

AI 模型首次揭示全球癌症生存率影响因素

本报讯 近日,一项发表于《肿瘤学年鉴》的研究利用人工智能(AI)工具,探究了全球癌症生存率存在巨大差异的原因。

科学家首次将机器学习用于确定与全球几乎所有国家癌症患者生存率相关的因素。通过分析 185 个国家的癌症数据和卫生系统信息, AI 模型揭示了诸如放疗可及性、全民健康覆盖水平、经济实力等因素与各国癌症患者生存率紧密相关。

该研究不局限于宽泛的比较,而是指出了哪些具体的政策调整或系统改进能够对癌症患者生存率产生最大影响。研究团队同时开发了一个在线工具,用户可查看一个国家国民财富、放疗可及性、全民健康覆盖水平等因素与癌症治疗结果之间的关联。

“我们希望创建一个可操作的、基于数据的框架,帮助各国确定最具影响力的政策杠杆,以降低癌症死亡率并缩小公平差距。”研究负责人、美国纪念斯隆-凯特琳癌症中心(MSK)的 Edward Christopher Dee 强调了该研究的重要性。

在研究中,团队利用机器学习技术,对交互式平台——全球癌症观察站(GLOBOCAN 2022)185 个国家的癌症发病率和死亡率数据进行了分析,并将其与世界卫生组织、联合国机构、放射



AI 揭示了各国癌症患者生存率差异巨大的关键原因。
图片来源: Shutterstock

治疗中心目录等卫生系统数据相结合。最终数据集包括卫生支出占国内生产总值(GDP)的百分比、人均 GDP、全民健康覆盖水平、病理服务可及性、人类发展指数、每千人拥有的医疗人员数量、患者直接支付医疗费用比例等。

该研究采用的机器学习模型由论文第一作者、MSK 的 Milit Patel 开发。“我们选择使用机器学习模型,是因为它能帮助我们生成每个国家的具体估计值及相关预测。尽管人口层面数据存在局限性,但我们希望这能为全球癌症系统规划提供指导。”Patel 说。

该模型计算了癌症死亡率与发病率比(MIR),这代表了导致死亡的癌症病例比例,可作为衡量一个国家癌症治疗效果指标。为了展示单个因素如何影响估计值,研究人员使用了 SHAP 模型预测方法,该方法可量化每个变量对模型预测结果的贡献。

“我们的模型不只是描述差异,还为

政策制定者提供了切实可行、基于数据的路线图,明确指出每个国家的哪些医疗系统投资能产生最大影响。”Patel 表示,随着全球癌症负担加重,这些见解有助于各国以最公平、最有效的方式优先配置资源,缩小患者生存差距。国际组织、医疗服务提供者和倡导者也可以利用这个工具,投资重点领域,尤其是在资源有限的环境中。

研究结果表明,对癌症患者生存率影响较大的因素因国家而异。模型显示,巴西全民健康覆盖水平与 MIR 的改善呈强正相关性。而其他因素,如病理服务可及性、每千人拥有的医疗人员数量目前发挥的作用较小。研究人员表示,这意味着巴西优先考虑提升全民健康覆盖水平,获益最大。

在波兰,情况则有所不同。该国的放疗服务可及性、人均 GDP 及全民健康覆盖水平对癌症治疗结果影响很大。这表明波兰近期扩大医疗保险和

医疗服务可及性带来的效果强于总体卫生支出。

相较上述国家,日本、美国和英国呈现出更广泛的模式,几乎所有卫生系统因素都与更好的癌症预后有关。在日本,放射治疗中心密度的影响最为突出;在美国和英国,人均 GDP 影响最大。这些发现为每个国家的政策制定者指明了可能取得最大成效的方向。

但 Patel 强调,这并不意味着其他影响较小的因素不值得关注或可以被忽视,因为上述结果是根据模型和现有数据得出的结论,存在一定局限性。比如该分析依赖于国家层面的数据而非个体患者的记录,数据质量差异很大,尤其是在许多低收入国家,而且国家层面的趋势可能掩盖了国内的差异。此外,该研究无法证明专注于某个特定因素会获得更好的癌症治疗结果,只是表明此类努力与改善的结果有关联。

即便如此,该研究仍为确定行动的优先次序提供了参考。“随着全球癌症负担加重,这一模型有助于各国在资源有限的情况下实现效益最大化。它将复杂的数据转化为政策制定者易于理解且可操作的建议。”Dee 说。

(徐锐)

相关论文信息:<https://doi.org/10.1016/j.annonc.2025.11.014>

深度思维新模型可预测 DNA 变异影响

本报讯 深度学习模型 AlphaGenome 于近日在《自然》发布。该模型能预测 100 万碱基对的 DNA 序列的功能,以及 DNA 序列变异如何影响不同的生物过程。此外,AlphaGenome 还能用于理解遗传疾病、改进基因检测,并为开发新疗法提供信息。

基因变异会影响生物学过程并可能引发疾病,但理解 DNA 序列变化如何影响其功能一直是个难题。约 98% 的变化发生于非编码区域,即不编码蛋白质但影响基因表达的 DNA 区域,使得预测其影响变得非常困难。

要解决这一问题需要计算模型,现有方法在序列长度和预测强度上必须做出取舍,但 AlphaGenome 却能在长 DNA 序列中做出高分辨率预测。

谷歌深度思维(DeepMind)的 Ziga Avsec、Natasha Latysheva、Pushmeet Kohli 和同事展示了

AlphaGenome 的能力。这一模型通过人类和小鼠基因组的训练来学习 DNA 序列如何影响不同的生物过程。AlphaGenome 可以即时预测 5930 种人类或 1128 种小鼠遗传信号,这些信号与特定功能有关,如基因表达、剪接(基因组的切割与重组)和蛋白质修饰。在 26 项变异效果预测评估中,有 25 项的结果与现有顶尖模型表现相当或更优。作者指出,该模型优势在于能同时对多种遗传信号和生物学结果进行多重预测。

作者写到,进一步改进这一工具或能拓展其应用,例如增加涵盖的物种或拓展模型识别非编码序列的范围。他们总结说,AlphaGenome 有望深化对 DNA 序列变异引发的复杂生物学结果的理解。

(赵熙熙)

相关论文信息:

<https://doi.org/10.1038/s41586-025-10014-0>

本报讯 瑞士苏黎世大学的 Christian C. Ruff 团队提出了自适应心智化的神经特征。近日,相关研究成果发表于《自然-神经科学》。

心智化能力,即推断他人的情绪和意图的能力,对人类的社会互动至关重要。心智化能力在各种脑部疾病中会受到损害。虽然以往的神经科学研究主要集中在静态心智化策略上,但研究团队对大脑如何在不同时刻适应性选择策略却知之甚少。

研究团队通过互动策略游戏中的计算建模,发现尽管存在显著的个体差异,但大多数参与者能根据对手不断变化的复杂程度调整自己的策略。基于模型的功能性磁共振成像分析,研究团队确定了一个分布式的大脑网络,其活动和连接可跟踪这种心智化信念的适应过程。人们对他人复杂程度的看法的更新程度,可以从神经活动的样本中预测,从而提供了自适应心智化的神经特征。

(柯讯)

相关论文信息:

<https://doi.org/10.1038/s41593-026-02219-x>

研究提出自适应心智化的神经特征