

DOI: 10.13376/j.cblls/2022106

文章编号: 1004-0374(2022)08-0957-08



高亦博, 国家自然科学基金委优秀青年科学基金项目获得者, 国家重点研发计划首席科学家, 中国医学科学院北京协和医学院博士生导师。现为国家癌症中心/国家恶性肿瘤临床医学研究中心/中国医学科学院肿瘤医院转化医学实验室主任, 胸外科主治医师、副研究员、实验室主管, 深圳医院中心实验室主任。

清华大学、北京协和医学院临床医学、肿瘤学博士, 美国杜克大学医学中心访问学者。主要开展食管癌、肺癌等胸部肿瘤研究, 多组学测序与基因功能研究发现食管癌表观遗传基因的突变, 免疫和预后相关分子分型; 构建肺肉瘤样癌、食管小细胞癌的首个多组学分子分型。在 *Nat Genet*、*Lancet Oncol*、*Nat Commun* 等发表第一/通讯作者 SCI 论文 40 余篇。

受邀担任 27 个 SCI 期刊和美国医疗信息学会 AMIA 同行评审专家; 在中国抗癌协会整合肿瘤学分会、中国医疗保健国际交流促进会健康大数据和数字化医疗分会、中国医药生物技术协会基因检测技术分会等担任委员。担任中国科学院大学、北京协和医学院任课教师。作为主要完成人, 获得国家科技进步奖二等奖和教育部科技进步奖一等奖。

## 人工智能在肿瘤领域的作用与潜力

郭昕哲<sup>1</sup>, 高亦博<sup>2\*</sup>, 赫捷<sup>2</sup>

(1 中国医学科学院肿瘤医院, 国家癌症中心, 转化医学实验室, 北京 100021;

2 中国医学科学院肿瘤医院, 国家癌症中心, 胸外科, 北京 100021)

**摘要:** 科技的进展使得人们有越来越有效的手段应对癌症。人工智能技术是高速发展的计算机科学技术的结晶, 其在肿瘤相关领域的应用有望进一步改善肿瘤的诊断和治疗。人工智能技术能够通过病灶识别进行肿瘤早期筛查, 有效预防肿瘤疾病; 能够通过提取病灶特征并加以演算处理的方式, 无创鉴别肿瘤的良恶性; 还能够辅助计算体内肿瘤负荷, 判断肿瘤的突变状态。然而, 人工智能算法在临床实践中的应用尚不成熟, 在医生、患者中的接受度均有待提高, 并且在肿瘤诊疗过程中存在许多医学伦理学问题。该文分别从肿瘤筛查、突变判断、病灶分析、疗效评估和患者随访等方面介绍了人工智能在其中发挥的作用和展现的潜力, 并讨论了目前的人工智能肿瘤辅助系统亟待解决的若干问题。

**关键词:** 人工智能; 肿瘤学; 癌症; 深度学习; 精准医学

**中图分类号:** TP18; R730.4

**文献标志码:** A

## Usage and potential of artificial intelligence in oncology

GUO Xin-Zhe<sup>1</sup>, GAO Yi-Bo<sup>2\*</sup>, HE Jie<sup>2</sup>

(1 Laboratory of Translational Medicine, National Cancer Center/National Clinical Research Center for Cancer/Cancer Hospital, Chinese Academy of Medical Sciences and Peking Union Medical College, Beijing 100021, China;

2 Department of Thoracic Surgery, National Cancer Center/National Clinical Research Center for Cancer/Cancer Hospital, Chinese Academy of Medical Sciences and Peking Union Medical College, Beijing 100021, China)

收稿日期: 2022-06-15

基金项目: 科技创新2030—重大项目(2020AAA0109500); 广东省重点领域研发计划(2021B0101420005)

\*通信作者: E-mail: gaoyibo@cicams.ac.cn

**Abstract:** Advances in technology during the past few decades have enabled people to have more and more effective ways to cope with cancer. The emergence of artificial intelligence technology results from the rapid development of computer science and technology, and the application of artificial intelligence in oncology is expected to further improve the diagnosis and treatment of tumors. Artificial intelligence technology can perform tumor early screening through lesion identification, in order to prevent tumor diseases; it is a non-invasive method of identifying benign and malignant tumors by extracting their characteristics and performing arithmetic processing; it can also assist in the calculation of tumor load *in vivo* and detect the mutation status of tumors. However, there are still problems to be solved with the application of artificial intelligence in clinical practice. The acceptance among doctors and patients needs to be improved, and there are ethical issues during tumor diagnosis and treatment. This article reviews the current role and potential of artificial intelligence in tumor early screening, mutation detection, prognosis analysis, therapeutic evaluation, and patient follow-up, and discusses major challenges of the usage of artificial intelligence in clinical practice.

