

重症大数据应用中国专家共识(2022)

中国卫生信息与健康医疗大数据学会重症医学分会 北京肿瘤学会重症医学专业委员会

通信作者: 隆云, 中国医学科学院北京协和医学院北京协和医院 疑难重症及罕见病国家重点实验室 重症医学科, 北京 100730, Email: ly_icu@aliyun.com

【摘要】 随着科技的发展与进步, 重症医学的诊疗融入了越来越多的监测手段与治疗方法。这些多元化的监测数据具有实时性、连续性、动态性、多源性等特点, 当前临床诊疗对其利用程度相对较低。重症患者病情瞬息万变, 对多元数据的及时处理与整合变得至关重要。因此, 近年来重症大数据应运而生。为了更好地推动重症大数据在中国的规范化发展, 在借鉴国内外相关文献和指南的基础上, 中国卫生信息与健康医疗大数据学会重症医学分会组织专家就本共识包括重症大数据的概念、意义与必要性; 重症大数据临床研究关注的临床科学问题; 重症大数据的建立、标准与原则; 重症医学大数据问题解决途径与方法; 重症大数据的临床应用五个方面制定了本共识, 为临床医生及致力于重症大数据的科研工作者提供参考。

【关键词】 重症监护; 重症医学; 大数据

基金项目: 中央高水平医院临床科研专项(2022-PUMCH-B-115, 2022-PUMCH-D-005); 中国医学科学院医学与健康科技创新工程项目(2021-I2M-1-056); 北京自然科学基金(7222134); 首都卫生发展科研专项(2020-2-40111); 2020年北京临床重点专科重症医学科卓越项目(ZK128001); 首都临床诊疗技术研究及转化应用(Z201100005520051)

Chinese expert consensus on the application of intensive care big data (2022)

*Intensive Care Medicine Branch of China Health Information and Health Care Big Data Society;
Intensive Care Medicine Professional Committee of Beijing Oncology Society*

Corresponding author: Long Yun, Department of Critical Care Medicine, Peking Union Medical College Hospital, Chinese Academy of Medical Sciences, Beijing 100730, China, Email: ly_icu@aliyun.com

【Abstract】 With the development and progress of science and technology, the diagnosis and treatment of intensive care medicine has incorporated monitoring tools and treatment methods. These multi-source monitoring data have the characteristics of large-scale, multi-heterogeneous, variably dynamic, high-speed and real-time acquisition, which are relatively underutilized in current clinical treatment. As the state of illness of critically ill patients changes rapidly, the timely processing and integration of multiple data becomes crucial. Therefore, intensive care big data has emerged in recent years. In order to promote the standardized development of intensive care big data in China, based on the relevant literature and guidelines at home and abroad, the Intensive Care Medicine Branch of China Health Information and Health Care Big Data Society organized experts to discuss the concept, significance and necessity of intensive care big data; clinical scientific issues of clinical research on intensive care big data; the establishment, standards and principles of intensive care big database; ways and methods to solve big data problems in intensive care medicine; clinical application of intensive care big data. This consensus was developed for clinicians and researchers working on big data in intensive care.

【Key words】 Intensive care; Intensive care medicine; Big data

DOI: 10.3760/cma.j.cn112137-20221008-02098

收稿日期 2022-10-08 本文编辑 吕相征

引用本文: 中国卫生信息与健康医疗大数据学会重症医学分会, 北京肿瘤学会重症医学专业委员会. 重症大数据应用中国专家共识(2022)[J]. 中华医学杂志, 2023, 103(6): 404-424. DOI: 10.3760/cma.j.cn112137-20221008-02098.



Fund program: National High-Level Hospital Clinical Research Fouding (2022-PUMCH-B-115, 2022-PUMCH-D-005); CAMS Innovation Fund for Medical Sciences (2021-I2M-1-056); Beijing Municipal Natural Science Foundation (7222134); Capital's Funds for Health Improvement and Research (2020-2-40111); Excellence Program of Key Clinical Specialty of Beijing in 2020 for Critical Care Medicine (ZK128001); Beijing Municipal Science and Technology Commission (Z201100005520051)

重症医学发展离不开多元化的监测数据,正是这些数据将重症患者的临床表现具体地呈现出来。这些数据说明了一定的临床现象,而现象背后代表的是疾病发生的本质。重症医学自其诞生之日起,便与数据紧密结合。可以说,数据承载了重症患者临床信息中最有价值的内涵。与其他专业数据相比较,重症医学数据具有实时性、连续性、动态性、多源性等特点,更加能体现重症患者病情的瞬息万变并能够提供治疗的信息。纵观国外重症大数据的发展方兴未艾,为紧跟学术发展前沿,2019年成立了中国卫生信息与健康医疗大数据学会重症医学分会与标准委员会。在过去的几年中学会秉承使命在重症大数据的推广、应用和研究方面发挥了巨大作用。鉴于在既往工作中遇到的问题,为解决相关专业人员对于本领域概念认识的一致性,并加快重症大数据的发展,特制定本共识,以期提高重症大数据领域研究质量并促进未来多学科相互合作。

