

药物非临床安全性评价毒性病理学深度学习方法和应用概述

张 頔¹, 闫振龙², 杨艳伟¹, 屈 哲¹, 霍桂桃¹, 李双星¹, 林 志¹, 吕建军³

(1 中国食品药品检定研究院, 国家药物安全评价监测中心, 北京市重点实验室, 北京 100176;

2 益诺思生物技术南通有限公司, 南通 226133; 3 湖北天勤鑫圣生物科技有限公司, 武汉 430207)

[摘要] 药物非临床安全性评价毒理学试验组织病理学数据在药物发现和开发过程中非常关键, 是监管部门在实验动物中评估药物相关毒性及其对临床试验人类安全性影响所必需的。随着全切片图像、数字病理学和算法研究的进步, 人工智能已广泛应用于人类临床医学实践。但机器学习, 尤其是深度学习在非临床毒理学试验毒性病理学的进展比较缓慢。本文简要综述了毒性病理学在药物发现和开发中的作用、人工智能和深度学习方法概述、深度学习方法在毒性病理学中的应用以及在毒性病理学中应用深度学习方法所面临的挑战, 以期为我国药物非临床安全性评价毒理学试验组织病理学检查应用深度学习方法提供一定参考。

[关键词] 药物; 非临床安全性评价; 毒性病理学; 人工智能; 深度学习

[中图分类号] R99 **[文献标志码]** A **[文章编号]** 1003-3734(2023)21-2147-08

Overview of approaches and applications of deep learning in toxicologic pathology during nonclinical safety evaluation of drugs

ZHANG Di¹, YAN Zhen-long², YANG Yan-wei¹, QU Zhe¹, HUO Gui-tao¹, LI Shuang-xing¹,
LIN Zhi¹, LÜ Jian-jun³

(1 National Center for Safety Evaluation of Drugs, National Institutes for Food and Drug Control, Beijing Key Laboratory, Beijing 100176, China; 2 Innostar Biotechnology Nantong Co., Ltd., Nantong 226133, China; 3 Hubei Topgene Xinsheng Biotechnology Co., Ltd., Wuhan 430207, China)

[Abstract] Histopathology data of toxicology studies during nonclinical safety evaluation of drugs are critical during the drug discovery and development process that is necessary for regulatory bodies to assess drug-related toxicity in laboratory animals and its impact on human safety in clinical trials. Artificial intelligence (AI) has been widely used in clinical medical practice with progress in the whole slide image (WSI), digital pathology and algorithms. However, the progress of machine learning (ML)-in particular, deep learning (DL)-has been rather slow in toxicologic pathology of nonclinical toxicology studies. This paper briefly reviews the role of toxicologic pathology in drug discovery and development, overview of AI and approaches of DL, applications of approaches of DL in toxicologic pathology, as well as the challenges in implementation of approaches of DL in toxicologic pathology, in order to provide reference for implementation of approaches of DL in toxicologic pathology for toxicology studies during

[基金项目] 中国食品药品检定研究院学科带头人课题(2021X2)

[作者简介] 张頔, 女, 本科, 主要从事药物非临床安全性评价工作。E-mail: 1374390958@qq.com。共同第一作者: 闫振龙, 男, 硕士, 主要从事药物非临床安全性评价毒性病理学诊断工作。E-mail: zlyan-1@innostar.cn。

[通讯作者] 林志, 女, 博士, 研究员; 主要从事药物非临床安全性评价毒性病理学诊断工作。E-mail: linzhi@nifdc.org.cn。吕建军, 男, 博士, 主任药师, 主要从事药物非临床安全性评价毒性病理学诊断工作。E-mail: jianjunlv@hotmail.com。

nonclinical safety evaluation of drugs in China.

[Key words] drug; nonclinical safety evaluation; toxicologic pathology; artificial intelligence; deep learning

在药物非临床安全性评价中,为了检测受试物所致的各种毒性反应,通常会制备大量的组织切片进行病理学评价。这些切片的阅片工作量巨大、所需时间较长,通常一项毒理学试验可能产生数千张组织切片^[1]。一般而言,常规毒理学试验中大多数组织处于正常范围内,而异常组织可能包含多种不同的形态学改变。毒性病理学家需要多年的培训和实践,才能将不同种属、品系、年龄、性别的常用实验动物特有的背景病变与受试物相关病变进行准确鉴别。因此,使用人工智能(artificial intelligence, AI)系统进行组织病理学评价可在一定程度上缩短非临床毒理学试验的时间,加速药物开发过程^[2]。迄今为止,基于深度学习(deep learning, DL)的方法多被开发用于人类活检样本中肿瘤诊断和生物标志物特征的描述。2013—2019年发表使用不同DL方法的130多篇文章中,几乎都与人类癌症相关^[3],这些应用方法在临床肿瘤学实践或发现生物学中意义重大^[4]。然而,目前在毒性病理学领域仍缺少此类方法。虽然对数字病理学(digital pathology, DP)的不同方面有详细和全面的综述、评论、意见和调查,但其中大部分都是关于DL的临床诊断应用^[3,5-12]。本文对药物非临床安全性评价毒性病理学DL方法和应用进行综述,主要包括毒性病理学在药物发现和开发中的作用、AI和DL方法概述、DL方法在毒性病理学中的应用以及在毒性病理学中应用DL方法所面临的挑战,以期为在我国药物非临床安全性评价毒理学试验的组织病理学检查中应用DL方法提供一定参考。