**Key words:** artificial intelligence; oncology; cancer; deep learning; precision medicine

近年来,人工智能为医学带来了广泛而深远的影响。肿瘤多学科诊疗医学影像分析辅助系统能够辅助基层医生提高诊疗水平,节省医生培训费用。多种应用模型对肿瘤的早期发现、精准诊断、复发监测,还可将肿瘤防治关口前移,避免对低风险患者的过度干预,降低病患医疗费用,避免医疗资源浪费。同时,随着社会老龄化的加重,医疗系统需要从一次性、短期的疾病治疗向长期健康和疾病管理转变,长期的健康和疾病管理需要更加先进的技术来降低成本提升效率,人工智能为肿瘤影像分析等领域带来的便利正好满足了这一点需求。本文将就人工智能在肿瘤诊治过程中各个阶段的研究进展和发展前景作一综述。

## 1 人工智能在肿瘤诊治过程中的作用与潜力

人工智能在肿瘤诊治过程中的早期筛查、病灶分析、突变判断、疗效预测和患者随访等方面均有较大的应用潜力。在肿瘤早期筛查方面,成熟的计算机辅助探测(CAD)技术在面对大量的肿瘤筛查和临床诊断影像时,可以减少人类医生的读片疏忽,并防止漏诊的发生;而使用下一代测序(NGS)技术与微小残留病变(MRD)分析技术的新型人工智能早筛工具能够更早地检查到肿瘤病灶。肿瘤特征提取技术覆盖了病灶分析与突变判断,包括病灶与正常组织的自动分割、影像学参数测量、良恶性鉴别诊断和肿瘤分期,乃至使用影像学表型推断肿瘤的基因型和生物学特征,预测临床结局。在疗效预测方面,人工智能能够计算肿瘤负荷并评估治疗反应,通过自动配准和比较,对肿瘤的演变和可能发生的耐药进行追踪,对肿瘤侵袭转移和复发进行预测。

接下来将就以上几类临床任务阐述最新的研究进展。在肿瘤随访方面,人工智能是医生们潜在的有力助手,能够辅助评估病灶风险,指导后续治疗方案。

上述每个方面都涉及医生与患者做出的关键决策,并且人工智能能够在其中形成辅助作用<sup>[1]</sup>。同时,涉及到的技术细节将根据诊治阶段和癌症亚型而有所不同。人工智能在这些方面有潜力应用于大量未满足的需求,并且有潜在的大型数据集支持。在使用监督机器学习的条件下,这些数据集需要强大而准确的注释来形成一个可靠的“金标准数据集”,人工智能系统可以在该数据集上进行训练。

### 1.1 肿瘤筛查:人工智能通过病灶识别进行肿瘤早期筛查,形成有效肿瘤预防

目前,人们尚未找到能够有效检测肺癌早期病变的方法。美国全国肺癌筛查试验(NLST)的结果显示,通过低剂量CT(LDCT)筛查吸烟者或有吸烟史的高危人群可降低20%的全因死亡率<sup>[2]</sup>。与发现时有临床症状的患者相比,早期肺癌患者手术的机会更大,总生存期也有优势,因为后者的疾病通常已进展到中晚期<sup>[3]</sup>。虽然免疫检查点抑制剂或靶向治疗可以有效改善部分患者的生存状况,但并非所有患者都能从中受益;因此,早期肺癌筛查可以有效提高患者的生存率,并尽量避免高强度治疗。根据NSLT,每年LDCT已成为肺癌高危人群的推荐筛查方案,其优先级仅次于一级预防(戒烟)。LDCT可有效降低肺癌患者的死亡率,对戒烟但高危的人群尤其有效。尽管NLST已经确定了使用LDCT进行早期肺癌筛查的全因死亡率的获益情况,但它仍然存在一些局限性,计算机分析技术可

以改善这一缺点<sup>[2]</sup>。

计算机辅助检测 (CADe) 诞生于 20 世纪 90 年代。这种用于解释乳房 X 线照相术的计算机化系统已在临床上使用了数十年<sup>[4-7]</sup>。乳腺癌影像筛查的处理存在很多问题, 如结构噪声影响正常组织背景掩模、图像搜索模式不完整、病灶太小或太复杂、医生疲劳、注意力分散、图像数据过多、图像处理不佳等。CADe 系统为这些问题带来了转机, 计算机快速显示大量图像数据, 检测出可疑病变, 然后将发现的病变发送给医生阅读。它显著降低了上述诸多问题的影响, 可以看作是人工智能系统在乳腺癌筛查领域的雏形。随着 CADe 系统的快速发展, 人们也在结合深度学习技术, 拓展自动化乳腺癌筛查在 3D 超声、乳腺 MRI 和乳腺断层合成图像等平台上的应用<sup>[8-11]</sup>。