一、共识的编写方法

本共识由中国卫生信息与健康医疗大数据学会重症医学分会发起和负责制订,由兰州大学健康数据科学研究院、中国医学科学院循证评价与指南研究创新单元、世界卫生组织指南实施与知识转化合作中心提供方法学支持。本共识已在国际实践指南注册平台(Practice Guideline Registration for Transparency, PREPARE, <http://www.guidelines-registry.org>)进行注册,注册号:PREPARE-2022CN566。该共识制订小组由共识专家组、秘书组以及外审专家组构成,共识专家组成员为29人。入选标准为具有丰富重症大数据应用经验的专家。专家具有地域、学科的代表性,专业涵盖重症医学、信息科学、数学、指南方法学等方面,其主要职责是对共识进行整体把握。工作组成员共41人,入选标准为对重症大数据试验有一定初步实践的专家,其主要职责是筛选相关文献并初步凝练推荐意见。秘书组成员为2人,全面负责共识的协调、管理和统稿校审工作。上述成员均由中国卫生信息与健康医疗大数据学会重症医学分会、

标准委员会常委会提名、酝酿,并通过投票选出。本共识纳入来自指南、共识、系统评价、荟萃分析、随机对照试验(randomized controlled trial, RCT)等证据,检索PubMed数据库,以及与重症大数据相关的学会/协会网站文件,检索时间为建库至2022年6月,检索不限定发表语言。最后采用相应的工具评价纳入研究的质量并确定文献证据。共识专家组成员基于指南、共识、系统评价、荟萃分析、RCT的证据,分成不同的领域,根据预先确定的范围,初拟了共识意见和证据与解释。工作组成员整合共识专家组文件,撰写共识全文。秘书组整合不同问题的共识意见,通过线上会议与共识专家组成员开展共23次讨论与修改。2022年7月和9月通过邮件形式在中国卫生信息与健康医疗大数据学会重症医学分会全体委员中进行2轮德尔菲调查,调查总人数为68人,问卷有效回收率为89.7%、92.6%。调查问卷设计、问卷内容均由秘书组成员在中国医学科学院循证评价与指南研究创新单元指导下完成,由专家组成员审核通过并发表。问卷内容主要包括针对每个推荐意见的李克特量表评分及可自由填写的意见和建议区域。针对每一条推荐意见,专家采用李克特量表评分,满分7分,7分表示非常同意,6分表示同意,5分表示一般同意,4分表示不确定,3分表示不太同意,2分表示不同意,1分表示完全不同意。本共识设定:针对单条推荐意见,评分 ≥ 6 分的专家超过70%,则为该条推荐意见达成共识。本共识共凝练出31条拟推荐意见,除缺失值处理及离群值处理两个推荐意见未达成共识外,其余29条推荐意见均达成推荐意见。专家推荐程度以“共识度”标注,共识度=评分 ≥ 6 分的专家/总参评专家人数 $\times 100\%$ 。

二、重症大数据的概念、意义与必要性

推荐意见1:重症大数据概念:重症大数据指重症患者或潜在重症患者在诊疗全过程中形成的大规模、多源异构、可变动态、高速实时获取、低价值密度且难以用传统方法分析的各种指标归纳而形成的有逻辑内涵的数据集合^[1-4]。(共识度97%)

推荐意见2:重症大数据具有多模性、海量性、

动态性、连续性、客观性的特点,其正确获取可为重症预警和临床诊疗提供辅助证据支持。(共识度 98%)

解读与证据:重症病房使用的监测手段丰富,联合使用多种监测设备获得的数据具有多模态特征^[5-7]。可根据需求,在不同层面不同时间进行多参数采样,获得大量连续数据,因此重症大数据具有海量性^[5]、动态性、连续性、准确性的特点^[1,8]。正确有效的数据处理对重症诊疗及护理过程具有指导和预警作用^[4]。最近,基于云的重症监护病房(intensive care unit, ICU)管理系统 Epimed Monitor System[®]部署预测 ICU 住院时间的工具,该系统包含超过 250 万例巴西全国 ICU 入院患者的数据,这些数据可以预测新入院患者长期 ICU 住院风险,有助于规划床位使用^[9]。Komorowski 等^[10]使用强化学习技术指导脓毒症患者液体或血管活性药物使用,外部验证表明该模型比重症监护医生做出了更好的治疗选择。在使用数据挖掘技术建立预测模型方面, Nemat 等^[11]证明“人工智能(artificial intelligence, AI)脓毒症专家”可用于实时数据处理,以预测 4~12 h 内新发脓毒症。尽管大数据研究展现出广阔的应用前景,但在现阶段,临床随机试验数量不多,各种技术模型还需在临床中进行前瞻性测试,以证明其有效性和安全性。鉴于重症患者还存在病情个体化、差异化的特征^[12],现阶段重症大数据还无法为临床提供成熟的指导意见,可作为辅助支持工具使用。

推荐意见 3:中国重症大数据库的建立应遵循多中心、多病种和自动抓取的原则,为大数据的应用和 AI 的开发提供可靠准确的数据支撑。(共识度 92%)