1 毒性病理学在药物发现和开发中的作用

药物发现和开发是一个复杂且长期的过程,从一个化学实体或化合物的靶标识别和验证、安全性评价到最后临床应用,整个过程平均需要10~12年^[13]。其目的是通过在非临床和临床安全性及有效性严格测试后,将安全有效的药物上市销售。毒性病理学家在一个高度矩阵化、多学科的环境中工作,从早期发现到监管批准,甚至上市后的研究,支持着药物开发的整个过程^[13]。

毒性病理学是一门综合了病理学和毒理学2门学科的交叉学科,主要研究受试物(如药物、医疗器械或化学品)对实验动物组织的影响。毒性病理学

家对多个种属实验动物的毒理学试验结果进行形态学评估,准确诊断并报告组织病理学结果。试验的周期越长,毒性病理学检查的结果越为重要。毒性病理学回答受试物造成病理损伤部位、程度、性质和可恢复性等基本问题,因此毒性病理学是药物非临床毒理学试验最重要的方法之一,也是非临床药物安全性评价工作的最基本内容。组织病理学是在光学显微镜水平上观察组织的形态学变化,其结果常常构成毒理学试验中未观察到有害作用剂量(no observed adverse effect level, NOAEL)的基础,也是药物、化学品、生物制品和医疗器械危害识别和风险评估的关键组成部分。

2 AI和DL方法概述

对于病理学家而言,组织病理学评价的基本组成部分是细胞。病理学家基于多年的严格训练和经验,对细胞类型和组织结构进行显微镜下观察,从而得出准确的诊断。病理学家将AI系统拟人化,并假设其与人类“看到”细胞和组织的方式相同,从而做出初步诊断,甚至期望AI对其他研究中看到的结果也有一定程度的“记忆”。但AI系统的运行方式存在一些关键差异,计算机以二进制(0和1)方式运行,“看到”的图像是以像素为单位的,单个像素并不能提供有意义的信息。但随着数学算法和计算能力的进步,计算机可以像人类一样接受训练,以更高的抽象水平查看图像中的特征,例如:AI系统可以使用形状和纹理的像素级别尺寸“学习”细胞核等特征,还可以从其训练示例中存在的任何特征上学习^[5]。尽管近10年在先进的AI算法开发方面取得了显著进步,并已应用于许多行业。但其在药物研发工作中仍处于早期阶段^[14-15],尚未在毒性病理学学科中得到广泛应用。目前AI系统在毒性病理学中的应用主要包括以下几方面。

2.1 切片数字化

将AI应用于毒性病理学,第1步就是切片数字化,全切片图像(whole slide image, WSI)的进步和DP系统的显著改进为病理工作流程完全数字化奠定了基础。在人类病理学和动物病理学中已被证明,基于WSI的数字化病理工作流程的诊断与传统的基于玻璃切片的诊断相比具有以下优势:① WSI存档简单和高效。② 可追溯性,可简单和快速检索

案例。③ 便捷性,可及时在全球范围内进行同行评议和疑难切片会诊。④ 可对比性,可在同一屏幕上观察多个图像进行比较,可以避免诊断漂移或解决动物之间的细微差异。值得注意的是,由于2020年全球新型冠状病毒肺炎大流行,WSI用于远程同行评议被全球毒性病理学家广泛接受^[16]。

2.2 DL 方法

在介绍 DL 一些关键方法之前,应首先理解架构(architecture)、模型(model)和算法(algorithm)以及三者间的差异。架构是指 DL 任务的一般方法和该方法的参数化,如不同层的数量和大小或每个层的类型。模型是某一给定架构的一个具体实例,或者在某一给定数据集上训练的多个架构的模块化实例。算法是实现架构时需要遵循的一组规则^[17]。在病理学中,基于 DL 的方法通常有 4 个主要步骤:

- ① 建立数据库。
- ② 预处理。
- ③ 模型选择和训练。
- ④ 后处理和评价。

2.2.1 建立数据库

DL 研究首先要在计算机系统中创建一个以电子方式存储的各种图像数据库。虽然人类医学中有多个公开的病理学数据库,但非临床研究中只有少数数据库(见表 1)。值得注意的是,这些非临床数据库只是 WSI 存储库,没有经过病理学家审查的参考标准注释,也不用于 DP。虽然收集没有基本注释的 WSI 存储库相对容易,但对于大多数 AI/DL 应用程序,这些数据库需要有病理学家的注释。创新药物倡议(BIGPICTURE 项目)等旨在建立人类和非临床不同物种数字切片的中央存储库,从而尝试创建 AI 工具,这些尝试会在不久的将来加速其在毒性病理学领域的发展^[18]。

