

· 专家论坛 ·

本文亮点:

- (1) 提出面向严重烧伤与创面修复的全程动态诊疗框架。
- (2) 强调按使用者分层界定人工智能的临床增益。
- (3) 主张通过前瞻性人机比较验证人工智能辅助诊疗的临床有效性。

Highlights:

- (1) A full-course dynamic diagnostic and treatment framework is proposed for severe burn and wound repair.
- (2) The clinical benefit of artificial intelligence in severe burn and wound care is defined according to user roles.
- (3) Prospective comparisons between clinicians alone and clinicians assisted by artificial intelligence are advocated to validate clinical effectiveness.



人工智能赋能严重烧伤与创面修复患者 全程诊疗:从风险预测到因果推断

聂为之¹ 谢克亮²¹天津大学电气自动化与信息工程学院,天津 300072;²天津医科大学总医院重症医学科,天津 300052

通信作者:谢克亮,Email:mkz2011@126.com

【摘要】严重烧伤与复杂创面的诊疗需要统筹局部组织损伤、全身病理生理变化及连续临床干预。现有诊疗路径能够支持医师基于床旁证据进行判断并随患者病情变化调整治疗方案,但在院前与基层评估的一致性、多模态纵向信息整合以及风险预测结果对临床决策的支持等方面仍存在不足。该文结合严重烧伤与复杂创面的诊疗需求及相关人工智能研究证据,分析现有诊疗路径的能力边界、不同使用者可能获得的临床增益,以及风险预测进一步用于受约束因果推断所需的条件。在此基础上,提出“创面评估—全身监测—临床干预”的全程动态诊疗框架:通过标准化图像分析评估创面面积、深度、坏死、感染及愈合趋势;通过时间序列模型整合生命体征、实验室检验、炎症与感染指标、代谢与凝血状态以及器官功能等信息,连续监测患者全身状态;同时将液体复苏、抗感染、清创与创面覆盖、营养支持和康复等临床干预映射至统一时间轴,以保留患者状态变化、干预时机

与临床结局之间的时间序列关系。风险预测用于估计现有诊疗路径下可能发生的临床结局,而反事实推演必须建立在明确的因果假设、适用条件、不确定性估计和责任边界之上,其结果不能直接作为治疗处方。人工智能的临床价值应通过前瞻性研究直接比较医师独立判断与医师结合人工智能判断的效果得出;评价指标除区分度外,还应包括评估耗时、医师间一致性、校准度、预警提前时间、警报与工作流程负担、决策改变情况、患者安全及临床结局。该文围绕现有诊疗路径的能力边界、智能创面评估、全身风险监测、因果推断以及转化应用与治理,讨论人工智能参与严重烧伤与创面修复患者全程诊疗的路径及边界。

【关键词】 人工智能; 烧伤; 创伤和损伤; 伤口愈合; 重症监护; 风险评估; 因果推断

基金项目:国家自然科学基金面上项目(62671416)

DOI:10.3760/cma.j.cn501225-20260718-00265

收稿日期 2026-07-18

引用本文:聂为之, 谢克亮. 人工智能赋能严重烧伤与创面修复患者全程诊疗:从风险预测到因果推断[J]. 中华烧伤与创面修复杂志, 2026, 42(9): 1-8. DOI: 10.3760/cma.j.cn501225-20260718-00265.

Nie Weizhi, Xie Keliang. Artificial intelligence-enabled full-course diagnosis and treatment for patients with severe burns and wounds requiring repair: from risk prediction to causal inference[J]. Chin J Burns Wounds, 2026, 42(9): 1-8. DOI: 10.3760/cma.j.cn501225-20260718-00265.



Artificial intelligence-enabled full-course diagnosis and treatment for patients with severe burns and wounds requiring repair: from risk prediction to causal inference

Nie Weizhi¹, Xie Keliang²

¹School of Electrical and Information Engineering, Tianjin University, Tianjin 300072, China; ²Department of Critical Care Medicine, Tianjin Medical University General Hospital, Tianjin 300052, China

Corresponding author: Xie Keliang, Email: mkz2011@126.com

【Abstract】 The diagnosis and treatment of severe burns and complex wounds require the coordinated management of local tissue injury, systemic pathophysiological changes, and sequential clinical interventions. Existing diagnostic and treatment pathways support bedside evidence-based decision-making and allow clinicians to adjust treatment as a patient's condition evolves. However, limitations remain in the consistency of prehospital and primary-care assessments, the longitudinal integration of multimodal information, and the use of risk predictions to support clinical decisions. Based on the clinical requirements of severe burn and complex wound care and the available evidence on artificial intelligence, this expert commentary examines the capability boundaries of current care pathways, the potential clinical benefits for different users, and the conditions required to extend risk prediction to constrained causal inference. We propose a full-course dynamic diagnostic and treatment framework comprising wound assessment, systemic monitoring, and clinical intervention. Standardized image analysis is used to assess wound area, depth, necrosis, infection, and healing trajectories. Time-series models integrate vital signs, laboratory measurements, inflammatory and infectious markers, metabolic and coagulation status, and organ function to continuously monitor the patient's systemic condition. Clinical interventions including fluid resuscitation, antimicrobial treatment, debridement and wound coverage, nutritional support, and rehabilitation, are aligned along a unified timeline to preserve the temporal relationships among changes in patient status, intervention timing, and clinical outcomes. Risk prediction estimates outcomes that may occur under the current care pathway, whereas counterfactual simulation requires explicit causal assumptions, clearly defined applicability conditions, uncertainty estimates, and accountability boundaries, and should not be directly interpreted as a treatment prescription. The clinical value of artificial intelligence should be established prospectively by comparing clinicians working independently with clinicians assisted by artificial intelligence. In addition to discrimination, evaluation should include assessment time, inter-clinician agreement, calibration, warning lead time, alert and workflow burden, changes in clinical decisions, patient safety, and clinical outcomes. The article discusses the

