEA 使用 CNN 来预测非编码变体的染色质可及性、组蛋白标记和转录因子结合。它实现了~0.9 的高 AUROC 并集成了细胞类型特异性特征，但需要大量的计算资源进行大规模分析。这使其成为调节变异的功能注释和精细定位研究的理想选择。LINSIGHT 将进化守恒与功能基因组数据相结合，根据非编码变异的适应性影响对其进行优先级排序。它具有~0.91 的高曲线下面积（AUC）值和低计算要求，对于基于人群的研究来说非常高效且可扩展。然而，它缺乏组织特异性分辨率，并且在快速进化的基因组区域中挣扎，限制了它的多功能性。MiRLog 专门从事 miRNA 特异性分析，优先考虑影响 miRNA 结合和基因表达的变异。它整合了多个有害评分，并实现了>85%的 miRNA 变体优先级 AUC。它的低计算要求使其易于访问，但其通用性受到小型训练数据集的限制，限制了其用于 miRNA 相关研究。

每种工具都解决了 GRF 分析的不同方面：在资源允许的情况下，DeepSEA 是最具通用性的调控和编码变体分析工具，LINSIGHT 是非编码基于适应性的优先级排序的理想选择，MiRLog 在 miRNA 研究中表现出色。工具的选择取决于研究重点、计算资源和数据集范围。DeepSEA 是最强大的整体工具，可以通过集成 LINSIGHT 中缺失的功能来改进基于适应性的过滤，从而减轻其局限性。通过使用 LINSIGHT 工具进行高通量过滤，使用 MiRLog 进行 miRNA 特异性变体过滤，可以进一步优化下游分析。在应用 DeepSEA 进行详细分析之前，使用 LINSIGHT 对变体进行预过滤，可以减少计算资源需求。

任务 2: GRF 优先级

FINEMAP 设计用于使用贝叶斯概率模型在 GWAS 研究中进行因果 SNP 鉴定。它在预测因果变异方面达到了 99%的准确率，使其成为精细映射 GWAS 关联的最可靠工具之一。FINEMAP 需要适度的计算资源，可以有效地处理大型 GWAS 数据集，整合 LD 信息，并有效地探索因果配置。但是，其准确性在很大程度上取决于 LD 矩阵和 GWAS 汇总统计的质量。此外，FINEMAP 无法直接集成功能注释，例如调控或表观遗传数据，

从而限制了其范围。RegVar 识别并优先考虑与非编码致病性变异相关的组织特异性调节变异。它在调节变异预测中提供 96% 的 AUC，擅长结合基因组、表观遗传学和进化数据进行稳健的组织特异性影响预测。RegVar 支持常见和罕见变体，特别适用于非编码变体分析。然而，它是计算密集型的，并且严重依赖于高质量的组织特异性数据。此外，它难以预测对邻近地区的影响，这可能会限制其在更广泛分析中的表现。DeepWAS 将功能注释与 GWAS 结果相结合，以优先考虑 SNP 并将其与表型或细胞类型特异性机制联系起来。它实现了 90% 的 GWAS 信号恢复准确率，有效地桥接了 GWAS 和功能数据，减轻了与传统 GWAS 研究相关的多重测试负担。然而，DeepWAS 的计算要求很高，并且依赖于函数数据的完整性，例如 ENCODE。它的适用性在能力不足的 GWAS 数据集中可能受到限制，这阻碍了它在样本量较小或注释不完整的研究中的应用。CADD 使用集成学习对孟德尔和复杂特征的有害和因果变体进行优先级排序。CADD 的致病性变异预测准确率高达 98%，集成了 60 多种基因组特征，结合了功能和进化注释，以实现全面的优先级排序。它支持编码和非编码变异分析，提供可靠的基因组评分。与 RegVar 和 DeepWAS 不同，CADD 需要适度的计算资源，并且可以扩展以进行全基因组分析。然而，它不能预测组织特异性影响或长期调节作用，其中间评分缺乏可解释性。

在比较这些方法时，FINEMAP 在 GWAS 中的因果 SNP 识别方面脱颖而出，提供高精度和适度的计算要求，但仅限于没有功能注释的统计推断。相比之下，RegVar 利用多组学数据提供稳健的组织特异性调控预测，但需要更高的计算资源和高质量的组织数据。DeepWAS 有效地桥接了 GWAS 和功能注释，提供了对细胞类型特异性机制的见解，但共享 RegVar 的计算强度和对功能数据的依赖性。CADD 虽然不是特定于组织的，但可在具有中等计算需求的编码和非编码区域之间提供广泛的适用性，使其成为一般优先级排序任务的平衡工具。对于 GWAS 精细映射，FINEMAP 是最有效的。对于组织特异性分析，RegVar 无与伦比，但计算成本高昂。DeepWAS 非常适合将 GWAS 与监管注释集成，而 CADD 最适合跨多种变体类型进行全面优先级排序。工具的选择取决于研究目标、数据集可用性和计算能力。在通用风险分析中，整合 FINEMAP 和 CADD 的结果既提供了精确的因果变异识别，也提供了广泛的变异优先级。