表 1 非临床数字病理学 WSI 存储库和临床 HE 染色切片存储库

存储库	临床/非临床	网址
TG-GATE	非临床	http://toxico.nibio.go.jp/english/index.html
CEBS	非临床	https://connect.niehs.nih.gov/cebs3/ui/
VMD	临床和非临床	http://www.virtualmicroscopydatabase.org/
TCGA	临床	https://portal.gdc.cancer.gov/repository
GTE _x	临床	https://www.gtportal.org/home/datasets
TMAD	临床	https://tma.im/cgi-bin/home.pl
KIMIA	临床	https://kimialab.uwaterloo.ca/kimia/index.php/data-and-code-2/
Camelyon	临床	https://camelyon17.grand-challenge.org/Data/
TUPAC	临床	http://tupac.tue-image.nl/node/3

2.2.2 预处理

WSI 具有千兆像素,任何使用 DL 模型的现代计算机都无法处理。因此,在训练任何模型之前,都需进行预处理,如数据管理、数据块提取、选择以及将数据拆分为训练集和测试集。WSI 作为拼接的图像存储在计算机上,这些拼接的图像可以通过网格采样提取成更小的、易于管理的小块。通常,这些小块会以所需的放大倍率提取,然后可以用于训练或推理,即在 DL 模型上进行预测以进行验证和测试。

2.2.3 模型选择和训练

目前 DL 方法主要包括监督学习(supervised learning)、无监督学习(unsupervised learning)和弱监督学习(weakly supervised learning)。

2.2.3.1 监督学习 监督学习是一种机器学习(machine learning, ML)模型,使用带有正确标注的

数据作为训练集,需要在训练期间进行监督。对于图像数据,由主题专家(subject matter expert, SME)以手工标注注释的形式提供监督。监督学习的目的是开发一个合适的数学模型,将图像映射到病理学家提供的相应标记,典型任务包括分类和回归。

2.2.3.2 无监督学习 无监督学习是对没有类别标记的样本进行学习,目的是寻找一种能够描述无标记图像模式的数学模型,但还处于非常早期的发展阶段^[18]。该学习方法不需要由专家将输入图像映射到预定义的带注释的标记,而是将图像转化为低维表示,也被称为潜在特征(用于分析数据的更简单表示),这些特征可以基于聚类等各种技术进行分析,以得出有意义的结论。

2.2.3.3 弱监督学习 弱监督学习是监督学习和无监督学习的衍生方法,其不需要监督学习所需要

的密集标记,但可以推断切片级别的标记来预测局部检测。弱监督学习的目的是使用较少或较弱的切片级别的标记训练模型,以预测更精细的像素/块级别的标记。实际上,这些切片级别的标记通常作为元数据使用,而不需要劳动密集型的像素级别标记。因此,弱监督学习是切片分类等任务的一种理想方法。在毒性病理学诊断中,如果 WSI 的任何部分包含病变,WSI 可以被标记为异常,并且这些切片级别的标记可以作为病理学报告中的表格数据。因此,弱监督学习在毒性病理学中可用于异常的筛选。

2.2.4 后处理和评估

经过训练的 DL 模型在投入实际使用之前需经过严格的性能评估,尤其是在决策支持系统中。DL 模型验证应包括将 DL 模型的结果与参考标准之间的比较,可以选择手动注释的测试数据集和/或病理学家做出的诊断作为判定标准^[9,19]。需注意的是,由于观察者之间和观察者内部的差异、病理学家对注释软件的熟悉程度以及视觉和认知陷阱等原因,手工注释的参考标准可能会受到用户诱发的干扰和偏差的影响^[20]。因此,参考标准最好由至少 2 名经验丰富的病理学家确定,如果存在分歧,最好由 3 名经验丰富的病理学家确定^[19,21]。

了解评价方法对于病理学家和 DL 开发人员非常重要,以便双方能够客观地定性和定量评估模型的性能。虽然病理学家对算法性能的主观视觉评估在某些应用中是可以接受的^[22],但模型性能的定量标准可能更适合于客观评价。

3 DL 方法在毒性病理学中的应用

近年来,随着企业研究人员和外部供应商合作以期不同的病理学任务提供基于 DL 的解决措施日益深入,AL 辅助分析方法和 DL 方法在非临床安全性评价和毒性病理学中应用不断增加。DL 方法在毒性病理学领域中的应用主要包含以下 4 类:① 计算机辅助质量控制。② 计算图像分析。③ 计算机辅助异常检测。④ 基于内容的图像检索。