解读与证据:中国重症大数据库的建立处于初步探索中,通过汲取国内外已建数据库经验,总结现有数据库的不足,可以使中国重症大数据库为国内临床决策、精准医疗实施及国内医疗政策制定等提供依据。目前,国外已建立多个重症大数据库,如第四代重症医学监护信息数据库(Medical Information Mart for Intensive Care-IV, MIMIC-IV)^[13]、eICU 协作研究数据库(eICU Collaborative Research Database, eICU-CRD)^[14]、ICU 高时间分辨率数据库(high time resolution ICU dataset, HiRID)^[15]及阿姆斯特丹大学医学中心数据库(Amsterdam University Medical Centers Database, AmsterdamUMCdb)^[16]等。这些数据库主要基于欧

美人种建立,数据体量大,数据类型丰富,但重要生命体征记录定时监测,非全自动抓取,危重症评分系统无数据自动采集及整合功能^[17],普遍缺乏在线初步数据分析软件嵌入,国外重症大数据库在数据实时性和可利用度方面有待提高。Komorowski 等^[10]基于 MIMIC 数据库对大数据进行深度挖掘构建临床模型,并利用 AI、机器学习,发现 AI 医师可以用于辅助临床决策,为危重症患者提供个性化临床最佳治疗决策,有助于改善患者预后。

近年来,国内一直在进行重症大数据库的探索,已经先后建立 ICU 感染患者数据库^[18],儿科重症监护数据库(Pediatric Intensive Care Database, PIC)^[19]及 HeartFailure 数据库^[20]等,但国内现有的重症大数据库起步晚,发展尚不成熟,均为单中心数据库,病种或人群单一,局限于早期数据的初步收集,不具备数据自动抓取与分析功能,数据整体质量偏低且利用效率不高,尚未充分利用大数据挖掘和 AI 技术发挥其效能^[21]。

国外主要重症大数据库的简要信息比较如表 1 所示,可见目前国内外已有数据库均以单中心为主。而各种疾病都有可能发展为重症且需要收入 ICU 进行治疗,提高数据的全面性亦很重要,因而目前这些已有的针对某类疾病建立的单中心或者多中心数据库显然无法满足绝大多数 ICU 患者的需求。目前应用最广的数据库如 MIMIC-III 数据库中生命体征记录是每小时 1 次,但对于需要连续动态监测的重症患者,这种时间分辨率还远远不够。HiRID 具有比其他已发布的数据集更高的时间分辨率,数据储存每两分钟进行一次^[15],但其他数据库都还无法实现。综上所述,中国重症大数据库的建立应遵循多中心、多病种和自动抓取的原则,为 AI 的开发和应用提供数据支撑。

推荐意见 4:构建我国重症患者大数据库用于重症患者病情监测、重症相关临床药物研发及临床试验,以便为重症患者提供规范化、个体化的治疗。(共识度 97%)

解读与证据:理解重症大数据和重症临床的关系至关重要。重症大数据与临床的关系在于:通过数据的集成,给临床医生提供可管理、可解释、可操作、可用于制定治疗方案的数据,给予临床治疗一定参考;通过数据管理构建预测模型和预后模型,给患者提供更优质的个体化和精准化医疗保障;通过机器学习算法,为临床科研工作者提供简便、可靠的模型,为药物开发及探索提供更多支撑,最后

表 1 国外主要重症大数据库简要信息比较

数据库名称	纳入人群来源	国家	年份	患者例数
MIMIC-III	单中心,大样本,麻省理工以色列迪康医学中心	美国	2001—2012	46 520
MIMIC-IV	单中心,大样本,麻省理工以色列迪康医学中心	美国	2008—2019	383 220
eICU	多中心,中小型医院为主,非重症专科医师整理,美国 335 个 ICU 的患者	美国	2014—2015	139 367
HiRID	单中心,瑞士伯尔尼大学医院 ICU 患者	瑞士	2008 年 1 月至 2016 年 6 月	36 098
Amsterdam UMCdb	多中心,欧洲 23 172 个 ICU 患者	欧洲	2013—2016	20 109

注: MIMIC-III、MIMIC-IV 分别为第 3 代、第 4 代重症医学监护信息数据库; eICU 为 eICU 协作研究数据库; HiRID 为 ICU 高时间分辨率数据库; AmsterdamUMCdb 为阿姆斯特丹大学医学中心数据库

共同推进重症医学领域的发展。专家组认为可以通过机器学习等方式建模,对不同的临床问题及时作出预警或适时作出追踪和汇总,从而实现“总结过去经验、警示当下决策、预测未来进展”的目标。

2010 年在美国两家社区医院进行的 RCT 研究指出远程数据算法可以有效地提升重症患者医疗质量^[22],同时 2015 年一篇综述展示了大数据用于重症患者抗生素使用的策略,提出了“自动药代动力学”(AutoKinetics)的概念,为临床用药提供决策支持,并通过与电子病历(electronic medical records, EMR)的直接交互,为大数据的使用开辟了道路,为每例患者在正确的时间提供正确的剂量^[23]。Kindle 等^[24]和 Carra 等^[4]汇总了所有远程算法的开发结果,提示机器学习算法对于脓毒症监测、脓毒症管理、机械通气、减少错误预警和 ICU 预后有着重要意义。此外,重症大数据对于临床护理决策也有着重要意义,2022 年斯坦福医疗团队开发了一种无监督流程挖掘算法用于评估护理质量,最后结果队列的平均符合性得分为 0.36(即路径与所有可能路径的平均重叠率为 36%),最高为 0.64,最低为 0.20,结果说明大数据算法用于 EMR 数据挖掘的可靠性,该方案同时可用于评估其他疾病的护理质量^[25]。2022 年, Boss 等^[26]发布了“ICU Cockpit”——一种算法模式集成平台,实现了重症大数据对病情的预警功能。自 2016 年以来,该平台已经从 200 个信号和实验室分析中处理了来自 979 例患者超过 890 亿条数据,提出了一个基于基础设备的框架,用于部署和验证重症监护算法。它允许算法无缝集成到实时数据流,在临床实践中提供实时临床决策支持和预测。其次是重症大数据对于临床科研的指导。Taglang 和 Jackson^[27]和 Xu 等^[18]系统阐述了大数据探索在临床试验中的重要性。在近两年的大数据探索过程中,多项研究已经开展通过大数据构建个体化计算模型分析,指出重症患者高死亡率的危险因素^[28-30]。对于重症相关