任务 3：风险评估和致病性预测

为了通过风险分层进行早期干预，AI 算法利用各种技术，将不同位点的遗传风

险变异纳入其中，以估计个体对特定疾病的易感性。PRSice 等工具可处理基因分型和插补数据，提供稳健的 p 值，支持多遗传模型，并支持多性状分析。PolyPred 结合了 PolyFun-pred 和 BOLT-LMM 模型来计算 PRS，并提高了交叉群体的准确性。它整合了全基因组、功能知情的精细定位，并解决了群体特异性 LD 差异。PolyPred 的 AUC>0.92，在对不同人群的风险预测方面表现出高性能，尤其是在非欧洲队列中。然而，它需要高质量的 LD 面板、插补数据集和大样本量，这可能会限制资源受限设置中的可访问性。WEVar 利用基于集成的机器学习来集成来自多种致病性预测工具（如 CADD、SIFT 和 PolyPhen-2）的输出，从而生成复合致病性评分。WEVar 专注于 SNP 级别的优先级，可适应不同的数据集并减少单个预测因子的偏差。对于非编码功能预测，受试者工作特征下面积（AUROC）为 91%，它平衡了准确性和计算效率。然而，它高度依赖于预先计算的分数和集成工具的质量，因此对于现有数据库未涵盖的新变体效果较差。

在基因表达和细胞类型分析水平上，使用 VariantSpark、GRAM 和 MUSSEL 等机器学习方法来揭示遗传关联，检测表达调节变异，并进行 SNP 选择以预测各种疾病的风险。这些工具提供了对遗传变异的功能后果及其与疾病途径的关系的更深入见解。在蛋白质水平上，EVE 是一种生成模型，它通过对生物体之间的序列变异进行建模来预测没有标记数据的蛋白质变异的致病性。GRAM 专注于预测非编码变异对相关基因的表达调节作用。它使用随机森林和 LASSO 建模，整合了转录因子结合分数和细胞类型特异性基因表达谱来计算表达调节效应分数。GRAM 在调节活性预测方面具有 94% 的 AUC 性能，在基因水平非编码变异分析方面表现出色，为细胞类型调控提供高特异性。然而，它难以处理 3' 末端数据，需要基因表达谱和 SELEX 数据才能进行准确预测。它的计算量适中，需要~16GBRAM 和来自 GTEx 和 ENCODE 等资源的训练数据集。

在编码区，EVE 使用贝叶斯变分自动编码器来预测蛋白质编码错义变体的致病性。它对来自多个序列比对（MSA）的进化约束进行建模，并输出致病性评分和置信度分类。EVE 在 3219 种疾病相关蛋白中的 AUC 为~0.91，擅长识别致病罕见变异。然而，它仅限于蛋白质编码区域，需要高质量的 MSA，并且不处理剪接效应或变体组合。Xrare 使用梯度提升决策树整合遗传证据和表型数据，以确定变异的优先级，尤其是在罕见的孟德尔疾病中。将 ACMG/AMP 指南与表型相似性评分相结合，罕见变异检测

的平均准确率达到 95%。它对含噪的表型数据高度敏感，对已知和新的疾病基因都有效。然而，Xrare 主要关注编码和近端非编码区域，对深度非编码变体的效用有限。

总之，GRAM 和 PolyPred 专注于多基因和监管风险预测，而 WEVar 评估变异致病性。Xrare 擅长罕见病变异的优先排序，而 EVE 使用进化数据解释蛋白质编码错义变异。性能各不相同，PolyPred (AUC>0.92) 和 CADD (98%准确率) 在风险预测方面处于领先地位，而 EVE 和 Xrare 对特定变体类型表现出高疗效。计算要求范围从中等（例如 GRAM、PolyPred）到高（例如 RegVar、DeepWAS）。Xrare 和 WEVar 等工具的计算效率很高，可在标准设置上运行。优势包括 GRAM 的细胞类型特异性预测和 CADD 的广泛变体评分。限制源于对高质量数据集（例如 GRAM、RegVar）和受限变体范围（例如 EVE、Xrare）的依赖。选择取决于研究目标、数据集和计算能力，CADD 为全面分析提供了多功能性。最实用的通用风险分析方法整合了各种工具的结果，应用深度过滤以消除冗余，并结合了来自不同组学水平的多种预测方法的风险评分，以实现全面而稳健的变体优先级。