pathway and boundaries of artificial intelligence-enabled full-course care from the perspectives of current pathway capabilities, intelligent wound assessment, systemic risk monitoring, causal inference, and translational implementation and governance.

【Key words】 Artificial intelligence; Burns; Wounds and injuries; Wound healing; Critical care; Risk assessment; Causal inference

Fund program: General Program of National Natural Science Foundation of China (62671416)

严重烧伤并非单纯的皮肤损伤。皮肤屏障破坏、毛细血管渗漏、炎症失衡及高代谢反应可共同导致循环不稳定、感染易感及器官功能损害^[1-2]。因此,其诊疗既包括烧伤总面积和深度判断、清创及创面覆盖,也包括液体复苏、感染防治、营养支持、器官保护与长期康复^[3-5]。复杂创面的病因、组织构成及愈合条件同样具有明显异质性。临床医师不仅要判断创面当前状态,还要持续评估其变化趋势及其与患者全身状态的关系。

计算机视觉、机器学习与大模型已被用于烧伤面积估计、深度分类、创面分割、感染识别、愈合评价和临床结局预测。算法之间的性能比较只能说明技术差异,尚不足以证明其较现行诊疗路径具有增量价值。现有体格检查、严重程度评分、影像学检查、组织评估和多学科团队讨论,均由临床医师结合床旁证据进行解释;医师还可追溯判断依据,并随病情变化及时调整诊疗方案。要判断人工智能能否缩短评估时间、降低观察者差异、整合复杂信息或改善临床结局,必须将其与上述现行方法直接比较。作为问题导向的专家述评,本文主要探讨严重烧伤与创面修复全程诊疗尚存的缺口、不同经验层级和应用场景所需的智能支持,以及模型从风险预测走向因果推断所需的证据与治理条件。结合笔者团队的既往研究,全文梳理“创面评估—全身监测—临床干预”的技术路径及其临床边界。

1 现有临床诊疗的能力边界与智能化需求

严重烧伤与复杂创面的临床管理已有基本方案。院前与急诊阶段可依据致伤机制、生命体征、烧伤总面积与部位进行伤情分层和转诊;住院阶段可通过连续体格检查、实验室检验、微生物培养及其他病原学检查、器官功能评估和严重程度评分,识别休克、感染和多器官功能障碍。同时,临床医师需根据深度、组织灌注、坏死、感染及愈合趋势反复评估创面,并结合清创、创面覆盖、抗感染、营养

支持和康复需求制订多学科诊疗方案^[6-9]。这些现行方法以患者的病理生理变化为依据,并依赖医师的临床经验;医师能够据此解释判断依据,并随病情变化及时调整诊疗方案。这构成人工智能系统必须比较并嵌入的现行诊疗基准,而非应被否定或替代的对象。

现有诊疗路径的不足主要体现在评估时效、结果一致性、信息整合和风险解释等方面。院前、基层或大批量伤员场景中的专业经验和资源有限,难以及时、准确地评估烧伤总面积、深度及转诊紧迫性。创面外观、面积测量和感染判断存在观察者差异,连续记录还会受到拍摄条件、时间点和标注方式的影响。局部创面、生命体征、实验室检验结果、器官功能和治疗措施具有高维、异步和纵向等特征,人工逐项整合负担较重,部分异常趋势可能在单项指标越过阈值后才被识别。传统评分多用于静态分层,对风险由何种变化驱动、何时复核以及哪项干预可能改变轨迹的提示有限。人工智能应针对这些可量化的临床需求发挥补充作用,技术上的“更先进”本身不能构成应用依据。

人工智能的目标使用者和使用方式也应分层界定。院前人员和非烧伤专科医护人员更需要标准化的烧伤总面积与深度估计、危险部位识别和升级转诊提示,以减少漏判并提高评估一致性;青年烧伤专科医师和烧伤专科护士更需要连续测量、趋势显示、异常提醒及结构化记录,以提高复核效率;经验丰富的烧伤专科医师及其他多学科团队成员通常不需要临床预测或决策模型替代基础判断,其潜在增益主要在于整合跨模态纵向信息、识别不易察觉的轨迹偏离、量化不确定性和比较候选路径。所有层级均应保留原始证据、模型适用范围和人工否决通道,且高风险决策必须由具备相应资质的医师负责。