3.1 计算机辅助质量控制

WSI 的人工质量控制通常涉及多个步骤,例如:检查切片元数据以确定准确性和差异性;比较切片上的组织和数字图像以确定扫描质量;评估组织以识别数字伪影(如拼接或失焦区域)。由于非临床病理学实验室制作的切片质量较高,实验室的标准做法通常是审查预先确定一定比例的切片,例如每只动物总切片数的 10%。任何确定为不可接受的

切片都将重新扫描并重新检查质量。WSI 的人工质量控制是一项耗时费力的任务,在每天扫描大量切片的实验室环境中,很容易出现疲劳、人为错误和效率低下的情况。

美国国立组织技术学会和美国病理学家协会发起了名为 HistoQIP 的项目,以提高 WSI 的质量^[23-24]。计算机辅助的质量控制工作流程可以通过在扫描 WSI 时快速审查,并标记需要人工干预的 WSI,可显著改善这一过程。为此,正在为计算机化在线数字质量控制开发自动化工具。HistoQC 是一个开源工具^[25],用于评估 WSI 数据集的异质性,并识别组织切片或数字切片上的人工假象。

3.2 计算图像分析

计算图像分析包括 AI 辅助的客观毒性评分和形态评估,是毒性病理学中的一个新兴领域。长期以来,病理学家试图在不同软件和试验之间协调诊断阈值,但不同病理学家之间和病理学家本身的差异和偏差一直难以克服^[17,22,26]。半定量严重程度评分最常用于识别变化的程度,不同病理学家和机构之间的评分都存在差异和主观性。序数值被分配给不同的严重程度(0 代表正常,1 代表极轻度,2 代表轻度,3 代表中度,4 代表重度,5 代表极重度),数值本身不是比率数据,因此不适合进行参数统计分析^[26]。在传统的病理学实践中,通常会发生诊断漂移。基于 DL 的定量评估可能为病理评估带来一致性和客观性。文献报道 AI 不仅与病理学家的评估具有良好的一致性,而且通常具有更高的敏感性和效率^[27]。此外,DL 模型已被证明可以成功地对心脏^[28]、睾丸^[29-30]、卵巢^[31]、眼睛/视网膜^[32]和肝脏^[33]进行 HE 切片病变的量化形态学评价。基于 DL 的形态量化是一种很有前途的替代评价方法,并可减轻病理学家人工半定量分析的负担。这些结果可用于病理学家对病变性质和严重程度的标准化评价,以及分析与生物特征、分子、药物暴露或临床结果的潜在联系。

3.3 计算机辅助异常检测

异常分类的最终解决方案尚未建立,但它将是毒性病理学中基于 ML/DL 最有前景的应用。这种方法可以筛选出正常切片,以使病理学家有更多的时间关注异常切片。这种工作流程使病理学家能够参与关键的决策过程,确定具体诊断术语和病变的解释。一旦确认了某一异常改变,DL 可以进一步通过组间的标准化分级提供帮助。其在临床诊断病理

学的应用取得了相当大的进展,特别是在辅助癌症诊断和外科活检标本的免疫表型方面,可以加快工作流程并切实提高诊断准确性^[34-36]。

一种基于正常组织识别未知异常的方法被作为检测特定异常的一种替代方法。在这个方法中,模型仅在正常组织上训练,离群点将被标记为异常或病变,可能包含具有毒理学意义的发现,以供病理学家进一步审查。这种方法是以前正常组织的特征为基础。毒性病理学中计算机辅助异常检测主要专注于在正常组织的组织学图谱上训练神经网络。包含46个不同组织的1690张切片以6种放大倍数成像,并用于训练VGG-16, ResNet50和Inception-v3网络以识别组织学上不同的组织^[37]。这些经过训练的大鼠模型可作为非人类灵长类动物和小型猪组织学习的基础。此外,还开发了一种监督DL模型,用于检测选定器官中的选定病变(异常)^[38]。但是,目前的数据集不包括毒性病理学研究中可能检出的所有异常。因为缺乏对于正常的真正定义,毒性病理学更多的是依赖于正常范围内的背景变化。“正常”会随着实验动物年龄、性别、饮食、品系等变化,而“异常”变化则更多种多样,其严重程度范围更加广泛。这些模糊和主观的定义将增加毒性病理学家和DL模型之间的差异。因此,若DL模型能识别出它以前从未见过的异常类型,可能具有巨大的挑战性。

3.4 基于内容的图像检索

如今,人们可以对照片进行索引和挖掘,以识别人、物体和位置,从而允许用户搜索数千张数字照片,这项技术非常成熟,目前正被应用于组织病理学^[39]。然而,与照片不同,WSI通常为千兆像素,约为100 000×100 000像素,这使得索引和挖掘成为一项具有挑战性的任务。基于内容的图像检索,允许根据与查询图像或特征或自然语言查询文本的相似性,从数字切片数据库中搜索和检索图像。