任务 4：风险解释和网络分析

图形网络方法为探索遗传机制、变异效应和复杂的性状关联提供了强大的功能。MecCog 采用图形模型来分析疾病机制，将生物学文献与真实世界数据相结合，以提供对基因和通路的多组学见解。MecCog 具有 90%的生物标志物识别准确率，计算效率高，只需要一台标准笔记本电脑。但是，它对领域专业知识进行结果解释的依赖可能会限制经验不足的用户的可访问性。它最适合于专注于了解疾病机制的非编码、多组学分析。DeepHisCoM 使用 DL，特别是 CNN，基于多组学数据和 GWAS 集成来分析通路。它提供精确的通路分析，准确率高达 94%，使其对识别疾病相关通路非常有效。但是，该方法需要大量的训练数据和高计算资源（>32GBRAM 和 GPU）。DeepHisCoM 非常适合在需要可靠通路见解的研究中进行非编码、SNP 和通路水平分析。SNPMap 专注于使用统计学习和自动化管道对 SNP 效应进行可视化和解释。它为 SNP 级分析提供了直观的可视化框架，准确率达到 85%，使研究人员只需最少的计算基础设施即可使用。然而，SNPMap 仅限于预先存在的注释，限制了它在新数据集或探索性分析中的应用。其计算要求低，使其适用于编码和非编码变体可视化。GREEN-DB 是一个多组学数据库，它整合了多组学信息以优先考虑监管变体，利用机器学习进行变体注释和优先级排序。它具有 90%的注释可靠性，可提供对非编码区域的深入

洞察，但需要全面且高质量的数据集，从而限制了其可扩展性。GREEN-DB 计算适中（~16GBRAM），是一种用于非编码变体优先级和多组学数据集成的多功能工具。VariantSpark 使用随机森林集成方法来识别复杂的性状关联，从而有效地处理 SNP 交互和高维数据集。它对复杂特征的预测准确率达到 85%，但由于集成方法的性质，它可能缺乏可解释性。VariantSpark 需要大量的计算资源（>32GBRAM），最适合涉及多组学数据集和跨编码和非编码区域的 SNP-基因相互作用的研究。

MecCog 和 DeepHisCoM 等工具在疾病机制解释和通路分析方面表现出色，而 GREEN-DB 和 VariantSpark 则专注于多组学整合和复杂性状分析。SNPMap 专注于直观的 SNP 效应可视化。DeepHisCoM 在精度（94%）方面领先，其次是 MecCog（90%生物标志物准确率）和 GREEN-DB（90%的注释可靠性）。MecCog 和 SNPMap 等轻量级工具可在标准笔记本电脑上运行，而 DeepHisCoM 和 VariantSpark 需要高 RAM 和 GPU。每种工具都有独特的优势：DeepHisCoM 擅长通路分析，MecCog 擅长疾病机制，GREEN-DB 擅长多组学整合，VariantSpark 擅长处理高维数据，SNPMap 擅长用户友好型可视化。限制包括 MecCog 对领域专业知识的依赖、DeepHisCoM 的广泛训练数据以及 VariantSpark 的可解释性挑战。选择取决于研究重点、计算资源和数据集可用性，其中 DeepHisCoM 适用于通路分析，MecCog 适用于疾病机制，GREEN-DB 用于非编码变体优先级。在一般风险分析中，解释遗传风险决策的影响需要多个视角。结合 MecCog 等工具（用于疾病机制洞察）、DeepHisCoM（用于通路水平分析）和 GREEN-DB（用于非编码变异优先级），提供了一种可靠的方法。将他们的结果与 SNPMap 集成以实现直观的 SNP 可视化，确保在不同的生物医学系统中进行全面的风险评估，从而推进精准医疗。

实施 GRF 分析的实用路线图和工作流程

根据我们的比较分析和排名分数，我们提出了一个实用的路线图，以指导生物信息学家选择和实施最有效的 15 种方法，如图 1 所示。我们通过为每个任务集成排名前（5-4）的工具，为在不同级别的组学分析（变异、基因、蛋白质）实施通用风险分析提供了指南。这种方法有助于丰富遗传风险分析和精准医学应用的工作流程。该路线图强调强大的数据集成、准确的预测模型和有效的结果解释，以优化患者治疗效果并推进个性化医疗保健。无花果。图 5 说明了分析 GRF 的三个主要管道：调

节分析（变异水平）、组织特异性分析（基因水平）以及功能和致病性分析（蛋白质水平）。工作流从原始变体文件开始，经过五个阶段，最终以全面的网络分析结束。这些阶段包括预处理和预注释、功能注释、遗传风险优先级、风险评估以及风险网络解释和可视化。突出显示了每种方法的计算成本：红色表示高计算需求，黄色表示中等，绿色表示低资源需求。