评价人工智能的临床增益,需将现行诊疗路径作为对照。应在不同经验层级、不同中心和真实工作流程中开展研究,比较医师单独判断与医师结合人工智能判断的表现;指标需对应具体临床需求,包括评估耗时、医师间一致性、校准度、预警提前时间、假阳性和工作流程负担、转诊或复核是否及时、决策是否改变以及患者安全和临床结局。只有当模型产生可重复的增量获益,且未增加不可接受的伤害和工作负担时,才具备临床应用价值。严重烧伤与创面诊疗中现有临床方法与人工智能辅助方

法的比较见表 1。

2 人工智能从局部创面识别拓展至全身风险预测

2.1 人工智能支持局部创面识别与动态评估

烧伤总面积是判断是否需要液体复苏、估算初始复苏量及进行转诊分层的重要依据,烧伤深度及部位则主要影响创面处理与手术策略,三者的临床评估均可能受到观察者经验、患者体位和创面形态影响^[3-4]。既往系统评价显示,机器学习已被用于烧伤面积预估、深度判断、植皮需求预测和瘢痕评估,但不同研究的数据来源、标签标准和验证方法差异明显,多数系统仍处于早期阶段^[10-11]。一般创面的图像人工智能研究还覆盖创面定位、边界分割、组织分类和愈合趋势分析,但商业化测量工具不能等同于已在烧伤患者中完成外部验证的临床决策系统^[12]。因此,人工智能系统的算法输出必须由临床医师结合床旁体格检查结果、致伤时间、创面部位、组织灌注状态和后续演变进行解释,不能脱离临床情境单独用于决策。

国内外研究已将卷积神经网络用于烧伤深度识别和面积估计。一项国内单中心诊断试验纳入 221 例烧伤患者伤后 48 h 内的 484 张创面照片,由 3 名具有 5 年以上专科经验的烧伤整形外科主治医师标注浅 II 度、深 II 度和 III 度烧伤,进而构建 ResNet-50 分类模型^[13]。另一项单中心回顾性研究从患者病历中筛选 2 591 幅烧伤创面图像,由 2 名烧伤外科医师形成一致性标注,采用 U-Net 和 Mask R-CNN 分割创面,并结合手掌图像估算烧伤总面积^[14]。这些研究提示图像算法在创面量化方面的技术可行性,但其数据按图像或图像块划分训练集、验证集和测试集,仍受到单中心来源、参考标准一致性、选择性排除、同一患者多图像相关性及独立外部验证不足等限制。创面图像分析模型还可能将标尺、敷料、拍摄背景、拍摄角度、拍摄光照乃至医院特有流程当作诊断依据。模型若依赖这些与病理特征无关但与标签相关的线索,可能在内部测试中获得较高准确率,却不能代表对真实病理特征的可靠识别。

笔者团队在胸部 X 线图像分类中从因果角度分析图像特征,利用结构因果模型和后门调整削弱混杂因素的影响^[15];进一步将工具变量学习方法与电子病历语义信息相结合,以减少影像噪声与伪相关^[16]。笔者团队还探索了先验知识引导的强化学

表1 严重烧伤与创面诊疗中现有临床方法与人工智能辅助方法的比较

Table 1 Comparison of current clinical and artificial intelligence-assisted approaches in severe burn and wound care

诊疗任务	现有临床方法	现有临床方法 主要优势	现有临床方法 主要瓶颈	人工智能合理增益	人工智能主要使 用者	应比较的临床终 点
院前或急诊 伤情分层 与转诊	致伤史、生命体征、 烧伤总面积与部 位评估	快速、可解释、便 于转诊沟通	评估较依赖临床经验； 在非烧伤专科人员 参与或大批量伤员 场景下，评估结果的 一致性不足	辅助开展烧伤面积 与深度的标准化 量化评估，并提示 关键高风险部位	院前人员、非烧伤 专科医护人员	评估耗时、评估结 果一致性、伤情 漏判率、转诊及 时效性
局部创面动 态评估	连续体格检查、带 标尺的标准化摄 影、组织灌注评 估及微生物培养	可结合触感、气味 及床旁背景即 时纠偏	记录格式和采集条件 不一致，观察者差异 明显，不同时间点的 变化难以量化	自动分割、配准、变 量化及图像质 量提示	青年烧伤专科医 师、烧伤专科护 士及负责复核 的专家	测量误差、重复 性、复核时间、 错误提示率及 治疗时机改变
全身严重程 度与风险 判断	生命体征、实验室 检验、器官功能、 严重程度评分及 临床查房	机制较明确，可随 病情即时修正	静态阈值难以捕捉异 步多变量轨迹，整合 负担高	融合纵向数据并显 示风险趋势和不 确定性	包含烧伤、重症、 感染和康复等 专科医师在内 的多学科团队	校准度、预警提前 时间、警报负 担、响应率及器 官结局
治疗与修复 决策	指南、专科经验及 多学科团队讨论	可纳入解剖条件、 可用医疗资源、 患者意愿及未 结构化信息	复杂方案比较依赖经 验，群体层面的证据 难以直接迁移到个 体患者	汇总证据、识别异常 轨迹并给出可核 查候选路径	负责最终决策的 经验丰富的烧 伤专科医师	决策改变情况、临 床净获益、并发 症发生及功能 结局
因果推断和 反事实比 较	机制知识、临床研 究及专家判断	临床责任主体明 确，可结合具体 临床情境解释 干预与结局之 间的关系	观察性数据易受混杂 因素影响，且个体在 未接受干预下的反 事实结局无法直接 观测	仅在明确因果图、适 用条件及不确定 性前提下开展研 究性模拟	包含烧伤专科医 师在内的多学 科专家团队	外部验证、前瞻性 验证、敏感性分 析、心理伤害、 知情同意及责 任追溯