临床药物研发,目前尚缺乏足够基于大数据方法的实践。鉴于近年来大数据在其他专科药物研发中的新进展^[31],推荐在重症医学亦可将大数据结合药物研发。因此可考虑推荐应用重症大数据监测临床病情变化,但需要更多数据库、算法以及大型 RCT 研究共同推进该领域的发展。

三、重症大数据临床研究关注的临床科学问题

推荐意见 5: 建议采用机器学习方法建模,对脓毒症、急性呼吸窘迫综合征(acute respiratory distress syndrome, ARDS)、急性肾损伤(acute kidney injury, AKI)做出早期预警。(共识度 94%)

解读与证据: 有关脓毒症、AKI 及 ARDS 预警模型的研究日益增多,大部分模型能做出早期预警,灵敏度和特异度均表现良好。不同模型的预测能力及可推广性尚需进一步验证。因此,可采用机器学习方法建模,对 ICU 患者脓毒症, AKI 和 ARDS 的发生风险进行早期预测,提升早期应对能力,从而减少其发生概率,改善患者预后。

脓毒症预警模型与人工筛查和评分相比,能早期准确地做出预测,并实现了外部验证。2020 年一项脓毒症预测模型的荟萃分析显示,单个机器学习模型可以早期准确预测脓毒症[受试者工作特征曲线下面积(area under the receiver operating characteristic curve, AUROC): 0.68~0.99],能替代传统评分,但研究间的异质性限制了对结果的评估^[32]。2022 年一项研究用梯度增强监督学习方法开发脓毒症筛查工具,发现在指示时间 1 h 和 24 h,基于梯度增强监督学习方法构建的模型对患者发生脓毒症的预测灵敏度和特异度均高于传统评分方式^[33]。2021 年一项对照研究开发了一种算法,能提前 12 h 准确预测脓毒症(AUROC 为 0.94,灵敏度为 0.87,特异度为 0.87)^[34]。2021 年一项多中心研究显示,使用转移学习算法使脓毒症外部验证数据集的有效性成为可能^[35]。

重症患者 AKI 预警模型能早期准确地做出预

测,但很少同时具有外部验证、临床可解释性和高预测性能的模型^[36]。Xiao 等^[37]研究结果提示,AKI 的早期预警模型 AUROC 为 0.88,能提前 6 h 预测 AKI。2020 年一项多中心研究显示 AKI 预警模型能提前 48 h 预测 AKI,在内部和外部验证中均表现出色(AUROC 分别为 0.86、0.85 和 0.86)^[38]。2020 年一项研究建立的基于 EMR 数据的连续预测模型,能在住院期间实时预测 AKI,性能显著优于单次预测模型(AUROC 分别为 0.724、0.653)^[39]。

ARDS 预警模型能早期高效预测 ARDS,部分模型能实现外部验证,部分纳入了影像学变量。2020 年 Le 等^[40]使用 XGBoost 梯度增强树模型能提前 48 h 准确预测 ARDS (AUROC 为 79.0%)。2020 年 Mayampurath 等^[41]对前瞻性研究数据进行二次分析,使用放射学报告文本建立模型,表现良好并实现了外部验证(C 指数,0.78;95%CI:0.72~0.84)。ARDS 的诊断对影像学依赖性很强,但诊断过程中,影像学信息不一定可及,或其解释存在不确定性,这种信息称为特权信息和不确定性标签,但纳入影像学变量的模型更接近临床。2022 年, Sjoding 等^[42]比较了基于 X 光的机器学习识别 ARDS 方法与基于病历识别 ARDS 的方法。发现 AUROC 分别可达到 0.83、0.80,而将两者结合后 AUROC 可达 0.88。

推荐意见 6: 基于机器学习预测模型能够对 ICU 内潜在器官功能损伤高危患者进行有效的风险预测。(共识度 89%)

解读与证据: 早期预警评分系统提出后使得医护人员能够更好地识别潜在的危重症患者,实现早识别、早干预,从而改善患者预后。但这一评分系统在患者发生明显的恶化前可能无法进行识别。2019 年的一项系统评价发现采用统计建模的早期预警评分相较于加权的早期预警能够更准确地识别高危患者(AUROC 分别为 0.80、0.73),其发现 1 例真正的阳性病例分别需 4.9 次与 7.1 次报警事件^[43]。与此类似的 2021 年的一项系统评价亦显示,基于机器学习的临床状况恶化预警系统能够更准确地预测在 ICU 环境中患者生存恶化的风险,该模型的 AUROC 为 0.57~0.97^[44]。