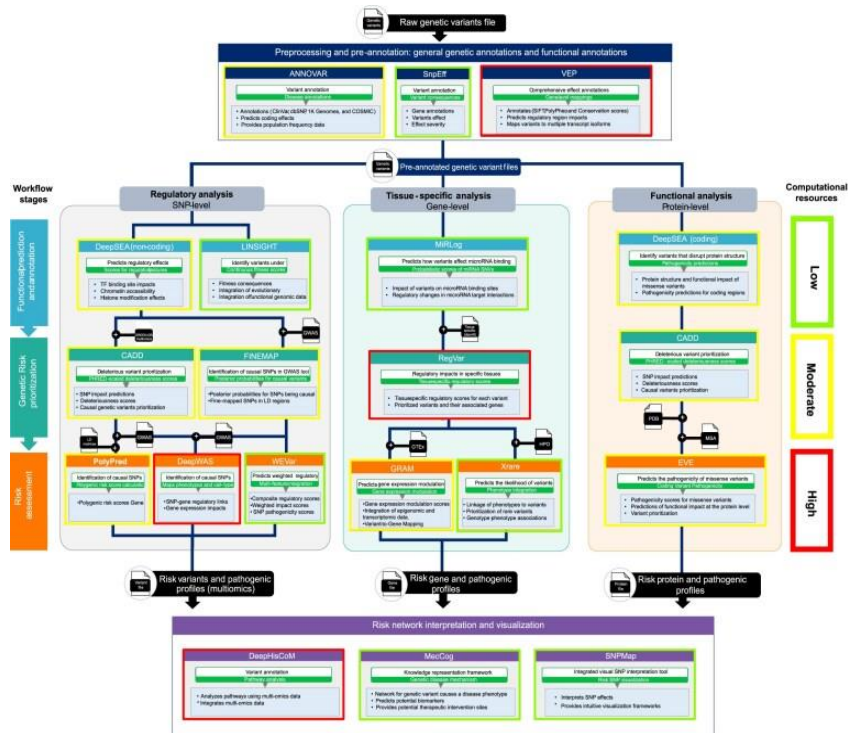


图 5 GRF 分析的实用工作流程

GRF 分析的三个层次的综合工作流程：调控分析、基因水平功能预测和蛋白质水平功能注释。这些工作流程说明了高级 AI 工具、计算方法和多组学数据集的集成，以优化对精准医疗的洞察。

变体级别：监管分析工作流程

在监管工作流程中，DeepSEA 用于预测监管效应，例如转录因子结合位点影响和染色质可及性。LINSIGHT 根据种群和进化环境中的适应性影响来细化和过滤变异，专注于分析位于非编码区的遗传变异。GREEN-DB 提供丰富的多组学数据，提供有关受控基因、组织和相关表型的详细信息，以解释变异在多组学水平上的影响。对于风险优先级排序，CADD 和 FINEMAP 识别致病 SNP 和有害变体，并提供其潜在影响的排名。在风险评估中，PolyPred、DeepWAS 和 WEVar 用于对单个变体的风险贡献进行评分和比较。在最后阶段，使用 DeepHisCoM、MecCog 和 SNPMap 分析整合的变体谱。

这些工具更深入地了解了变异在疾病机制中的重要性以及变异影响的多层次途径，从而全面解释了它们在遗传风险和疾病进展中的作用。

基因水平：组织特异性分析

MiRLog 用于预测 microRNA 区域中变异的影响及其对 microRNA-靶标相互作用的影响。在优先排序阶段，RegVar 评估 SNP 在特定组织中的调节影响。对于风险评估，GRAM 通过整合表观基因组和转录组数据（例如 GTEx）来预测基因表达调节，而 Xrare 通过使用人类表型本体论（HPO）术语结合基因型-表型关联来预测变异致病性的可能性，从而将表型与罕见变异联系起来。此外，Xrare 还提供基因型-表型数据可视化，以促进解释。该阶段的集成特征为风险评估工具提供信息，以便在基因水平上进行更全面的解释。该工作流程可识别关键基因水平的关联，并加深我们对多因素疾病中变异致病性的理解。

蛋白质水平：功能和致病性分析

该工作流程侧重于分析编码区中的错义和其他变异，以评估它们对蛋白质结构和功能的影响。DeepSEA 用于识别变异并预测由编码变异引起的结构破坏及其致病性。在优先级排序阶段，采用 CADD 来识别风险变异及其对蛋白质功能的影响。对于风险评估，生成模型 EVE 利用源自 PDB 的蛋白质结构来预测变异对蛋白质结构和功能的致病性影响。在最后阶段，将风险评估的结果与 DeepHisCoM 和 MecCog 相结合，以促进通路分析和构建包含有害变异的蛋白质网络，为疾病机制提供有价值的见解。这种全面的方法有助于识别潜在的生物标志物并提出治疗靶点，从而推进精准医疗策略。

AI 优化风险因素在精准医疗中的应用

本节重点介绍 AI 优化型 GRF 在精准医疗中的实际应用和实际影响。通过研究具体示例和案例研究，我们重点介绍了这些先进的 AI 方法如何通过提供针对个体患者需求的定制治疗来增强精准医疗。在表 3 中，表 1 和表 2 中选定的 AI 方法已通过临床或实验室测试进行了实验验证，强调了 GRF 对疾病预测、预防和个性化治疗策略产生重大影响的关键医疗应用。本节展示了将 AI 增强的 GRF 整合到临床实践中的切实好处和潜力，展示了精准医疗的有效性。此外，图 1 中的信息图 16 介绍了 AI 驱