肿瘤筛查是能够利用人工智能辅助完成的窄任务中比较重要的一项。癌症筛查指在一个特殊的时间点输入和评估数据, 以确定是否需要额外的诊断测试和诊断过程。数据流的形式可以是血清标志物、医学影像、视觉检查、内窥镜检查。这些模式中的每一种都为整合人工智能以改善癌症的预测提供了机会。早期研究表明, 对于前列腺特异性抗原 (PSA) 等标志物, 机器学习算法在不同时间点结合其他血清标志物对 PSA 进行建模, 可能比仅用 PSA 做指标能更好地预测前列腺癌的存在<sup>[12]</sup>。与其他任何应用相比, 人工智能在医学影像筛查中应用的作用也许更大。目前已有应用 CT 在肺癌和应用钼靶 X 射线在乳腺癌中定位病变和预测恶性肿瘤风险的模型, 并且被证明与肿瘤专家的表现不相上下甚至更优<sup>[13]</sup>。在这些应用中, 原始图像数据被输入深度学习卷积神经网络, 该网络以放射科医生标记的真值作为训练集。重要的是, 虽然这些算法在曲线下面积、敏感度和特异度方面的结果令人印象深刻, 但它们并不直接评估癌症死亡率、医疗费用或生活质量等临床终点。除去医学影像, 人工智能在结肠直肠癌内镜筛查中可用于指导组织活检位置的选取<sup>[14-15]</sup>。此外, 人工智能也潜在地可用于改善其他癌症的诊断水平, 因为通常而言, 这些癌症的筛查较为困难。人工智能可以通过重新分析现有数据流 (如 MRI 或 CT 检查结果) 或整合多模态数据流 (如基因组数据和 EHR) 来实现筛查效果的增强。尽管现在美国预防服务工作组 (USPSTF) 对于许多癌种未提出筛查建议<sup>[16]</sup>, 但将人工智能纳入筛查标准有可能使更多癌种的早期筛查变得可行。

2018 年, 中国 60% 的癌症都是在晚期 (III 或 IV 期) 被诊断出来的, 因此当时的癌症死亡率相对较高 (183.89/100 000, 对比美国 146.2/100 000)。有效的早期检测能够显著改善预后, 降低治疗费用。2017 年, 我国的体检率为 35.8%, 低于美国 (74.2%)、日本 (73.4%) 等发达国家。衡量筛查手段的主要标准为灵敏度和阳性预测值 (PPV), 现在临床上的早筛标准方案灵敏度尚可, 但特异度和阳性预测值较低, 因此许多人需进行额外的检验, 这增加了大众的检测成本。这一标准方案从进行检查到医生判定筛查结果共需约一周时间。

液体活检技术是新一代的早筛技术, 主要方法为基于 NGS 的 ctDNA 甲基化组测序。使用 AI 聚类算法分析测序结果, 可以有效地区分癌性病灶和非癌性病灶。相比于现标准方案, 液体活检能够在保证较高灵敏度的基础上提升 PPV, 并且更加方便 (仅需要血样): 我国研发的一款多癌种 AI 算法已经能够达到 80.6% 的灵敏度和 98.3% 的特异度, 与传统早筛方法相比有巨大的优势。另外, AI 辅助的液体活检早筛可同时检测多个癌种, 也具有价格上的优势 (单癌种筛查价格仅为传统方法的约 1/5)。

MRD 检查具有评估复发风险、指导辅助治疗、疗效预测、判断患者体内肿瘤负荷等多种用途, 与液体活检早筛相似, 其同样基于 ctDNA 测序来开展。MRD 与传统肿瘤评估方法的最大区别在于超前预测。虽然癌胚抗原 (CEA) 等肿瘤标志物或 PET-CT 等影像学检查提高了检测的准确性和范围, 但它们只能反映存在肿瘤或肿瘤异常时的情况, 而 MRD 却可以在肿瘤形成之前做出预测。研究表明, MRD 检查比 CT 检测到病灶至少早 72%。因此, MRD 在癌症预防和早期筛查中具有不可替代的作用。

## 1.2 病灶分析: 人工智能提取病灶特征, 鉴别肿瘤良恶性

肺癌有多种生物学行为, 有些进展缓慢, 有些则更具侵袭性, 需要积极治疗。一些预后因素与更好的生存率相关 (如女性、早期、EGFR 突变、无局部淋巴结受累)<sup>[17]</sup>, 而另一些与较低的生存率相关 (如肺功能低下、心血管疾病、男性、吸烟、高龄和晚期癌症)<sup>[18-21]</sup>。然而, 上述预后因素难以描述临床应用中的肿瘤的异质性和动态特征。肿瘤病灶处于不断进化分化的过程中, 改变自身的基因组和表型, 甚至通过转移改变其空间位置。化疗药物可能因其致突变性加速病灶的变化过程, 迫使肿瘤