在毒性病理学中,化学品或药物引起的病变有不同的形态和严重程度,有时这些病变可能与背景或自发性改变重叠或加重背景或自发性改变。由于诊断是应用于整个图像或图像集,找到代表性诊断的图像块可能具有挑战性。基于内容的图像检索也可以帮助成为病理学家从大型存储库中快速搜索和发现图像,而不必依赖于对历史案例的记忆。因此,大规模、相对结构化的图像数据库是高效、强大的基于内容的图像检索系统的基础。大多数从事毒性病理学评价的非临床药物安全性评价机构产生或存档

了大量的切片,如果对这些切片进行扫描,就可以构建大量的结构化图像数据,可用于培训和开展大规模基于内容的图像检索系统。

4 毒性病理学中应用DL方法所面临的挑战

尽管AI在DP中有着令人印象深刻的具体案例和应用,但目前毒性病理学中应用DL方法仍面临一些挑战,主要包括监管环境、基础设施、图像格式的大小和放大倍数、注释图像的数量和质量等。

4.1 监管环境

毒性病理学评价通常是监管条件下遵循药物非临床研究质量管理规范(good laboratory practice, GLP)毒理学试验的重要组成部分及金标准。与目前用于GLP研究的其他系统一样,基于DL的软件和方法也需要进行验证。但是,基于DL的软件和方法的验证可能是一项艰巨的任务,因为不仅用于生成WSI的DP系统需要验证,基于AI的软件和算法也需要验证。在早期,美国FDA将WSI系统归类为Ⅲ类医疗设备,即“最高风险”类别,需要全面性控制(质量体系法规、良好生产程序)和上市前批准^[40]。2015年,美国FDA发布了技术性能指南,并在2017年首次批准了一个完整的WSI系统,用于外科病理学的诊断。此次批准后,美国FDA将WSI系统重新归类为Ⅱ类医疗设备,要求制造商通过510(k)符合性技术审核,证明其与先前批准的设备“实质等同”^[41]。

药物非临床安全性和毒性病理学评价是高度监管的实践,因此必须对AI和DL方法做出的任何决定进行稽查和验证。此外,必须设计合适的解释方法来理解预测结果,并为解释和监管目的提供证据。从监管角度来看,不附带任何硬件的用于医疗目的的软件是一种医疗设备的软件(software as medical device, SaMD),而基于DL的SaMD还面临其他挑战,因为其在批准后能够适应或更改。2021年,美国FDA发布了一份讨论文件,对基于AI/ML软件的拟议监管框架征求公众意见^[42],之后根据建议发布了一份基于AI/ML的SaMD行动计划^[43]。2021年7月8日我国国家药品监督管理局(National Medical Products Administration, NMPA)颁布《人工智能医用软件产品分类界定指导原则》用于规范AI软件的注册审批^[44]。

4.2 基础设施

为了能够从基于DL的毒性病理学应用中受益,首先必须建立DP的工作流程。但基础设施是

毒性病理学中实施 DP 工作流程的最大障碍之一^[45],要使 DP 在常规的非临床毒性病理学评价中被接受和采用,该过程必须与使用显微镜评估组织一样或更有效、准确和方便。此外,DP 的实施必须具有成本效益合理,并且获得预算批准。硬件、数据存储系统、网速、加载 WSI 的速度、监视器的刷新率和图像管理软件必须为病理学家提供无延迟或停机的无缝体验。

4.3 图像格式、大小和放大倍数

根据扫描仪的不同,WSI 可能会有不同的图像格式。文件格式的差异可能会导致在 DL 模型开发之前需要额外步骤统一文件格式。虽然医学数字成像和通信(digital imaging and communications in medicine, DICOM)格式提供了互操作性和远程病理诊断的好处,但在 DP 的这个阶段,这还没有被采纳为标准。基于转换的方法通常使用开源或商业程序,以高通量或串联方式将专有文件转化为统一格式^[46]。切片的数字化会迅速产生大量数据,需要经过精心设计的大型存储和缓存数据网络。扫描的放大倍数取决于找到正确的平衡/权衡,以获得最佳的分辨率和图像大小(数据存储考虑因素)。

4.4 注释图像的数量及质量

任何 DL 任务成功与否的一个重要因素是训练数据的丰度。在监督学习中,WSI 本身不足以进行训练,必须进行适当的区域化和标记,以便成功地利用监督学习技术。普通图片的标记信息很容易获取,不需要任何特殊的专业知识。然而,只有经验丰富的病理学家才能准确地标记病理图像。因此,即使有大量可用的图像存储库,也可能没有足够的标记数据(参考标准)用于训练。此外,只有在任务、目的、染色、放大级别和分辨率类似的情况下,手工标注的公共数据集才有用。