危重症患者除原发病进展外,可能出现多种威胁患者生命的合并症、并发症。(1)循环功能紊乱:2020 年有研究建立了一个循环衰竭事件预警系统,能提前 2 h 以上识别有循环衰竭风险的患者,并成功在患者队列中进行了外部验证^[15]。校正 QT

间期(corrected QT interval, QTc)延长与尖端扭转型室性心动过速有关,Tisdale 等^[45]的研究提示通过使用数据挖掘、建模、机器学习等技术开发了风险量化方法,对 QTc 延长进行预测。该 QTc 风险评分在预测性能上表现良好,有较好的灵敏度(高风险为 74%,中风险为 67%)、特异度(高风险为 77%,中风险为 88%)、阳性预测值(高风险为 79%,中风险为 55%)和阴性预测值(高风险为 76%,中风险为 88%)。(2)水电解质的紊乱:西班牙 Broch Porcar 等^[46]通过回顾重症监护患者低钠血症的管理,开发了一种完善标准化的低钠血症诊断及治疗算法。(3)消化道出血:Levi 等^[47]基于重症监护信息系统的机器学习分析算法与协作数据分析库分析 5 h 后患者发生消化道出血风险,模型 AUROC>0.80,有较好的预测性。(4)血栓形成:深静脉血栓形成(deep vein thrombosis, DVT),具有较高的发病率和死亡率,并导致医疗费用的增加。Ryan 等^[48]开发了梯度增强的机器学习算法来预测患者在发病前 12 h 和 24 h 出现 DVT 的风险,机器学习预测因子获得的诊断院内 DVT 的 AUROC 分别为 0.83 和 0.85,早期康复训练的开展及防范压力性损伤也都对 ICU 内预后产生影响。

推荐意见 7: 建议采用机器学习的方法建模对住院患者进行早期筛查,为临床医师早期干预、减少疾病的重症化提供帮助。(共识度 88%)

解读与证据: 与普通患者相比,重症患者往往预示着住院时间较长、花费较多、预后较差,早期发现患者病情变化趋势、及时干预对于阻止疾病的进展具有良好的意义。机器学习的方法可有利于疾病的早期筛查,并可促进对疾病的及时处置。但对于不同研究对象在模型应用前应注意异质性的校正,否则容易导致错误的临床指导。建议采用机器学习的方法建模对重症患者进行早期筛查,为临床医师早期干预、减少疾病的重症化提供帮助。

2020 年发表的一项研究使用 5 折交叉验证对几种机器学习方法进行了评估,并且应用 XGBoost 算法构建脓毒症 AI 预测模型,验证结果显示其准确度为 82%±1%;灵敏度为 65%±5%;特异度为 88%±2%。XGBoost 算法及序贯器官衰竭评分(sequential organ failure assessment, SOFA)预测脓毒症的 AUROC 分别为 0.890 和 0.596。可见,基于机器学习算法,可以帮助临床医生更早地部署适当的治疗方案,对 AI 算法的早期和精确响应将降低成本,改善疗效,使医疗系统、医务人员和患者受

益^[49]。2020 年一项多中心真实世界数据研究证实,在临床环境中应用预警模型后,脓毒症患者住院病死率平均下降 39.5%,住院时间减少 32.3%,脓毒症相关住院 30 d 再入院率下降 22.7%^[50]。除脓毒症外,机器学习方法也已被用于其他重症疾病的早期筛查中,2021 发表的一项研究利用四种机器学习方法构建的模型(随机森林、XGBoost、GLMBoost 和 Lasso-GLM)预测小儿多脏器功能障碍,结果显示所有方法的早期预测模型在 AUROC 都达到了 0.91,通过风险为基础的患者监测进行早期预测可以为小儿多脏器功能障碍发生提供超过 22 h 的提前期,这将为患者预后的改善起到重要作用^[51]。2021 年 Romero-Brufau 等^[52]开发了一项能够在 24 h 内预测患者病情恶化的预警模型(复苏呼叫、需转运至重症监护病房、eCART 呼叫),灵敏度为 73% 时,梅奥早期预警评分(MC-EWS)每天每 10 例患者产生 0.7 次警报,比国家早期预警分数低 45%。在内部和外部验证数据集中都能准确预测(C 指数分别为 0.913 和 0.937),实现高精度识别高危患者。但是亦有研究提示临床医师在预备应用相关模型时应首先根据患者异质性对模型进行校正而后再行应用,避免造成误判影响临床决策^[32]。然而,临床工作中应首先充分提高临床一线工作人员对于已有评分系统的重视程度并监督实际应用情况,否则仅单纯地从模型性能角度进行提升而不提高临床应用及响应速度将是徒劳^[53]。

推荐意见 8: 推荐将重症患者的影像数据纳入重症数据库,以提供更加全面、准确和及时的诊断信息,从而通过相关算法指导临床决策。(共识度 92%)