学家需要不断调整方案以适应其变化。

与此相反,基于影像的生物标志物能够在多个时间点无创地提取肿瘤的影像特征,从而确定肿瘤的病理生理学特性。肿瘤大小的测量值(例如最长直径)经常被用于肿瘤分期及疗效评估,这往往是由于此类测量值较易获取。然而,基于肿瘤大小特征以及相应分期的方法具有局限性:其受临床结局及疗效反应测量结果的可变性影响较大。目前已有许多研究成功将语义性特征及可自动化获取的影像组学特征应用于预测肺癌患者预后。例如,CANARY工具包(计算机辅助肺结节评估及风险计算工具)通过语义学影像特征计算风险,从而识别具有较高侵袭性的肺腺癌病例。前期研究已显示AI能够自动化定量提取肿瘤的影像学特征,而此类特征能够有效预测多种肿瘤的预后,例如肺癌( $P < 3.53 \times 10^{-6}$ );另外,此类特征与肺腺癌的远隔转移( $P < 1.79 \times 10^{-17}$ )、肿瘤病理亚型( $P = 2.3 \times 10^{-7}$ )以及体细胞突变和基因表达特征等部分生物学特性显著相关<sup>[22]</sup>。

在乳腺癌诊断中,也有类似的特征识别技术应用。通过人工智能方法对胸围数据进行分析,可以定量获得肿瘤的大小、形状、形态、纹理和动态图像特征。例如,使用来自乳腺MRI造影剂成像的动态图像,研究人员可以量化肿瘤异质性并推断空间和动态图像特征的表型。作为另一个例子,熵是对随机性的数学描述,它可以提供关于肿瘤图像异质性的信息,进而可以描述DCE-MRI中肿瘤内血管记录(即对比记录)的异质性。这种成像特性可能反映了肿瘤中血管生成的空间异质性和治疗敏感性,如NCI的癌症基因组图谱(TCGA)乳腺癌表型组所示<sup>[23]</sup>。

人们在计算机辅助诊断中比较了传统的特征工程算法和深度学习算法。研究表明,传统的特征工程算法和深度学习算法在区分乳腺良恶性病变方面都表现良好,这两种方法的“融合”算法可以显著提高性能<sup>[24-25]</sup>。在3种乳腺成像方式(690例DCE-MRI病例、245例FFDM病例和1125例超声病例)中,“融合”算法表现最好,这表明传统特征工程和深度学习算法均可用于乳腺癌诊断<sup>[22]</sup>。另一项研究使用2282个数字化电影图像和FFDM数据来呈现迁移学习CNN,然后在324个乳房断层合成立体图像上提取肿瘤特征<sup>[26]</sup>,展示了在不同图像成像方法之间传输图像信息的能力。

乳腺密度增加是乳腺癌的危险因素,它的遮蔽

效应也掩盖了病变,因此研究人员使用基于面积或基于体积的密度测定来估计乳房X线的密度。乳腺密度反映了乳腺纤维腺体组织相对于脂肪组织的含量,这些组织类型在全视野数字乳房X线摄影(FFDM)上更容易区分,因为X射线在纤维腺体组织中比在脂肪组织中衰减得多。FFDM是乳腺的2D投影图像,因此可用于估计3D百分比密度值。除此之外,致密组织的空间分布特征等证据能够表明乳腺实质图像的变化,也与乳腺癌风险相关。通过使用偏度来描述乳房X光照片上的密度分布,然后应用基于图像的纹理分析,研究人员能够从FFDM区域内灰色区域的空间分布中提取特征。其他研究人员使用纹理分析和深度学习来区分BRCA1/BRCA2突变携带者(或患有对侧乳腺癌的患者)与患乳腺癌风险低的人。一项对近500例病例的研究表明,乳腺癌高风险女性的乳房较致密,实质较粗,对比度较低<sup>[27]</sup>。进一步的研究将纹理分析应用于乳房断层合成图像以描述实质模式,最终用于乳腺癌风险评估。初步结果表明,与数字乳房X线照相术相比,乳房断层合成图像中的纹理特征与乳房密度的相关性更强<sup>[28]</sup>。

此外,在动态对比增强(DCE)MRI上,乳腺实质模式的特征提取方法也扩展到了乳腺实质增强(BPE)的图像特征。在50个BRCA1/BRCA2携带者的数据集中,BPE的定量特征与乳腺癌的存在相关,并且其百分比的相对变化预测了输卵管卵巢切除术后患者的乳腺癌进展<sup>[29]</sup>。深度学习方法越来越多地用于评估乳房密度和实质特征,如在一项研究中,研究人员应用迁移学习根据BRCA1/BRCA2状态区分乳腺癌高风险女性<sup>[27]</sup>。

### 1.3 突变判断:人工智能辅助计算肿瘤负荷,判断肿瘤突变状态

基于RECIST标准的定量成像评分的成功应用为人工智能在肿瘤成像中的发展铺平了道路,因为这种成像评分被广泛用于大量临床研究,作为生存的早期替代指标产生大量带有临床元数据的CT图像数据集。此类数据集的重要性在于它们可以满足训练和验证AI算法所需的海量数据需求,而从头开始构建相同大小的数据集在时间和成本方面是不切实际的。由于RECIST的广泛应用,定量CT分析已成为当前肿瘤学的重要工具<sup>[30]</sup>,为AI预测模型提供了转化研究的基石。