图像数据的质量问题主要是不平衡的数据集,因为有些类别很少出现。在毒性病理学中,这一问题尤其在异常检测中更明显,因为大多数组织是正常的,异常可能是高度可变和罕见的。通过扩大代表性不足的类别可能会缓解这种不平衡,但这可能较困难,有时甚至难以实现。其次,就是切片染色差异以及人工假象^[47]。各种人工假象可能会无意中被引入到图像中,从而增加意外噪音并影响 AI 模型的性能。这些人工假象也会干扰切片的显微镜评价。其他一些人工假象仅出现在 WSI 中,与扫描不良有关,如拼接伪影、失焦和组织缺失。在这些情况

下,重新扫描可能有助于提高 WSI 的质量。

5 总结和展望

药物非临床毒理学试验是受试物被批准用于人类临床试验之前评价其安全性的必要阶段。虽然毒性病理学评价目前是评估候选药物靶器官、药物相关病变以及剂量-反应关系的金标准,但人工检查数千张组织切片以期发现受试物相关病变极其繁琐和耗时^[48-50]。本文对 AI 和 DL 方法在毒性病理学中的应用进行了综述,AI 和 DL 在 DP 工作流程中应用的第一步是数字切片质量控制,以帮助准备用于虚拟显微镜和图像分析的高质量 WSI。AI 和 DL 还可用于计算图像分析,并可减少来自毒性病理学家主观评价的诊断漂移。最重要的是,AI 系统可以快速发现与受试物相关的潜在异常,具有极大的实用价值。不仅可以显著缩短毒性病理学家检查数以千计的组织样本所花费的时间,从而缩短药物发现和开发周期,而且可以帮助毒性病理学家进行数据挖掘并快速检索相关图像。

尽管目前 DL 在算法上取得一定的成功,但毒性病理学界对 DL 方法的总体反应还是喜忧参半。当前 DL 方法的一个主要缺点是这些算法的局限性,以及其无法完全模仿和学习毒性病理学家所展示的各种技能和能力。此外,验证 DL 模型和在 GLP 环境中使用此类软件也面临诸多挑战。在毒性病理学的背景下,无监督学习和弱监督学习的概念非常有意义,因为毒性病理学家面对的是可能具有多种病变的多个种属实验动物的各种组织。总之,随着非临床数字病理学数据库不断建立和完善,以及 AI 各种算法和 DL 方法的日趋成熟和不断改进,通过跨学科团队密切合作,AI 和 DL 方法在毒性病理学领域的应用将很快取得长足进步,并将有力促进我国新药的发现和开发。

[参 考 文 献]

- [1] CFSAN. Center for Food Safety and Applied Nutrition, Office of Food Additive Safety. Toxicological Principles for the Safety Assessment of Food Ingredients, Redbook 2000: Chapter IV. C.4. a. Subchronic Toxicity Studies with Rodents; November, 2003[EB/OL]. [2022-04-22]. <https://www.fda.gov/regulatory-information/search-fda-guidance-documents/redbook-2000-ivc4a-subchronic-toxicity-studies-rodents>.
- [2] MARBLE HD, HUANG R, DUDGEON SN, et al. A regulatory science initiative to harmonize and standardize digital pathology and machine learning processes to speed up clinical innovation to patients[J]. *J Pathol Inform*, 2020, 11: 22.
- [3] SRINIDHI CL, CIGA O, MARTEL AL. Deep neural network models for computational histopathology: A survey[J]. *Med Im-*