动的应用程序，这些应用程序利用生物医学资源来确定风险因素，从而进一步增强精准医疗。

表 3. AI 方法在精准医疗中的整合

医疗任务	疾病	相关遗传因素	AI 方法	PMID
患者分层筛查	COVID-19，哮喘， 心血管和肾脏疾病	ACE2、 TMPRSS2、HLA- DQB1	EnsemblVEP ，DSNetwork	38855114、 34880287
高级诊断	复杂疾病，癌症， 类风湿性关节炎	多组学、 MAPK9、TNF、 JAK 抑制剂	Geneformer 、PASNet	34579788
风险评估	心血管，2 型糖尿 病，癌症	PCSK9、APOB、 LDLR、TCF7L2、 PANK1	PolyPred、 CADD、 DeepSurv	37925478、 3825622438872215
生物标志物预 测	冠状动脉疾病，癌 症	安里尔、 ARRDC3、SHBG	SNPnexus、 RegBase、 DeepSurv	36597873、 3041224138396861
药物发现和开 发	癌症、阿尔茨海默 病、囊性纤维化、 家族性高胆固醇血 症	PIK3CA、 BRCA1、BRCA2、 APOE、CFTR、 PCSK9、LDLR	CADD、 SNP2TFBS、 DeepVariant 、PASNet	38822093、 33834176
药物基因组学	克罗恩病，高血压	HLA、SEPHS1、 GFRA2	SNPnexus、 RegBase、 DeepSurv	36597873、 3041224138396861
发现风险基因	阿尔茨海默病，免 疫介导的疾病	PRL、PLCG2、 ABI3、TREM2、 LZTFL1、 ATP2B1、NAV3、 ITSN1、MARK2、 SCAF1HNRNPUL2	DeepSEA、 GWAVA、 DeepVariant	30820047、 28714976、 3881651435982159
了解疾病机制	阿尔茨海默病，免 疫介导的疾病	GRAMD1B、 FUT2、 ADAMTS4、 HBEGF、WNT3、 TSPAN14、 DHODH、ABCB9、 TNIP1	CADD、 SNP2TFBS、 MecCog	37841839、 36648426

表 1. 在精准医疗中优化 GRF 的 AI 方法列表

方法	方法论	输入	输出	生物医学应用	临床验证	实现	PMID
Deep Sea	卷积神经 网络	基因组序列	预测的染色质 特征、转录因	鉴定非编码变 异的功能和染	是（实验验 证）	平台	26301843

方法	方法论	输入	输出	生物医学应用	临床验证	实现	PMID
			子结合和变异效应	色质效应			
Deep WAS	卷积神经网络	基因型和调控信息（染色质特征、细胞类型）	与表型相关的 SNP	精细绘制 GWAS 关联并确定疾病机制	否（仅计算验证）	Github	32012148
变体 Spark	随机森林	基因型数据、复杂表型数据	与复杂表型相关的变异	鉴定与复杂表型相关的基因组变异	否（仅计算验证）	Github	32761098
Xrare	梯度提升决策树	基因序列、临床表型	致病变异预测	优先考虑罕见病相关变异	是（真实的临床数据集）	不适用	30675030
WEVar	多种评分方法的加权集成	非编码调节变体（SNV）	功能和致病性预测	预测非编码变体的监管影响	是（多个风险位点验证）	Github	34021560
MecCog	图形符号	遗传变异数据、生物医学文献	疾病机制的综合表示、生物标志物的识别、治疗干预部位	了解遗传疾病机制并确定潜在的生物标志物和治疗靶点	是（真实的临床数据和专家反馈）	网页服务器	34117883
Deep SNP	具有基于注意力定位的神经网络	全基因组 SNP 阵列数据	在大基因组窗口中检测断点	识别癌症和其他疾病中与疾病相关的 CNV	否（仅计算验证）	Github	30585743
RegVar	深度神经网络	非编码基因组变异、基因表达数据	组织特异性调控影响预测，优先变异	优先考虑各种组织中的非编码调节变异	否（仅计算验证）	网页服务器	34973416
SNPMap	自动化管道、数据集成	文本挖掘的 SNP 相关信息，GWAS 目录	SNP 的视觉解释，优先变异	精准医学中单核苷酸多态性（SNP）的解读和可视化	否（仅计算验证）	不适用	36061173
尤金	卷积神经网络	基因组序列	调控序列的预测分析，序列特征的识别	预测监管活动并理解顺式监管代码	否（仅计算验证）	Github	38177592
GREEN-DB	各种 ML 和统计方法	全基因组测序数据	带注释的监管变体，优先变体	优先考虑非编码监管变体	是（实验验证）	Github	35234913
PolyPred	PolyFun-pred 和 BOLT-LMM 的组合	基因分型和插补数据	PRS，跨群体预测精度	遗传风险预测和解决连锁不平衡差异	是（实验验证）	Github	35393596
CADD	SVM、集成方法	遗传变异（SNV、插入缺失）	指示变体有害性的缩放和原始 CADD 分数	确定因果变异的优先级并确定孟德尔疾病的有害变异	是（实验验证）	Github	30371827
PRS-ML 系列	结合 SNP 选择和 XGBoost	遗传变异（SNP）、PRS）	PRS	预测复杂疾病的遗传风险	是（实验验证）	Github	35995843