AI的优势非常适合应对当前靶向和免疫疗法的挑战,从而带来RECIST难以实现的明显临床益

处。RECIST 的图像端点设计依赖于有效治疗反映在肿瘤体积缩小上的假设, 而一些图像端点是基于肿瘤直径的假设, 即肿瘤呈球形并且在治疗后均匀收缩的假设。靶向治疗和免疫治疗导致的治疗反应模式不同于放疗和化疗等传统方法, 导致当前基于 RECIST 的终点表现不佳, 这可能导致更高的临床试验失败率和更高的药物开发成本。人工智能使肿瘤尺寸之外的指标能够被量化, 并指导实际临床应用。

目前, 靶向治疗和免疫治疗的疗效预测方法大多基于 PD-L1、EGFR 等肿瘤免疫微环境的生物标志物和突变状态。此类信息需要侵入性活检, 难以在多次随访中重复, 并且受限于肿瘤采样位置。同时, 现有的 PD-L1 表达水平的预测能力也相对有限。例如, 在 KEYNOTE-189 试验中, 无论 PD-L1 表达水平如何, 派姆单抗免疫疗法联合标准化疗对所有患者都有生存获益。即使在 PD-L1 肿瘤比例评分低于 1% 的患者中也实现了生存获益——上述评分本意味着对应患者群体生存获益的可能性很低<sup>[31]</sup>。

越来越多的证据表明, 人工智能可以通过识别与治疗反应相关的放射标记来评估对免疫治疗的反应。该研究证实, 以 PD-1 抗体纳武利尤单抗作为二线治疗的 NSCLC 患者的总生存期与影像学特性显著相关。在这项研究中, 治疗前测量了 2 个显著预测总生存期 (OS) 的放射组学特征, 包括区域差异和熵。这些特征表明, 增强 CT 上原发肿瘤的程度分布模式是高度异质的。另一项关于肺癌的研究表明, 通过将基因组学和放射组学信息整合到临床模型中, OS 的预测能力显著提高, 表现为 95% 置信区间宽度从 0.65 提高到 0.73。这些发现表明, 放射学特征和基因组生物标志物是互补的, 为人工智能解释两者之间可预测的关系提供了潜力。尽管可以通过使用机器学习方法分析微阵列的基因表达模式来对肺癌进行基因分型, 但使用机器学习技术预测组学 - 基因组学数据与临床结果相关性的研究尚无明确结果<sup>[32]</sup>。

使用人工智能对图像数据进行定量分析, 也可以提高预测靶向治疗效果的能力。用贝伐单抗治疗 NSCLC 导致放射性荧光标记的脱氧葡萄糖摄取减少, 这在检测反应良好的患者中比传统 CT 标准更有效 (73% 对 18%); 在该研究中, PET 或 CT 结果与 OS 没有显著相关性。目前, 人们主要是通过分析靶基因的突变来实现靶向治疗效果的预测。人工智能模型可以通过识别与突变状态相关的图像特征

来有效地帮助预测。这种方法具有非侵入性的优点, 可以反复评估整体肿瘤突变状态, 而不仅仅是活检部位, 因此可以避免由于肿瘤异质性和某些转移灶获得耐药性而降低预后预测能力。这种方法得到了对接受吉非替尼治疗的 NSCLC 患者的定量成像研究的支持: 研究表明, 放射性特征 Laws-Energy 可以有效地预测 EGFR 突变状态。

当使用生物标志物作为性能评估标准时, 它们的测量过程必须是客观的和可重复的。由于人工智能不同于人类医生, 前者没有与视觉评估相关的观察者间变异, 故其对肿瘤成像特征的复杂模式的识别和测量是高度客观的。因此, 了解放射学特征的测量误差对于提高人工智能预测模型的可重复性非常重要。不同的肿瘤分割算法会引入偏差并影响图像属性的计算, 进而影响人工智能算法的性能。因此, 需要半自动分割技术来避免这种负面影响<sup>[33]</sup>。图像环境 (CT 扫描仪、切割厚度和重建方法等) 的差异也会影响放射组学特征的计算<sup>[34-35]</sup>。这些环境变化客观存在于临床实践和临床试验中, 可能会影响人工智能提取的图像特征的预测能力和再现性。CNN 算法构建中的训练和验证过程可以削弱这种影响, 主要是通过选择可重复性强的预测特征, 丢弃不同图像集数据之间差异较大的预测特征, 但仍需要进一步的研究。尽管人工智能方法的发展极为迅速, 但在数据标准更新 (如从 RECIST 1.0 更新至 1.1) 之前, 传统图像端点在各种临床试验中提供的数据格式的一致性为大数据算法提供了保证。