- age Anal*, 2021;67:101813.
- [4] SALTO-TELLEZ M, MAXWELL P, HAMILTON P. Artificial intelligence-the third revolution in pathology[J]. *Histopathology*, 2019, 74 (3) : 372 – 376.
- [5] TIZHOOSH HR, PANTANOWITZ L. Artificial intelligence and digital pathology: challenges and opportunities[J]. *J Pathol Inform*, 2018, 9: 38.
- [6] BERA K, SCHALPER KA, RIMM DL, *et al*. Artificial intelligence in digital pathology-new tools for diagnosis and precision oncology[J]. *Nat Rev Clin Oncol*, 2019, 16(11) : 703 – 715.
- [7] ZHOU XM, LI C, RAHAMAN MM, *et al*. A comprehensive review for breast histopathology image analysis using classical and deep neural networks [J]. *IEEE Access*, 2020, 8: 90931 – 90956.
- [8] MADABHUSHI A, LEE G. Image analysis and machine learning in digital pathology: challenges and opportunities[J]. *Med Image Anal*, 2016, 33: 170 – 175.
- [9] GAUTHIER BE, GERVAIS F, HAMM G, *et al*. Toxicologic pathology forum: opinion on integrating innovative digital pathology tools in the regulatory framework[J]. *Toxicol Pathol*, 2019, 47 (4) : 436 – 443.
- [10] ACS B, HARTMAN J. Next generation pathology: artificial intelligence enhances histopathology practice [J]. *J Pathol*, 2020, 250(1) : 7 – 8.
- [11] DJURIC U, ZADEH G, ALDAPE K, *et al*. Precision histology: how deep learning is poised to revitalize histomorphology for personalized cancer care[J]. *NPJ Precis Oncol*, 2017, 1(1) : 22.
- [12] PARWANI AV. Next generation diagnostic pathology: use of digital pathology and artificial intelligence tools to augment a pathological diagnosis[J]. *Diagn Pathol*, 2019, 14(1) : 138.
- [13] VAN TONGEREN S, FAGERLAND JA, CONNER MW, *et al*. The role of the toxicologic pathologist in the biopharmaceutical industry[J]. *Int J Toxicol*, 2011, 30(5) : 568 – 582.
- [14] SCHUHMACHER A, GATTO A, HINDER M, *et al*. The upside of being a digital pharma player[J]. *Drug Discov Today*, 2020, 25(9) : 1569 – 1574.
- [15] YOSHIKAWA T, HORAI Y, ASAOKA Y, *et al*. Current status of pathological image analysis technology in pharmaceutical companies: a questionnaire survey of the Japan Pharmaceutical Manufacturers Association [J]. *J Toxicol Pathol*, 2020, 33 (2) : 131 – 139.
- [16] HANNA MG, REUTER VE, ARDON O, *et al*. Validation of a digital pathology system including remote review during the COVID-19 pandemic [J]. *Mod Pathol*, 2020, 33 (11) : 2115 – 2127.
- [17] TURNER OC, AEFNER F, BANGARI DS, *et al*. Society of toxicologic pathology digital pathology and image analysis special interest group article: opinion on the application of artificial intelligence and machine learning to digital toxicologic pathology[J]. *Toxicol Pathol*, 2020, 48(2) : 277 – 294.
- [18] MOULIN P, GRÜNBERG K, BARALE-THOMAS E, *et al*. IMI-bigpicture: a central repository for digital pathology[J]. *Toxicol Pathol*, 2021, 49(4) : 711 – 713.
- [19] ZURAW A, STAUP M, KLOPFLEISCH R, *et al*. Developing a qualification and verification strategy for digital tissue image analysis in toxicological pathology [J]. *Toxicol Pathol*, 2021, 49 (4) : 773 – 783.
- [20] AEFNER F, WILSON K, MARTIN NT, *et al*. The gold standard paradox in digital image analysis: manual versus automated scoring as ground truth[J]. *Arch Pathol Lab Med*, 2017, 141 (9) : 1267 – 1275.
- [21] LIU Y, KOHLBERGER T, NOROUZI M, *et al*. Artificial intelligence-based breast cancer nodal metastasis detection: insights into the black box for pathologists [J]. *Arch Pathol Lab Med*, 2019, 143(7) : 859 – 868.
- [22] HORAI Y, MIZUKAWA M, NISHINA H, *et al*. Quantification of histopathological findings using a novel image analysis platform [J]. *J Toxicol Pathol*, 2019, 32(4) : 319 – 327.
- [23] National Society For Histotechnology. HistoQIP whole slide image quality improvement program (HQWSI) [EB/OL]. (2022 – 01 – 25) [2022 – 04 – 22]. <https://www.nsh.org/learn/histoqip/histoqip-wholeslideimage>.
- [24] BARISONI L, LAFATA KJ, HEWITT SM, *et al*. Digital pathology and computational image analysis in nephropathology [J]. *Nat Rev Nephrol*, 2020, 16(11) : 669 – 685.
- [25] JANOWCZYK A, ZUO R, GILMORE H, *et al*. HistoQC: an open-source quality control tool for digital pathology slides [J]. *JCO Clin Cancer Inform*, 2019, 3: 1 – 7.
- [26] GIBSON-CORLEY KN, OLIVIER AK, MEYERHOLZ DK. Principles for valid histopathologic scoring in research [J]. *Vet Pathol*, 2013, 50(6) : 1007 – 1015.
- [27] AZAM AS, MILIGY IM, KIMANI PK, *et al*. Diagnostic concordance and discordance in digital pathology: a systematic review and meta-analysis [J]. *J Clin Pathol*, 2021, 74 (7) : 448 – 455.
- [28] TOKARZ DA, STEINBACH TJ, LOKHANDE A, *et al*. Using artificial intelligence to detect, classify, and objectively score severity of rodent cardiomyopathy [J]. *Toxicol Pathol*, 2021, 49 (4) : 888 – 896.
- [29] XU J, LU HD, LI HX, *et al*. Computerized spermatogenesis staging (CSS) of mouse testis sections via quantitative histomorphological analysis[J]. *Med Image Anal*, 2021, 70: 101835.
- [30] CREASY DM, PANCHAL ST, GARG R, *et al*. Deep learning-based spermatogenic staging assessment for hematoxylin and eosin-stained sections of rat testes [J]. *Toxicol Pathol*, 2021, 49 (4) : 872 – 887.
- [31] CARBONI E, MARXFELD H, TUOKEN H, *et al*. A workflow for the performance of the differential ovarian follicle count using deep neuronal networks [J]. *Toxicol Pathol*, 2021, 49 (4) : 843 – 850.
- [32] DE VERA MUDRY MC, MARTIN J, SCHUMACHER V, *et al*. Deep learning in toxicologic pathology: a new approach to evaluate rodent retinal atrophy [J]. *Toxicol Pathol*, 2021, 49(4) : 851 – 861.
- [33] PISCHON H, MASON D, LAWRENZ B, *et al*. Artificial intelligence in toxicologic pathology: quantitative evaluation of compound-induced hepatocellular hypertrophy in rats [J]. *Toxicol Pathol*, 2021, 49(4) : 928 – 937.
- [34] RACITI P, SUE J, CEBALLOS R, *et al*. Novel artificial intelligence system increases the detection of prostate cancer in whole slide images of core needle biopsies[J]. *Mod Pathol*, 2020, 33 (10) : 2058 – 2066.
- [35] CAMPANELLA G, HANNA MG, GENESLAW L, *et al*. Clinical-grade computational pathology using weakly supervised deep learning on whole slide images [J]. *Nat Med*, 2019, 25(8) : 1301 – 1309.
- [36] LITJENS G, SÁNCHEZ CI, TIMOFEEVA N, *et al*. Deep learning as a tool for increased accuracy and efficiency of histopathological diagnosis[J]. *Sci Rep*, 2016, 6: 26286.
- [37] HOEFLING H, SING T, HOSSAIN I, *et al*. HistoNet: a deep learning-based model of normal histology [J]. *Toxicol Pathol*, 2021, 49(4) : 784 – 797.
- [38] KUKLYTE J, FITZGERALD J, NELISSEN S, *et al*. Evaluation of the use of single- and multi-magnification convolutional neural networks for the determination and quantitation of lesions in non-clinical pathology studies [J]. *Toxicol Pathol*, 2021, 49 (4) : 815 – 842.
- [39] SCHAUMBERG AJ, JUAREZ-NICANOR WC, CHOUDHURY SJ,