解读与证据: 早已有研究将 AI 用于肿瘤、感染灶影像的筛查和诊断,并证实其在快速处理大量影像资料、前移诊断“关口”、避免漏诊误诊等方面的优势^[54-55]。重症患者疾病状态和影像表现更加复杂多样,使用 AI 进行影像诊断的最佳时机和场景有待更多验证。专家组认为 AI 辅助 ICU 患者影像诊断具有良好的应用前景,推荐进行相关探索,以提高诊断效率和准确性,并为临床决策提供参考。

一项研究对密歇根大学医院 500 例 ICU 患者的 3 078 次胸片影像采用方向性测量和深度机器学习特征进行建模诊断 ARDS,最终可达到 83% 的准确度和 0.79 的 AUROC 值^[56]。神经网络模型分析卧位胸片诊断肺炎和胸腔积液的 AUROC 值分别为 0.737 和 0.740,与影像专家诊断结果相似(AUROC

分别为 0.779 和 0.698)^[57]。在新型冠状病毒肺炎暴发流行中,AI 辅助影像诊断表现优异,各类机器学习方法不仅能从海量的肺部 CT 影像中快速鉴别出新型冠状病毒肺炎的 CT 影像(AUROC 为 0.951~0.980),还能对患者的重症转化做出预测(AUROC 为 0.848)^[58]。经典图像处理 and 深度学习相结合的机器学习法分析 110 例硬膜下血肿重症患者的 CT 影像,显示样本查全率、查准率分别为 78.61%、76.12%,基于血肿体积的严重度判断特异度为 92.31%,可帮助医生减少决策时间^[59]。

除放射影像外,AI 在其他 ICU 床旁影像诊断中也有应用研究。2019 年的 1 项研究显示神经网络模型检测床旁肺部超声 B 线的灵敏度和特异度可达 0.871 和 0.930^[60]。一项研究应用机器学习的方法分析 136 例插管患者 SBT 试验前和结束前的床旁肺部、心脏和膈肌超声数据,发现其判断拔管后发生呼吸窘迫的 AUROC 值分别为 0.972 和 0.920^[61]。2021 年的 2 项研究表明神经网络模型利用超声影像早期诊断脓毒症患者发生 AKI 的准确度和灵敏度均高于专业影像医师^[62-63]。电阻抗断层扫描(electrical impedance tomography, EIT)仅能粗略显示肺部各区域通气和血流分布,无法量化成床边监测指标。通过深度学习训练的神经网络模型可以从 EIT 信号中推算出肺容积、气流速、标准化气道压甚至跨肺压等信息^[64]。

推荐意见 9: 推荐通过聚类分析等方式,将脓毒症、ARDS、AKI 患者分成临床结局及治疗反应不同的表型,识别最可能受益于特定治疗策略的患者。(共识度 91%)

解读与证据: 聚类分析可以在异质性人群中识别相对同质的群体。某些治疗手段仅针对特定人群有效。利用聚类技术将重症患者分为不同表型,其在合并症、实验室指标、生命体征、临床结局、治疗反应性等方面具有显著差异,识别出对特定疗法受益的人群。目前表型的识别在脓毒症、AKI、ARDS 中已有研究进展,但表型的准确性及可推广性仍需进一步的验证。专家组推荐通过聚类分析等方式,将重症患者分成不同表型,识别最可能受益于特定治疗策略的患者。

使用临床和(或)宿主反应数据和机器学习(如潜类别分析和 K 均值聚类),将脓毒症、AKI、ARDS 等重症患者细分为不同的表型^[65-68]。2021 年一项 RCT 研究通过 K 均值聚类确定了 4 种基于凝血的脓毒症表型,并使用机器学习方法确定了哪种表型

将受益于重组人血栓调节蛋白^[69];另一项 RCT 研究通过聚类分析确定了 4 种脓毒症临床表型。这些表型在人口学特征、实验室异常、器官功能障碍模式方面存在差异,并且与传统的患者分组(如感染部位、器官功能障碍模式或疾病严重程度)不具有同源性^[70];2020 年一项对 AKI 队列进行的潜在类别分析,确定了具有不同临床结局的脓毒症 AKI 的两种表型^[71];2017 年一项前瞻性观察性队列研究通过无监督的共识聚类和机器学习对全血 RNA 表达谱进行分析,确定了 4 种脓毒症内表型(Mars 1~4),其中 Mars 1 与 28 d 死亡率相关。为了便于临床应用,该研究还为每种表型提取了可进行准确分类的生物标志物^[72]。有研究利用 ARDS RCT 数据,通过潜类别分析(latent class analysis, LCA)方法确定了两种不同的 ARDS 表型。这些表型具有不同的临床结局,并且已确定对呼气末正压策略^[73]、液体治疗^[74]和辛伐他汀^[75]的不同治疗反应。

推荐意见 10:在特定临床场景下,如拔除气管插管决策、重症药物治疗决策等,建议基于机器学习算法构建可用于临床治疗的决策模型。(共识度 74%)

解读与证据:ICU 患者病情通常疑难且危重,EMR 系统、监护仪、呼吸机等仪器设备均能生成海量生命信息数据,已远超出 ICU 医生持续处理和正确解读的能力,影响临床决策的效力和响应速度。AI 模型可以连续地对大量数据进行清晰、归类、分类、计算和关联,对病情变化进行预测,从而辅助临床决策、提升重症救治质量和效率。