方法	方法论	输入	输出	生物医学应用	临床验证	实现	PMID
基因形成者	变压器型号	单细胞转录组	特定于上下文的预测	发现关键网络调节因子、候选治疗靶点	是（实验验证）	不适用	37258680
克	LASSO, 随机森林	非编码变异、TF 结合评分、基因表达数据	非编码变体的表达调节作用	以细胞类型特异性方式预测非编码变体的分子效应	是（实验验证）	Github	31469829
DSNetw rk	集成多个预测分数	遗传变异	优先变体	确定变体预测的优先级并可视化	是（实验验证）	不适用	32010198
MiRLog 和 dbmiR	Logistic 回归	microRNASNV	有害性分数、功能注释	对 miRNA 序列变体进行优先排序和注释	是（实验验证）	Github	35583122
RegBase 数据库	梯度树提升	非编码调节变体（SNV）	致病和癌症驱动因素预测	功能注释、致病性预测	是（实验验证）	Github	31511901
IMPUTE2	隐马尔可夫模型，马尔可夫链蒙特卡罗	基因组变异	未分型 SNP 的插补基因型	GWAS 中的插补	是（真实的临床数据和专家反馈）	网页服务器	34117883

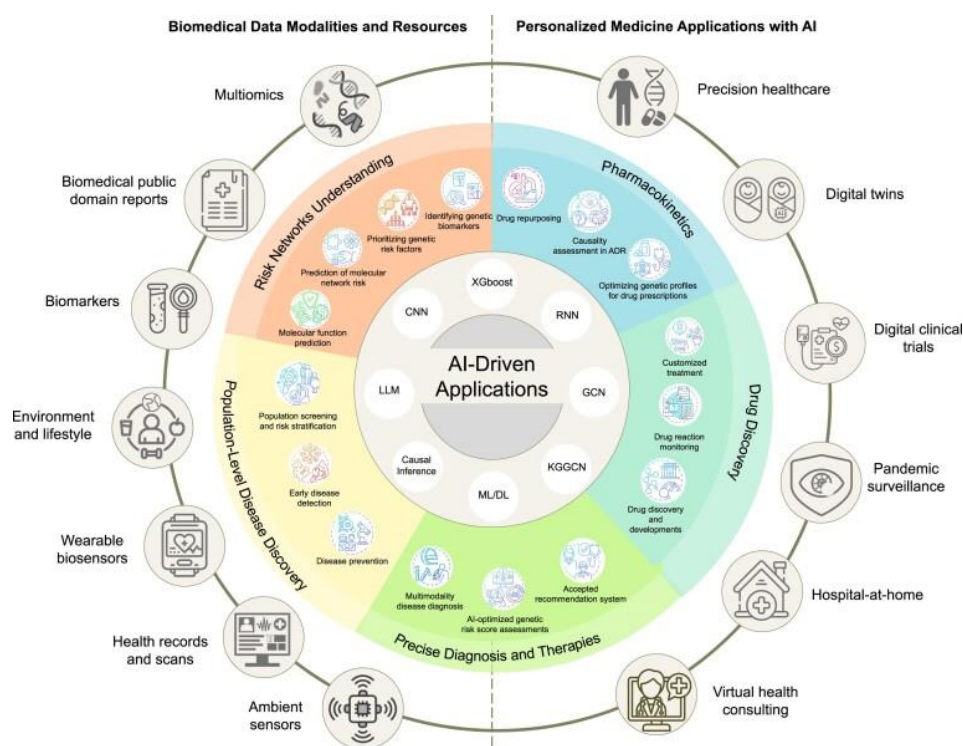


图 6 来自个性化医疗中不同模式和资源的风险因素的 AI 驱动应用

该图说明了来自各种生物医学模式的 GRF 在个性化医疗中的 AI 驱动应用。它强调了多个数据源的集成，例如多组学、健康记录、可穿戴生物传感器、环境和生活方式数据、虚拟健康咨询、数字孪生、大流行监测、家庭医院计划和数字临床试验，

以实现精准医疗保健。这些 AI 技术应用于多种医疗应用，包括人群筛查、风险分层、早期疾病检测、疾病预防、多模式疾病诊断和 AI 优化的遗传风险评估。

用于个性化医疗应用的生物医学数据综合分析

精准医学的基石在于各种生物医学数据模式的整合。例如，基因组数据提供了患者基因组的详细概况，电子健康记录提供了对过去诊断和治疗反应的见解，可穿戴生物传感器数据跟踪实时健康趋势。人工智能算法分析这些庞大的数据集，以发现可能逃避人类解释的模式，从而在诊断和个性化治疗计划方面取得重大进展。通过整合生物医学资源，人工智能有助于更准确地识别疾病的根本原因，并根据患者的基因图谱定制治疗方法，从而提高疗效。