#### 1.4 疗效评估: 人工智能综合影像学 with 临床评估, 有潜力改善边缘病例治疗评价

肿瘤疗效的评估通常由影像学评估和临床评估两部分组成。RECIST 和 RANO 等定量反应评估标准长期以来一直被确立为评估治疗反应的可重复方法 (尽管其有效性在靶向 / 免疫治疗肿瘤的时代受到质疑)<sup>[36]</sup>。然而, 随着靶向和免疫疗法进入临床实践, 由于肿瘤假性进展等现象的存在, 仅通过 RECIST 进行疗效评估还不够<sup>[37]</sup>。详细的疗效评估通常需要大量的人类专业知识、经验和时间, 并且不同医生之间阅片的差异对于疗效评估是额外的挑战。此外, 尽管这些标准经常被审查和修订, 但在评价边缘病例方面不甚有效, 如在接受免疫治疗的患者中, 病变反应是动态的。目前深度学习已显示出能够自动进行疗效评估的潜力, 具体而言包括正接受免疫治疗的患者的 RECIST 评估<sup>[38]</sup> 和自动 RANO 评估<sup>[39]</sup>。



### 1.5 患者随访：人工智能有望提升随访效率，辅助风险预测，指导后续治疗

人工智能在肿瘤学中应用的另一个潜在领域是开发随访辅助工具。肿瘤筛查和诊断使用的算法通常可以迁移到随访中，但需要对具体任务重新进行训练与验证。既往癌症治疗对数据流的影响往往会使问题发生重大改变，如从同一癌性病灶中提取的放射组学特征在治疗前后有显著差异<sup>[40]</sup>。这些发生改变的特征可用于预测患者复发风险，有助于调整后续治疗方案<sup>[41]</sup>。促进对患者的深度随访可以降低发病率，提升医疗资源利用效率；使用 EHR 数据的人工智能已被证明能够在治疗时预测有重症风险的患者，并建议使用更强的预防措施<sup>[42]</sup>。当患者的肿瘤复发被判断为无法治愈时，临终姑息治疗的重要性得到凸显。人工智能能够提前对高死亡风险患者进行分类，并鼓励医生与患者和家属交流其对治疗和生活质量的预期，指导姑息治疗方案的制定<sup>[43]</sup>。

## 2 人工智能在肿瘤领域临床应用中面临的挑战

尽管人工智能在肿瘤领域已经有了诸多研究进展，由于医疗行业及肿瘤疾病的特殊性，其在临床应用中仍面临一些挑战。这些挑战主要分为两类：有待提高的医生与患者接受度和可能的医学伦理学问题。

从医生的角度而言，目前仍没有评价人工智能用于肿瘤诊疗的量化指标，因此，临床医生对人工智能医疗产品的接受度并不高。比如在呼吸系统疾病领域，人工智能的所有临床应用还处在偏早期的试验探索的阶段。在肺癌领域，尽管胸部 CT 影像分析系统已经取得医疗器械许可证，但临床医生在使用胸部 CT 人工智能系统时还是十分保守和谨慎，主要有以下几种可能的原因。(1) 人工智能的结果可解释性相对低。由于医生不知道人工智能系统如何学习和训练以获得计算模型，他们也可能不知道系统如何推理得出结论。因此，系统输入数据与其输出结论之间缺乏“因果关系”，临床医生在接受系统的建议时持怀疑态度。(2) 人工智能系统的鲁棒性较差。面对复杂的临床诊断和鉴别诊断，系统没有足够的缓冲弹性，病历信息的细微差异可能造成超出其控制范围的决策偏差。然而当人工智能系统出错时，本身无法知道如何检测这些错误，并且这种错误很难检测到，故常会产生严重的事故。(3) 人工智能模型的建立和机器学习训练需要大量格式高度标准、高质量的数据，而获得这些数据极其困

难。繁重的临床工作将编写病历的需求降至最低，医疗数据精简化的采集原则也不利于临床实践中大数据的建立。人们无法得知建立在高标准、高质量数据之上的人工智能能否应对临床实践中数据采集简化、低纬度的情况。

在患者端，接受度低的问题同样有可能存在。由于多数人对肿瘤疾病的认知不足，经常会在疾病早期高估其严重性，因此他们常常更倾向于相信人类肿瘤科医生的决策。只有当患者准备好接受基于人工智能的肿瘤治疗时，人工智能的实施才能取得成功。但由于人工智能医疗系统本身的特性有限，可解释性不足，以及相关责任和伦理安全问题尚未解决，目前患者对人工智能医疗产品的接受程度有限。

另外，人工智能在肿瘤领域的应用面临着一些伦理安全问题。目前各类人工智能系统在临床应用中的定位仍然是辅助工具，这是因为人工智能系统在作用上的进一步扩展受到一些伦理问题的限制。比如，患者挂号就诊的目的是向有执照的肿瘤医生寻求帮助，如果患者被推荐到“人工智能肿瘤医生”处就诊，医患伦理就得不到一个合理的解释。如果肿瘤医生根据人工智能算法给出的建议为患者给出治疗方案或进行手术，但是后续过程中发生了医疗事故，在事故责任划分、医患矛盾处理上都存在较难以解决的伦理问题。