目前多项研究评估了应用机器学习等 AI 技术进行治疗决策后对临床的影响。将强化学习应用于脓毒症人群,AI 临床医生可以优化液体管理和血管活性药物治疗,降低病死率^[10, 76]。2019 年一项研究建立 XGBoost 模型预测 AKI 患者尿量,与传统的 logistic 回归模型相比,该模型能够更好地区分患者是否具有容量反应性。

AI 技术已被尝试应用于拔管决策和优化重症患者药物治疗等临床情况。2018 年一项回顾性研究使用机器学习来识别需要延长机械通气时间和气管切开高风险的患者^[77]。2021 年 Fabregat 等^[78]比较了 3 种分类学习方法[logistic 回归模型、XGBoost 和支持向量机(support vector machines, SVM)]预测拔管失败,目前可以降低 9% 拔管失败率。2021 年另一项研究通过机器学习算法建立了意外拔管的预测模型,其中随机森林算法得到最佳

AUROC 为 0.787^[79]。

应用机器学习优化重症患者抗凝、抗感染及镇静等治疗效果的研究尚处于探索阶段。Chen 等^[80]、Su 等^[81]、Li 等^[82]对比了不同机器学习方法预测抗凝药物(枸橼酸盐、肝素)的治疗效果,结果均显示浅神经网络模型准确率和 F1 得分总体优于其他模型。2022 年一项单中心回顾性研究使用机器学习和聚类分析的方法为重症患者的抗生素管理提供指导意见^[83]。2022 年另一项研究基于自注意力和残差结构卷积神经网络(convolutional neural network, CNN)对麻醉深度监测有良好预测效果^[84]。以上案例说明 AI 在指导危重患者的重要决策方面可能发挥作用。但绝大多数已开发的 ICU-AI 模型仍处于测试或原型设计阶段,只有少数在临床治疗实践中得到了实际评估。van de Sande 等^[85]发现目前尚无任何研究报道在常规重症诊疗过程中集成 AI 模型的结果。当前,AI 用于指导临床决策的研究结果,大多是根据回顾性和观察性数据集计算所得。因此,为实现 AI 直接指导临床决策,必须对源自此类 AI 系统的建议序列或策略进行全面分析,尚需更多高质量前瞻性研究。

推荐意见 11:建议采用机器学习方法,对重症患者的预后情况作出预测。(共识度 85%)

解读与证据:目前有关 ICU 重症患者死亡率的预测模型日渐增多,许多数据模型对于疾病的预测效能优于临床评分系统,部分预测模型仍依赖临床评分系统辅助,但目前重症医学中的 AI 模型主要是使用回顾性数据生成的,样本量小且结论可重复性低,缺乏足够的外部验证或前瞻性评估。

机器学习模型、算法多种多样,如 SVM、梯度提升决策树(gradient boosting decision tree, GBDT)、邻近算法(algorithm, KNN)、随机森林模型(random forest, RF)。Barchitta 等^[86]发现,相较于简化急性生理评分 II (Simplified Acute Physiology Scores II, SAPS II),SVM 模型可以用于识别人 ICU 7 d 内死亡风险较高的患者。进一步根据不同病种比较机器学习方法和传统评分系统的预测效能,结果如下:(1)脓毒症:2021 年的研究表明,GBDT 预测脓症患者死亡的准确率优于其他模型(LR、KNN、RF、和 SVM)^[87]。García-Gallo 等^[88]使用随机梯度提升(stochastic gradient boosting, SGB)等组装算法生成的脓毒症模型比 SAPS II、SOFA 或牛津急性疾病严重程度评分(Oxford Acute Severity of Illness Score, OASIS)等传统评分系统,



更能准确预测 1 年死亡率。(2)脑出血(intracerebral hemorrhage, ICH):Nie 等^[89]研究指出 RF 是预测接受 ICU 治疗的 ICH 患者死亡率的最佳模型,与 APACHE- II 评分相比,所有用于预测的机器学习算法均能更好地预测 ICU 死亡率。(3)重症急性胰腺炎(severe acute pancreatitis, SAP):Halonen 等^[90]的研究建立预测 SAP 严重程度的人工神经网络(artificial neural network, ANN)模型,结果优于 Rason 评分、Glasgow-imrie、APACHE- II 及 SOFA 评分。Ding^[91]等研究表明 ANN 模型能快速准确地急性胰腺炎早期阶段筛选出具有死亡风险高的患者。

推荐意见 12: 临床决策支持系统(clinical decision support system, CDSS)可用于改善重症患者诊治指南和临床路径实施的依从性。(共识度 86%)

解读与证据:以证据为基础的临床诊疗指南为危重患者的救治提供了规范化、同质化的诊疗策略。然而,在 ICU 日常诊疗中,依从临床指南进行临床诊疗的比例不高,导致原本可避免的患者死亡增加^[92-93]。CDSS 是一种帮助卫生保健工作者做决定的计算机程序。随着 CDSS 应用于临床,绝大部分研究显示 CDSS 可以辅助 ICU 医生决策并提高对诊疗指南的依从性,改善患者预后。但 CDSS 种类繁多,一个 CDSS 针对某一种疾病,开发成本高,以大数据为基础的 CDSS 已开始应用到临床辅助决策,但尚未用于改变指南依从性。不仅如此,CDSS 需与患者电子健康档案系统整合,由于不同地区或医院采用的电子健康档案系统不同,CDSS 在不同医院的推广和应用受限。因此,CDSS 可用于改善危重患者诊治指南依从性,但以大数据为基础的 CDSS 尚处于研发阶段,建议有条件的医院考虑以大数据为基础的 CDSS 的研发和临床应用,以提高指南依从性。