注：人工智能的对照对象应是现行诊疗路径，而非另一种算法；高风险决策须保留人工复核步骤和医师否决通道

习方法，通过视觉、语义和诊断智能体协同，缓解医学图像标注不足和持续学习困难等问题^[17]。在医学图像分割方面，因果启发的无监督领域适应可利用目标域风格信息改善跨域迁移性能^[18]。这些研究均未在烧伤创面人群中完成直接验证，只能为烧伤创面建模提供方法学启示，其可迁移性仍需在烧伤专科、多中心数据中检验。创面图像智能分析模型应尽可能学习组织损伤、组织灌注和感染相关特征，而非依赖数据采集环境；同时应主动应对不同中心、设备和人群之间的分布差异。

人工智能系统分析局部创面时需纳入时间维度。单一时间点的分类结果所含信息有限，临床决策更依赖坏死组织范围、组织灌注、感染征象和愈合轨迹的连续变化。连续图像可与热成像、荧光血管造影所反映的组织灌注信息、微生物培养结果、组织氧合和可穿戴传感信息联合，用于更新创面状态。创面图像分析模型除输出面积或类别外，还需报告变化幅度、可信度和可能的异常来源，为复查、清创和修复时机提供可核查依据。

2.2 人工智能支持严重烧伤患者全身风险预测

局部创面状态不能完整反映严重烧伤患者的整体风险。严重烧伤除造成局部组织损伤外，还可因体液丢失、全身炎症反应和感染等引起循环、呼吸、代谢及器官功能异常，病情较重者需要在重症监护环境中接受连续监测和器官支持。因此，对于需要重症监护的严重烧伤患者，人工智能风险预测不能仅依赖创面信息，还应整合生命体征、实验室检验、器官功能和治疗记录等纵向数据。然而，这类重症监护数据通常具有缺失率高、采样不同步、变量相关性强和病情变化迅速等特点。MIMIC-IV (Medical Information Mart for Intensive Care IV) 等开放数据库为时间序列建模提供了研究基础^[19]。笔者团队提出时空相关注意网络，可根据不同临床任务选择关键指标和关键时间点，用于死亡、住院时间、生理恶化和临床表型等预测^[20]。该研究表明，重症患者的死亡和生理恶化风险通常由多项生理变量在特定时间窗口内共同体现，难以通过单一指标完整概括。但需要指出的是，MIMIC-IV 主要包含

一般重症患者数据,其研究结果不能直接作为针对严重烧伤患者的有效证据,相关模型仍需在烧伤专科队列中进行校准和外部验证。

脓毒症预警是严重烧伤智能诊疗的重要应用场景,但也容易产生误判。烧伤后脓毒症的诊断会受到烧伤本身所致炎症反应、高代谢状态及器官功能波动的干扰^[5-6];烧伤创面感染的定义和评价指标在既往研究中差异显著,限制了研究间证据合并及模型标签的一致性^[8-9]。针对非烧伤 ICU 患者的系统评价也显示,现有脓毒症机器学习研究在定义、预测窗口、数据处理和验证策略方面高度异质,且多数仍停留在离线回顾性评价^[21]。笔者团队提出 CISepsis 框架,利用后门调整、工具变量和互信息约束处理临床文本及静态指标中的潜在混杂因素,并基于 MIMIC-IV 数据库开展脓毒症早期预测^[22],随后又通过序列化因果解耦,区分脓毒症相关因素、其他疾病因素和无关混杂因素,并分析不同时间窗口中的关键临床指标^[23]。然而,这些结果来源于非烧伤 ICU 数据,不能直接外推为烧伤患者中的有效性证据。

上述因果建模和序列化风险预测方法为烧伤重症预警提供了思路,但不能替代在烧伤人群中的独立验证。严重烧伤患者的感染定义、复苏阶段和生理基线具有特殊性,通用脓毒症模型可能出现校准偏差。一项纳入 28 项诊断研究、1 517 例住院烧伤患者的系统评价与荟萃分析显示,降钙素原诊断脓毒症的合并敏感度和特异度分别为 73% 和 75%,C 反应蛋白的合并敏感度和特异度分别为 86% 和 54%,白细胞计数的合并敏感度和特异度分别为 47% 和 65%;研究间患者构成、脓毒症定义和阈值亦存在异质性,因此单一生物标志物不足以单独用于脓毒症早期诊断^[24]。烧伤相关急性肾损伤等器官损害还可能由炎症、低灌注和微循环障碍等因素共同作用所致^[25]。因此,烧伤风险预测需要整合连续生理数据、创面变化和治疗过程,并在烧伤专科队列中重新评价区分度、校准度和临床净获益。