## 3 总结

肿瘤曾被认为是“无法治愈的疾病”，然而近年来的大量研究进展已经让人们看到了战胜肿瘤的曙光。随着计算机科学技术的快速发展，人工智能也逐渐体现出了其在肿瘤领域的应用潜力。人工智能能够潜在地推进肿瘤智慧医疗、分级诊疗、规范化、科学化和个体化等概念，缩小医疗水平差距，将肿瘤防治关口前移，避免治疗不足或治疗过度。同时，我们也要正视目前人工智能技术在肿瘤方面的诸多不足，意识到这一技术仍需要更多研发力量和法律法规的支持。人工智能在肿瘤诊疗中的应用推动了人工智能新技术与经济社会发展需求的深度融合，将人工智能技术创新应用于肿瘤多学科诊疗领域，有助于实现人工智能助力国家医疗卫生水平及经济社会水平发展的目标。

### [参 考 文 献]

- [1] Chambers DA, Vinson CA, Norton WE. Advancing the

- science of implementation across the cancer continuum[M]. New York: Oxford University Press, 2018
- [2] National Lung Screening Trial Research Team, Aberle DR, Adams AM, et al. Reduced lung-cancer mortality with low-dose computed tomographic screening. *N Engl J Med*, 2011, 365: 395-409
- [3] Rami-Porta R, Bolejack V, Giroux DJ, et al. The IASLC lung cancer staging project: the new database to inform the eighth edition of the TNM classification of lung cancer. *J Thorac Oncol*, 2014, 9: 1618-24
- [4] Lodwick GS. Computer-aided diagnosis in radiology. A research plan. *Invest Radiol*, 1966, 1: 72-80
- [5] Giger ML. Computerized analysis of images in the detection and diagnosis of breast cancer. *Semin Ultrasound CT MR*, 2004, 25: 411-8
- [6] Freer TW, Ulissey MJ. Screening mammography with computer-aided detection: prospective study of 12,860 patients in a community breast center. *Radiology*, 2001, 220: 781-6
- [7] Rao VM, Levin DC, Parker L, et al. How widely is computer-aided detection used in screening and diagnostic mammography? *J Am Coll Radiol*, 2010, 7: 802-5
- [8] Kooi T, Litjens G, van Ginneken B, et al. Large scale deep learning for computer aided detection of mammographic lesions. *Med Image Anal*, 2017, 35: 303-12
- [9] Ribli D, Horvath A, Unger Z, et al. Detecting and classifying lesions in mammograms with deep learning. *Sci Rep*, 2018, 8: 4165
- [10] Yap MH, Pons G, Marti J, et al. Automated breast ultrasound lesions detection using convolutional neural networks. *IEEE J Biomed Health Inform*, 2018, 22: 1218-26
- [11] Dalmis MU, Vreemann S, Kooi T, et al. Fully automated detection of breast cancer in screening MRI using convolutional neural networks. *J Med Imaging (Bellingham)*, 2018, 5: 014502
- [12] Nitta S, Tsutsumi M, Sakka S, et al. Machine learning methods can more efficiently predict prostate cancer compared with prostate-specific antigen density and prostate-specific antigen velocity. *Prostate Int*, 2019, 7: 114-8
- [13] Salim M, Wählén E, Dembrower K, et al. External evaluation of 3 commercial artificial intelligence algorithms for independent assessment of screening mammograms. *JAMA Oncol*, 2020, 6: 1581-8
- [14] Guo L, Xiao X, Wu C, et al. Real-time automated diagnosis of precancerous lesions and early esophageal squamous cell carcinoma using a deep learning model (with videos). *Gastrointest Endosc*, 2020, 91: 41-51
- [15] Zhou D, Tian F, Tian X, et al. Diagnostic evaluation of a deep learning model for optical diagnosis of colorectal cancer. *Nat Commun*, 2020, 11: 2961
- [16] USPSTF: A and B Recommendations. <https://www.uspreventiveservicestaskforce.org/uspstf/recommendation-topics/uspstf-and-b-recommendations.2021>
- [17] Schabath MB, Haura EB. Epidemiology of non-small cell lung neoplasms [M]//Cameron RB, Gage DL, Olevsky O. *Modern thoracic oncology. Volume 2: trachea, lung, and pleura*. Singapore: World Scientific Publishing, 2018: 61-6
- [18] Dela Cruz CS, Tanoue LT, Matthay RA. Lung cancer: epidemiology, etiology, and prevention. *Clin Chest Med*, 2011, 32: 605-44
- [19] Schabath MB, Thompson ZJ, Gray JE. Temporal trends in demographics and overall survival of non-small-cell lung cancer patients at Moffitt Cancer Center from 1986 to 2008. *Cancer Control*, 2014, 21: 51-6
- [20] Kachroo S, Tong L, Spitz MR, et al. Trends in prevalence of prognostic factors and survival in lung cancer patients from 1985 to 2004 at a tertiary care center. *Cancer Detect Prev*, 2008, 32: 101-8
- [21] Birim O, Kappetein AP, van Klaveren RJ, et al. Prognostic factors in non-small cell lung cancer surgery. *Eur J Surg Oncol*, 2006, 32: 12-23
- [22] Bi WL, Hosny A, Schabath MB, et al. Artificial intelligence in cancer imaging: Clinical challenges and applications. *CA Cancer J Clin*, 2019, 69: 127-57
- [23] Zhu Y, Li H, Guo W, et al. Deciphering genomic underpinnings of quantitative MRI-based radiomic phenotypes of invasive breast carcinoma. *Sci Rep*, 2015, 5: 17787
- [24] Huynh BQ, Li H, Giger ML. Digital mammographic tumor classification using transfer learning from deep convolutional neural networks. *J Med Imaging*, 2016, 3: 034501
- [25] Andropova N, Huynh BQ, Giger ML. A deep feature fusion methodology for breast cancer diagnosis demonstrated on three imaging modality datasets. *Med Phys*, 2017, 44: 5162-71
- [26] Samala RK, Chan HP, Hadjiiski L, et al. Mass detection in digital breast tomosynthesis: deep convolutional neural network with transfer learning from mammography. *Med Phys*, 2016, 43: 6654
- [27] Li H, Giger ML, Huynh BQ, et al. Deep learning in breast cancer risk assessment: evaluation of convolutional neural networks on a clinical dataset of full-field digital mammograms. *J Med Imaging (Bellingham)*, 2017, 4: 041304
- [28] Kontos D, Bakic PR, Carton AK, et al. Parenchymal texture analysis in digital breast tomosynthesis for breast cancer risk estimation: a preliminary study. *Acad Radiol*, 2009, 16: 283-98
- [29] Wu S, Weinstein SP, DeLeo MJ 3rd, et al. Quantitative assessment of background parenchymal enhancement in breast MRI predicts response to risk-reducing salpingo-oophorectomy: preliminary evaluation in a cohort of BRCA1/2 mutation carriers. *Breast Cancer Res*, 2015, 17: 67
- [30] Schwartz LH, Seymour L, Litiere S, et al. RECIST 1.1—standardisation and disease-specific adaptations: perspectives from the RECIST Working Group. *Eur J Cancer*, 2016, 62: 138-45
- [31] Gandhi L, Rodriguez-Abreu D, Gadgeel S, et al. Pembrolizumab plus chemotherapy in metastatic nonsmall-cell lung cancer. *N Engl J Med*, 2018, 378:

- 2078-92
- [32] Grossmann P, Stringfield O, El-Hachem N, et al. Defining the biological basis of radiomic phenotypes in lung cancer. *Elife*, 2017, 6: 323421
- [33] Balagurunathan Y, Gu Y, Wang H, et al. Reproducibility and prognosis of quantitative features extracted from CT images. *Transl Oncol*, 2014, 7: 72-87
- [34] Balagurunathan Y, Kumar V, Gu Y, et al. Test-retest reproducibility analysis of lung CT image features. *J Digit Imaging*, 2014, 27: 805-23
- [35] Zhao B, Tan Y, Tsai WY, et al. Reproducibility of radiomics for deciphering tumor phenotype with imaging. *Sci Rep*, 2016, 6: 23428
- [36] Villaruz LC, Socinski MA. The clinical viewpoint: definitions, limitations of RECIST, practical considerations of measurement. *Clin Cancer Res*, 2013, 19: 2629-36
- [37] Gerwing M, Herrmann K, Helfen A, et al. The beginning of the end for conventional RECIST — novel therapies require novel imaging approaches. *Nat Rev Clin Oncol*, 2019, 16: 442-58
- [38] Arbour KC, Luu AT, Luo J, et al. Deep learning to estimate RECIST in patients with NSCLC treated with PD-1 blockade. *Cancer Discov*, 2020, 11: 59-67
- [39] Kickingereder P, Isensee F, Tursunova I, et al. Automated quantitative tumour response assessment of MRI in neurooncology with artificial neural networks: a multicentre, retrospective study. *Lancet Oncol*, 2019, 20: 728-40
- [40] van Dijk LV, Langendijk JA, Zhai TT, et al. Delta-radiomics features during radiotherapy improve the prediction of late xerostomia. *Sci Rep*, 2019, 9: 12483
- [41] Chang Y, Lafata K, Sun W, et al. An investigation of machine learning methods in delta-radiomics feature analysis. *PLoS One*, 2018, 14: e0226348
- [42] Hong JC, Eclov NCW, Dalal NH, et al. System for high-intensity evaluation during radiation therapy (SHIELD-RT): a prospective randomized study of machine learning-directed clinical evaluations during radiation and chemoradiation. *J Clin Oncol*, 2020, 38: 3652-61
- [43] Ramchandran KJ, Shega JW, Von Roenn J, et al. A predictive model to identify hospitalized cancer patients at risk for 30-day mortality based on admission criteria via the electronic medical record. *Cancer*, 2013, 119: 2074-80