- et al.* Interpretable multimodal deep learning for real-time pan-tissue pan-disease pathology search on social media [J]. *Mod Pathol*, 2020, 33(11): 2169 – 2185.
- [40] LEVY JJ, AZIZGOLSHANI N, ANDERSEN MJ Jr, *et al.* A large-scale internal validation study of unsupervised virtual trichrome staining technologies on nonalcoholic steatohepatitis liver biopsies[J]. *Mod Pathol*, 2021, 34(4): 808 – 822.
- [41] ABELS E, PANTANOWITZ L. Current state of the regulatory trajectory for whole slide imaging devices in the USA [J]. *J Pathol Inform*, 2017, 8: 23.
- [42] U. S. Food and Drug Administration. Proposed regulatory framework for modifications to artificial intelligence/machine learning (AI/ML)-based software as a medical device (SaMD) [EB/OL]. (2020) [2022 – 04 – 22]. <https://www.fda.gov/media/135713/download>.
- [43] U. S. Food and Drug Administration. Artificial Intelligence/Machine Learning (AI/ML)-Based Software as a Medical Device (SaMD) Action Plan[EB/OL]. (2021) [2022 – 04 – 22]. <https://www.fda.gov/news-events/press-announcements/fda-releases-artificial-intelligencemachine-learning-action-plan>.
- [44] 国家药品监督管理局. 人工智能医用软件产品分类界定指导原则 [EB/OL]. [2022 – 04 – 22]. <https://www.nmpa.gov.cn/xxgk/ggtg/qtggtg/20210708111147171.html>.
- [45] SCHUMACHER VL, AEFNER F, BARALE-THOMAS E, *et al.* The application, challenges, and advancement toward regulatory acceptance of digital toxicologic pathology: results of the 7th ESTP international expert workshop (September 20 – 21, 2019) [J]. *Toxicol Pathol*, 2021, 49(4): 720 – 737.
- [46] CLUNIE DA. DICOM format and protocol standardization-a core requirement for digital pathology success [J]. *Toxicol Pathol*, 2021, 49(4): 738 – 749.
- [47] CHLIPALA EA, BUTTERS M, BROUS M, *et al.* Impact of pre-analytical factors during histology processing on section suitability for digital image analysis [J]. *Toxicol Pathol*, 2021, 49(4): 755 – 772.
- [48] 李一昊, 滕伊洋, 张亚群, 等. 毒性病理学中人工智能和机器学习的应用研究进展[J]. 中国新药杂志, 2023, 32(6): 598 – 604.
- [49] 宋佳敏, 王萌, 陈美玲. 高内涵技术在药物线粒体损伤毒性机制研究中的应用[J]. 中国新药杂志, 2021, 30(20): 1861 – 1866.
- [50] 田康, 屈哲, 吕建军, 等. 药物神经毒性非临床安全性评价方法和技术概述[J]. 中国新药杂志, 2021, 30(4): 306 – 311.

编辑:刘卓越/接受日期:2023 – 06 – 28