受试者操作特征曲线下面积只能反映风险预测模型的部分性能,不同算法在曲线下面积上的小幅差异也不能直接证明风险预测模型的临床价值。应在相同患者、相同信息和相同时间预算下评价人工智能风险预测模型的增量价值,直接比较医师单独判断与医师结合人工智能判断,观察评估耗时或预警提前时间、不同经验层级医师间的一致性、跨

中心校准、漏判与假阳性负担,以及复核、转诊或干预时机是否改变。若离线准确率的提高未改善复核、转诊或干预的及时性,亦未提高患者安全性或改善其临床结局,则只能说明技术可行性。风险预测用于判断患者可能出现的结局;因果推断则进一步考察风险形成的原因,以及改变干预后结局可能发生的变化。

3 人工智能推动风险预测走向因果推断

3.1 人工智能由相关性学习转向因果推断

真实世界医疗数据并非在随机条件下产生。病情越重的患者往往接受更积极的治疗,而治疗又会改变后续指标和结局。如果临床预测模型忽略这一过程,可能把治疗与不良结局的伴随关系误判为治疗导致不良结局,也可能把某种医院流程误认为疾病特征。笔者团队在接受冠状动脉旁路移植术患者的术后并发症预测中构建结构因果模型,利用后门调整处理术前信息和非因果性术中信息中的潜在混杂因素,并结合少样本学习处理标注数据有限的问题^[26]。该研究从方法学层面尝试减少混杂因素形成的伪相关并提取较稳定的预测特征,但基于观察性数据的混杂因素控制并不等同于已经识别出可干预的临床因果效应;能否识别此类因果效应,仍取决于因果图的合理性、关键混杂因素是否被测量,以及外部验证和前瞻性验证。

因果推断不能简化为在深度学习模型上增加一个因果模块。建模前需依据临床知识明确时间顺序和可能的因果路径,区分基线因素、暴露因素、治疗因素、中介因素和结局;随后判断哪些混杂因素可观测并可控制,同时标明仍然缺失的关键变量。反事实问题还应对应清晰、可执行的临床干预,笼统的患者状态变化无法形成可验证的估计目标。相关方法学研究同样主张区分面向可行动医疗的预测与因果估计:预测描述既有诊疗模式下可能出现的结果,因果估计考察改变某项措施后结果可能如何变化^[27]。

在严重烧伤及其他复杂创面的诊疗中,可围绕具体临床问题逐步建立因果图。烧伤总面积和全身炎症反应负荷密切相关,烧伤深度、部位和组织灌注状态主要影响创面处理与手术策略;患者年龄、伤前基础状态(包括既往基础疾病)、吸入性损伤情况,均可能同时影响治疗选择和结局。清创、创面覆盖和针对明确感染的治疗可通过控制感染

源和改善创面环境影响创面愈合及全身状态,低灌注、营养不足和免疫功能紊乱则可能导致创面恶化与器官功能损害。明确这些关系有助于确定模型应控制、保留和解释的变量,也能避免把“预测高风险”直接等同于“某项治疗有效”。

3.2 人工智能驱动的全病程推演

笔者建议用创面评估、全身监测和临床干预 3 个相互作用的层次来表示严重烧伤与创面全程诊疗。创面评估层记录面积、深度、坏死、组织灌注、感染、上皮化和瘢痕演变,全身监测层记录循环状态、呼吸功能、炎症与免疫反应、代谢与凝血状态以及器官功能,临床干预层记录液体复苏、抗感染、清创、手术覆盖创面、营养支持和康复措施。3 个层次应映射到统一时间轴,保留每次观察与干预的先后关系。诊疗模型据此持续更新患者状态及其可能的发展路径,避免仅保留一次性风险分数。

在这一框架下,人工智能可形成 3 类输出:状态估计利用不完整、异步和多模态数据实现创面评估和全身监测;风险推演预测当前治疗路径下创面感染、器官功能损伤、手术失败或创面延迟愈合的可能性;反事实推演则在明确适用条件和混杂假设后,估计不同干预时机或方案对应的可能轨迹。反事实推演结果属于需临床复核的模型估计,不能直接作为治疗处方。

笔者团队提出的 Organ-Agents 以不同智能体表示心血管、肾脏和免疫等器官或生理系统,利用重症时间序列数据学习多器官变量间的动态联系,并探索不同脓毒症治疗策略下的反事实轨迹^[28]。该成果目前在出版中,模拟结果主要说明方法学可行性,尚不能据此证实特定治疗对患者生理状态或临床结局的因果效应,也不能直接指导临床干预。其方法学意义在于将患者状态分解为多个相互作用的器官和生理系统,并据此研究系统间的动态联系。至于该框架能否纳入创面状态,仍需使用烧伤专科数据验证局部组织损伤、感染负荷与全身器官反应之间的联系,同时报告不确定性和失败条件。

医学数字孪生利用个体纵向数据构建可更新的虚拟表征,并在计算环境中比较潜在干预^[29]。现有研究已在心血管等领域讨论机制模型与数据驱动模型的互补关系^[30],近期工作进一步关注跨尺度数据整合、系统交互操作和临床界面^[31]。用于烧伤与创面时,数字孪生宜先处理可验证的局部问题,但现阶段尚缺乏完整复制患者情况的验证基础。