此外，将人工智能与数字孪生（患者的虚拟表示）和药代动力学数据（描述体内的药物行为）相结合，可以模拟治疗结果并优化针对个体遗传图谱的药物剂量，从而增强个性化医疗保健的预测建模。这种方法扩展到远程健康咨询和家庭医院服务，其中持续的患者监测和护理是远程管理的。这些 AI 方法将不同的生物医学资源转化为可作的应用程序和个性化治疗，体现了个性化医疗的潜力。

精准的疾病检测和预防

GRF 可以在临床适应症出现之前识别出某些疾病风险较高的个体，从而通过人群分层实现个性化医疗保健。遗传风险较高的个体受益于更频繁和更深入的筛查，优化了资源分配并加强了早期疾病检测。在神经退行性疾病中，GRFs 通过分析 APOE 基因等基因多态性来促进阿尔茨海默病的早期识别。DL 模型根据 APOE 多态性预测阿尔茨海默病的进展，并将 AI 与遗传和神经影像学数据相结合，有助于读取脑部扫描以识别阿尔茨海默病基因。AI 技术通过识别 GRF 并整合来自颈动脉超声成像和电子健康记录的数据来增强心血管健康。对于家族性高胆固醇血症，PCSK9 和 APOE 等遗传标志物指导早期干预。AI 还通过自动化细胞形态学、细胞遗传学和免疫表型来增强造血疾病诊断，从而提高准确性并减少专家工作量。在 2 型糖尿病中，AI 算法利用 TCF7L2 等 GRF 和生活方式数据来制定先发制人的策略。AI 整合了遗传和医学数据来了解早期 T2D 调控网络，并使用 GWAS 进行风险预测。在肾脏疾病中，ML 遗传编程算法通过对不平衡数据集的适应函数增强风险识别来改进早期慢性肾脏病的诊断。在肿瘤学中，基因筛查针对遗传性癌症，例如遗传性乳腺癌和卵巢癌综合征（HBOC）和 Lynch 综合征，使用 BRCA1 和 BRCA2 突变进行早期发现和提高生存率。对于肝细

胞癌，早期检测涉及循环肿瘤 DNA 和 microRNA 等生物标志物。AI 增强了癌症筛查并支持生活方式的改变以降低风险，从而提高了乳腺癌、宫颈癌和前列腺癌的检测率。

精确诊断和改善患者预后

一些医疗应用受益于将 GRF 与 AI 结合使用，提供高级诊断并改善患者预后。在癌症研究中，人工智能识别对定制治疗有反应的基因组生物标志物，从而提高预后能力。在药物靶点鉴定中，GRF 可以精确定位疾病相关基因或蛋白质，这对药物开发至关重要。例如，Cox-PASNet 在多形性胶质母细胞瘤和卵巢癌中识别临床上可作的遗传性状，检测关键通路和基因，如 PI3K 级联反应和 MAPK9，这些对肿瘤进展和潜在治疗至关重要。此外，AI 和 ML 分析基因数据，用于副作用较少的个性化疗法，适用于 IBD、结肠癌、AML 和乳腺癌等疾病。对于心血管疾病和 2 型糖尿病，AI 有助于患者护理并预测疾病进展，改善结果和成本效益。例如，模型 DIABLO 结合了遗传和生活方式数据来预测糖尿病的发作和进展。同样，在肾脏疾病中，人工智能平台 KidneyIntelX 使用 ML 算法分析来自电子健康记录和血液测试的数据，以预测和管理疾病进展，预防器官衰竭和个性化护理。人工智能的最新进展通过整合多模态数据（如病历、成像和新型生物标志物）改变了阿尔茨海默病的诊断，为疾病风险评估提供了一种全面的方法。例如，DeepAD 分析 MRI 扫描和 APOE 基因中的遗传数据，以改善阿尔茨海默病的早期诊断和进展监测。这些技术还支持创建临床决策支持系统，分析患者数据以进行个性化诊断和治疗。

靶向治疗和药物发现

将 GRF 与 AI 一起使用显著推进了各种疾病的治疗策略。例如，DeepCPI 用于鉴定作用于代谢紊乱药物靶点的化合物，并筛选治疗高血压的药物。在 RA 中，遗传变异指导靶向治疗的发展，如 TNF 和 JAK 抑制剂。阿尔茨海默病通过利用 APOE 基因中的遗传变异来指导针对淀粉样蛋白- β 斑块积累的治疗，从而受益于 AI。在囊性纤维化中，治疗是根据特定的囊性纤维化跨膜传导调节因子（CFTR）基因突变量身定制的，肿瘤学利用 GRF 对肺癌、乳腺癌和慢性髓性白血病进行靶向治疗。人工智能在神经遗传学中至关重要，它通过注释和优先考虑遗传因素来增强对阿尔茨海默病等复杂疾病的理解，从而为治疗靶点提供信息。在药物基因组学中，AI 注释影响药物代谢的遗传标记，以优化药物剂量并减少不良反应。此外，AI 根据 GRF 定制抗凝