例如,可先模拟创面面积和组织灌注变化与全身炎症、循环功能及肾功能的关系,再逐步加入清创、创面覆盖和抗感染等干预。每增加一个模块,均需明确数据来源、适用人群、不确定性和失败条件。

因果推断的最终目标仍是分层人机协同。院前人员和非烧伤专科医护人员可使用受限的标准化评估与转诊提示,但不应接收超出其处置能力的复杂治疗建议;青年烧伤专科医师可借助证据定位、风险趋势和复核清单提高判断一致性;经验丰富的烧伤专科医师及其他多学科团队成员则负责审查因果假设、识别未观测因素,并权衡功能、风险、可用医疗资源和患者意愿。尤其在清创范围、手术时机和抗感染方案等高风险决策中,临床决策模型应展示原始依据、适用条件、不确定性和替代解释,并允许医师否决或退出,而不能越过临床责任主体直接做出治疗决定。

4 人工智能临床转化面临的挑战与展望

多中心纵向数据库是全程诊疗模型的基础。数据需覆盖标准化创面图像、烧伤总面积、局部创面面积与深度、生命体征、实验室检验结果和微生物培养结果、器官功能、治疗时间及长期功能结局。图像采集应记录设备、距离、角度、光照和标尺,临床数据应保留真实时间戳和缺失原因。创面感染与愈合、器官功能障碍和康复结局还需采用可执行的统一定义。只有保证这些变量在不同中心之间可比,才能开展可靠的算法比较。

模型评价的重点应从内部准确率转向相对于现行诊疗路径的临床有效性。医疗人工智能转化研究指出,数据偏倚、外部验证不足、临床工作流程不匹配和监管责任不清,都会削弱模型的实际临床影响^[32]。脓毒症算法的部署研究也显示,预警呈现方式、医师响应、工作负担和实施环境会与模型性能共同影响最终效果^[33]。烧伤与创面模型需报告区分度、校准度、预警提前时间、不同人群表现和跨中心稳定性,并在真实工作流程中开展前瞻性比较,按医师经验分层评价医师单独判断与医师结合人工智能判断在评估耗时、一致性、决策改变情况、警报负担、患者安全和临床结局方面的差异。因此,无法改善决策或结局的系统应界定为测量或科研工具,不宜称为临床决策系统。

透明报告和前瞻性评价应贯穿人工智能研究全过程。TRIPOD+AI (Transparent Reporting of a

multivariable prediction model for Individual Prognosis Or Diagnosis+Artificial Intelligence) 已对预测模型研究的数据来源、样本处理、性能评价和可解释性报告提出更新要求^[34]; 涉及人工智能干预的临床试验还可参考 CONSORT-AI (Consolidated Standards of Reporting Trials-Artificial Intelligence) 和 SPIRIT-AI (Standard Protocol Items: Recommendations for Interventional Trials-Artificial Intelligence) 明确系统版本、输入输出、人工参与、错误处理及运行环境^[35-36]。据此, 烧伤与创面人工智能研究应在项目设计阶段纳入这些规范, 论文完成后的补充说明难以弥补方案缺陷。

数据隐私、模型偏倚和责任边界需要与技术研发同步处理。创面图像可能包含面部、纹身或其他可识别信息, 跨中心共享数据时应进行授权、脱敏和访问控制。患者在最初知情同意时往往无法预见数据未来会与何种模态链接、用于训练何种模型、由哪些机构再次使用, 因此一次性、笼统授权不足以覆盖持续演化的人工智能用途。可根据风险采用分层或动态同意、再次联系、退出机制和独立数据治理, 清楚说明二次使用、商业合作、模型更新和数据保存期限, 并对失去同意能力的危重患者设置代理同意与事后确认程序^[37]。

反事实推演还可能产生不同于常规风险预测的心理和关系风险。向患者展示“若更早治疗可能获得更好结局”等模拟结果, 可能诱发后悔、自责或对既往医师的不信任; 向医师展示不利反事实, 也可能导致防御性决策、自动化依赖或责任焦虑。此类结果不应作为确定事实或追溯责任的证据呈现, 而应同时说明模型假设、不可观测混杂因素、置信区间和替代解释, 并由经过培训的临床团队结合患者意愿决定是否、何时及如何沟通。部署前还应明确开发者、医疗机构和临床使用者在数据质量、模型更新、异常处置和最终决策中的责任, 持续监测不同肤色、年龄、部位和资源条件下的公平性与性能漂移情况, 保留人工复核、申诉、异常上报和模型退出机制^[27,38]。

现阶段应优先提高创面量化和重症风险预警的标准化程度, 在现行诊疗路径中完成多中心外部验证, 并前瞻性比较医师单独判断与医师结合人工智能判断的效果。在具备可靠验证基础后, 可进一步纵向融合创面图像、全身生理数据和治疗记录, 形成可解释、可校准、可审计的动态患者状态模型。