治疗剂量，从而降低与治疗相关的风险。在癌症研究中，人工智能创新包括检测异柠檬酸脱氢酶（IDH）基因突变和分析肺肿瘤病理图像的无创方法，从而提高治疗和诊断的精度。AI 的作用延伸到识别炎症性肠病和结肠癌等疾病的遗传原因，展示了其在药物靶点识别和治疗开发方面的多功能性。

药代动力学和药物反应

GRF 是精准医学不可或缺的一部分，可实现靶向治疗和精细药物发现。例如，HLA 变异显著增加了 RA 和狼疮等疾病的风险。囊性纤维化的早期基因检测可识别有风险的个体，从而减少潜在的器官衰竭。妙佑医疗国际使用基因组学和临床数据预测对甲氨蝶呤的反应，从而使用 AI 和药物基因组学来定制 RA 药物。遗传图谱为药物靶标识别提供信息并优化剂量，以增强治疗反应并最大限度地减少不良反应。药物代谢基因（ADME）的变异表明需要个性化剂量来减轻副作用并提高疗效。此外，遗传标记可识别药物不良反应风险较高的个体，例如对卡马西平等药物的严重过敏反应。人工智能与遗传数据的整合提高了药物代谢预测和治疗优化的精度。它有助于预测个体药物代谢变化并识别不良反应的标志物，从而提高患者安全性和治疗效果。

了解疾病机制和分子风险网络

AI 与 GRF 的整合显著增强了我们各种疾病的分子机制的理解。人工智能驱动的遗传风险变异分析提供了对广泛疾病易感性的见解，有助于风险评估和制定个性化的预防策略。AI 增强的 GRF 对于理解疾病机制、分子网络和进行功能基因组学研究至关重要，这些功能基因组学研究决定了研究重点并加深了我们对疾病发展的了解。人工智能驱动的工具，如预测网络，促进了基因相互作用网络的分析，丰富了功能基因组学研究，并有助于理解基因相互作用和疾病机制。将 GRF 与 AI 相结合，增强了疾病风险的预测模型，促进了早期干预和高效的医疗保健分配。STRING、DAVID、GeneMANIA 和 Cytoscape 等工具丰富了分子风险网络和疾病机制的知识，推动了疾病研究和个性化治疗。人工智能的利用帮助研究人员和临床医生揭示了遗传元件如何塑造疾病通路，这是制定精确、有针对性的治疗的一个重要方面。这两类遗传因素在精准医学中都发挥着关键作用，它为个体对一系列健康状况的遗传易感性提供了重要的见解。基因组学中的 AI 应用，例如通路富集分析和分子风险网络构建，揭示了受风险变异影响的基因，并有助于了解疾病发展中被破坏的细胞过程，从而有助于精准医疗。

讨论和结论

精准医疗的曙光标志着医疗保健领域的变革时代，其特点是从通用治疗方法转变为高度个性化的治疗策略。精准医疗在很大程度上依赖于来自各种资源的大量生物医学数据，包括可穿戴技术、具有不同模式的医疗数据存储库、多组学、表观遗传学和病因学数据。这种范式转变是通过对生物医学大数据的准确表示和综合分析来实现的，以提取有意义的见解。AI 在此过程中发挥着关键作用，它使用生成式 AI、自然语言处理、表示学习、因果模型、强化学习和计算机视觉等高级算法捕获特征并集成 GRF。本文对 GRF 及其在增强精准医疗中的作用进行了明确的定义。它展示了如何利用尖端的人工智能模型有效地将 GRF 集成到下游生物医学任务中。对前 20 种经过验证的方法进行了深入分析和批判性回顾，评估了它们的性能、应用、优势和局限性。全面比较确定了遗传风险评估任务的最有效工具，最终提出了跨组学水平利用前 15 种工具的路线图。该路线图强调强大的数据集成、准确的预测模型和有效的结果解释，以推进个性化医疗。本文还探讨了 AI 在精准医疗中的实际应用，通过信息丰富的图表提供见解，说明集成 AI、GRF 和个性化医疗保健的影响。它强调了 AI 驱动的 GRF 分析中的道德考虑因素，并提供了一个全面的框架，用于将高级计算工具整合到临床和研究工作流程中。通过解决变异解释、多组学整合和风险评估方面的挑战，本研究为精准医学的未来创新奠定了基础，强调了对可解释、可扩展的 AI 模型和公平获得基因组技术的需求。利用广泛的生物医学公共资源，人工智能促进了新的进步，为提高结果并提供针对个人基因图谱量身定制的个性化解决方案的医疗保健范式铺平了道路。

***注：原文和译文版权分属作者和译者所有，若转载、引用或发表，请标明出处。**