数字孪生和多智能体生理模拟涉及对不同于干预路径的比较, 只有在因果假设、伦理审查和前瞻性验证均得到落实后才适合探索。在明确使用者、临床需求和责任边界的条件下, 人工智能能否比现行诊疗路径更及时、更一致地整合信息, 并使关键时间窗内的决策和患者结局获得可重复的净改善, 是评价其临床增益的关键问题, 最终需通过可检验的研究结果予以回答。

利益冲突 所有作者声明不存在利益冲突

参考文献

- [1] Jeschke MG, van Baar ME, Choudhry MA, et al. Burn injury[J]. Nat Rev Dis Primers, 2020, 6(1): 11. DOI: 10.1038/s41572-020-0145-5.
- [2] Greenhalgh DG. Management of burns[J]. N Engl J Med, 2019, 380(24): 2349-2359. DOI: 10.1056/NEJMra1807442.
- [3] ISBI Practice Guidelines Committee, Steering Subcommittee, Advisory Subcommittee. ISBI practice guidelines for burn care[J]. Burns, 2016, 42(5): 953-1021. DOI: 10.1016/j.burns.2016.05.013.
- [4] Rowan MP, Cancio LC, Elster EA, et al. Burn wound healing and treatment: review and advancements[J]. Crit Care, 2015, 19: 243. DOI: 10.1186/s13054-015-0961-2.
- [5] Zhang P, Zou B, Liou YC, et al. The pathogenesis and diagnosis of sepsis post burn injury[J/OL]. Burns Trauma, 2021, 9: tkaa047[2026-07-18]. <https://pubmed.ncbi.nlm.nih.gov/33654698/>. DOI: 10.1093/burnst/tkaa047.
- [6] Dvorak JE, Ladhani HA, Claridge JA. Review of sepsis in burn patients in 2020[J]. Surg Infect (Larchmt), 2021, 22(1): 37-43. DOI: 10.1089/sur.2020.367.
- [7] Singer M, Deutschman CS, Seymour CW, et al. The third international consensus definitions for sepsis and septic shock (Sepsis-3) [J]. JAMA, 2016, 315(8): 801-810. DOI: 10.1001/jama.2016.0287.
- [8] Lee KS, Young A, King H, et al. Variation in definitions of burn wound infection limits the validity of systematic review findings in burn care: a systematic review of systematic reviews[J]. Burns, 2022, 48(1): 1-12. DOI: 10.1016/j.burns.2021.05.006.
- [9] Greenhalgh DG, Kiley JL. Diagnosis and treatment of infections in the burn patient[J]. Eur Burn J, 2024, 5(3): 296-308. DOI: 10.3390/ebj5030028.
- [10] Huang S, Dang J, Sheckter CC, et al. A systematic review of machine learning and automation in burn wound evaluation: a promising but developing frontier[J]. Burns, 2021, 47(8): 1691-1704. DOI: 10.1016/j.burns.2021.07.007.
- [11] Taib BG, Karwath A, Wensley K, et al. Artificial intelligence in the management and treatment of burns: a systematic review and meta-analyses[J]. J Plast Reconstr Aesthet Surg, 2023, 77: 133-161. DOI: 10.1016/j.bjps.2022.11.049.
- [12] Anisuzzaman DM, Wang C, Rostami B, et al. Image-based artificial intelligence in wound assessment: a systematic review[J]. Adv Wound Care (New Rochelle), 2022, 11(12): 687-709. DOI: 10.1089/wound.2021.0091.
- [13] 何志友, 王元, 张丕红, 等. 基于卷积神经网络的人工智能烧伤深度识别模型的建立及测试效果[J]. 中华烧伤杂志, 2020, 36(11): 1070-1074. DOI: 10.3760/cma.j.cn501120-201

- 90926-00385.
- [14] Chang CW, Lai F, Christian M, et al. Deep learning-assisted burn wound diagnosis: diagnostic model development study [J]. *JMIR Med Inform*, 2021, 9(12): e22798. DOI: 10.2196/22798.
 - [15] Nie W, Zhang C, Song D, et al. Chest X-ray image classification: a causal perspective[C]//Greenspan H, Madabhushi A, Mousavi P, et al. *Medical Image Computing and Computer Assisted Intervention-MICCAI 2023*. Cham: Springer, 2023: 25-35. DOI:10.1007/978-3-031-43898-1_3.
 - [16] Nie W, Zhang C, Song D, et al. Instrumental variable learning for chest X-ray classification[C]//2023 IEEE International Conference on Systems, Man, and Cybernetics. Piscataway: IEEE, 2023: 4506-4512. DOI: 10.1109/SMC53992.2023.10394601.
 - [17] Nie W, Zhang C, Song D, et al. Deep reinforcement learning framework for thoracic diseases classification via prior knowledge guidance[J]. *Comput Med Imaging Graph*, 2023, 108:102277. DOI:10.1016/j.compmedimag.2023.102277.
 - [18] Song J, Chen H, Lyu Y, et al. Causality-inspired unsupervised domain adaptation with target style imitation for medical image segmentation[J]. *IEEE Trans Circuits Syst Video Technol*, 2025, 35(10): 10175-10187. DOI: 10.1109/TCSVT.2025.3562650.
 - [19] Johnson AEW, Bulgarelli L, Shen L, et al. MIMIC-IV, a freely accessible electronic health record dataset[J]. *Sci Data*, 2023, 10(1):1. DOI:10.1038/s41597-022-01899-x.
 - [20] Nie W, Yu Y, Zhang C, et al. Temporal-spatial correlation attention network for clinical data analysis in intensive care unit[J]. *IEEE Trans Biomed Eng*, 2024, 71(2): 583-595. DOI: 10.1109/TBME.2023.3309956.
 - [21] Moor M, Rieck B, Horn M, et al. Early prediction of sepsis in the ICU using machine learning: a systematic review[J]. *Front Med (Lausanne)*, 2021, 8: 607952. DOI: 10.3389/fmed.2021.607952.
 - [22] Li Q, Li D, Jiao H, et al. CISEps: a causal inference framework for early sepsis detection[J]. *Front Cell Infect Microbiol*, 2024, 14:1488130. DOI:10.3389/fcimb.2024.1488130.
 - [23] Li Q, Li D, Nie W, et al. Temporal and spatial analysis in early sepsis prediction via causal disentanglements[J]. *IEEE Trans Knowl Data Eng*, 2025, 37(8): 4860-4872. DOI: 10.1109/TKDE.2025.3569584.
 - [24] Li AT, Moussa A, Gus E, et al. Biomarkers for the early diagnosis of sepsis in burns: systematic review and meta-analysis[J]. *Ann Surg*, 2022, 275(4): 654-662. DOI: 10.1097/SLA.0000000000005198.
 - [25] Schult L, Halbgebauer R, Karasu E, et al. Glomerular injury after trauma, burn, and sepsis[J]. *J Nephrol*, 2023, 36(9): 2417-2429. DOI:10.1007/s40620-023-01718-5.
 - [26] Zhang Q, Zhang W, Li Q, et al. Causal inference model for accurate medical diagnosis in coronary artery bypass graft operation[J]. *Artif Intell Med*, 2025, 167: 103150. DOI: 10.1016/j.artmed.2025.103150.
 - [27] Prosperi M, Guo Y, Sperrin M, et al. Causal inference and counterfactual prediction in machine learning for actionable healthcare[J]. *Nat Mach Intell*, 2020, 2(7): 369-375. DOI: 10.1038/s42256-020-0197-y.
 - [28] Chang R, Jiao H, Nie W, et al. Organ-Agents: virtual human physiology simulator via LLMs[EB/OL]. (2025-08-20) [2026-07-18]. <https://doi.org/10.48550/arXiv.2508.14357>. DOI:10.48550/arXiv.2508.14357.
 - [29] Björnsson B, Borrebaeck C, Elander N, et al. Digital twins to personalize medicine[J]. *Genome Med*, 2019, 12(1): 4. DOI: 10.1186/s13073-019-0701-3.
 - [30] Corral-Acero J, Margara F, Marciniak M, et al. The 'Digital Twin' to enable the vision of precision cardiology[J]. *Eur Heart J*, 2020, 41(48): 4556-4564. DOI: 10.1093/eurheartj/ehaa159.
 - [31] Li X, Loscalzo J, Mahmud AKMF, et al. Digital twins as global learning health and disease models for preventive and personalized medicine[J]. *Genome Med*, 2025, 17(1): 11. DOI: 10.1186/s13073-025-01435-7.
 - [32] Kelly CJ, Karthikesalingam A, Suleyman M, et al. Key challenges for delivering clinical impact with artificial intelligence[J]. *BMC Med*, 2019, 17(1): 195. DOI: 10.1186/s12916-019-1426-2.
 - [33] van der Vegt AH, Scott IA, Dermawan K, et al. Deployment of machine learning algorithms to predict sepsis: systematic review and application of the SALIENT clinical AI implementation framework[J]. *J Am Med Inform Assoc*, 2023, 30(7):1349-1361. DOI:10.1093/jamia/ocad075.
 - [34] Collins GS, Moons KGM, Dhiman P, et al. TRIPOD+AI statement: updated guidance for reporting clinical prediction models that use regression or machine learning methods[J]. *BMJ*, 2024, 385:e078378. DOI:10.1136/bmj-2023-078378.
 - [35] Liu X, Cruz Rivera S, Moher D, et al. Reporting guidelines for clinical trial reports for interventions involving artificial intelligence: the CONSORT-AI extension[J]. *Lancet Digit Health*, 2020, 2(10): e537-e548. DOI: 10.1016/S2589-7500(20)30218-1.
 - [36] Cruz Rivera S, Liu X, Chan AW, et al. Guidelines for clinical trial protocols for interventions involving artificial intelligence: the SPIRIT-AI extension[J]. *Lancet Digit Health*, 2020, 2(10): e549-e560. DOI:10.1016/S2589-7500(20)30219-3.
 - [37] World Health Organization. Ethics and governance of artificial intelligence for health: WHO guidance[M]. Geneva: World Health Organization, 2021.
 - [38] Char DS, Shah NH, Magnus D. Implementing machine learning in health care - addressing ethical challenges[J]. *N Engl J Med*, 2018, 378(11):981-983. DOI:10.1056/NEJMp1714229.