早在 2011 年,CDSS 如“流程图”,可在床旁实时监测患者各项参数,早期筛查出脓毒症患者并根据 SSC 指南提出一系列指令性治疗措施^[94]。CDSS 的应用可以显著提高 SSC 指南复苏集束化策略的依从性,缩短抗生素使用时间,降低医院病死率^[95]。小潮气量肺保护性通气的临床实施通过采用 CDSS 指导医务人员进行呼吸机模式、支持水平的设置,肺保护性通气的依从性提高,停用 CDSS 后,潮气量的水平显著增加^[96]。一项关于谵妄管理的研究发现在采纳个体化预防 ICU 谵妄 CDSS 系统后,谵妄

发生的持续时间明显缩短,昏迷时间减少,脑功能改善^[97]。在另一项评估 AKI 指南依从性的前瞻性观察性研究中,将 AKI 的 CDSS 整合入 ICU 的重症信息系统,显著降低从 1 期 AKI 恶化的患者比例、依诺肝素剂量使用不当的比例、AKI 患病率显著降低^[98],但目前尚无基于大数据的重症 CDSS 应用提高指南依从性的报道。有待今后进一步研究证实。

四、重症大数据库的建立、标准与原则

推荐意见 13: 建议构建重症医学数据库与数据分析平台。(共识度 98%)

解读与证据:重症医学数据库能够为临床医学研究提供良好的数据基础及新的思路,进而提高对疾病的认识。例如,Sepsis 1.0 将脓毒症定义为感染导致的全身炎症反应综合征(systemic inflammatory response syndrome, SIRS)^[99],虽然在 Sepsis 2.0 中各种诊断指标更加完善^[100],但仍延续了 Sepsis 1.0 的标准。但依据感染和 SIRS 的诊断标准并不能准确描述患者的疾病特点,如原发病不同,患者的症状和死亡率等也不相同。在 2016 年,以大数据分析为主要依据的 Sepsis 3.0 将脓毒症定义为机体因感染的失控反应所导致可以威胁生命的器官衰竭,即感染和器官功能诊断模式,使脓毒症的定义更适应于病理生理学,并且在临床实践中更容易实现^[101]。可以说,国外已经构建的 MIMIC 和 eICU 协作研究数据库^[13],在临床诊疗中发挥的作用逐渐凸显。目前国内也加快了建立重症大数据平台的脚步,但多是局限于各个医院的单独数据库,在数据交流和影响力方面还存在一定不足,因此,我们推荐构建属于中国人的重症医学数据库及数据分析平台,以加强学科建设,进而提高对重症患者的救治水平。

推荐意见 14: 推荐形成标准规范的重症数据集。(共识度 97%)

解读与证据:标准规范化的数据集是大数据应用的基础,并且有利于不同区域研究中心的合作。ICU 设备仪器获取信息众多,合理地纳入数据并做好数据分类,能够更加充分便利地利用数据。目前,国内外已有多个重症医学大数据信息系统,该类信息系统将临床数据以特定的分类标准分为不同数据元,之后使用特定的数据采集方式获取并分析数据。参考国家 EMR 基本架构与数据标准、北京市地方标准(DB11/T 1866-2021,《重症医学数据集患者数据》)以及在医学研究领域广泛应用的重症医学数据库,推荐标准数据集需包含以下数据

集:(1)患者基本信息数据;(2)患者诊断信息数据;(3)患者监护数据;(4)患者药物使用数据;(5)患者实验室信息数据;(6)患者出入量数据;(7)患者影像数据;(8)患者病原学数据。详见表2。

同样推荐可结合实际情况如医院病种情况、信息中心、实验室检测项目等客观条件进行调整,例如基于ARDS、脓毒症、AKI等重症医学常见疾病构建专病数据库。构建专病数据库,需提高和丰富专病相关的重症信息采样频率及类别,如ARDS数据库需要对呼吸机参数、体外膜肺氧合参数、肺部超声参数、肺部影像学参数、相关生物标记物等进行进一步采集;脓毒症数据库需要对血管活性药物、病原学采集、器官功能评估等进行进一步采集。

推荐意见 15: 建议对于客观数据首先选择进行自动采集。对于暂时不可自动采集的数据应结合研究需求、数据来源及数据类型进行有针对性的采集。(共识度 92%)

解读与证据: 数据采集过程需遵循全面、多维、高效、及时的原则进行,鉴于ICU内数据来源多、数据结构丰富,为避免人为误差影响后续数据使用,推荐采用自动化数据采集技术实现数据采集过程。

ICU 中的数据可大致分为“表型数据 phenotypic data”和“生理数据 physiologic data”。表型数据包括人口统计、年龄、性别、实验室值以及医生和护理记录。表型数据可以从EMR中查询和提取。可以通过爬虫或应用程序接口等方式获取相关内容,并从中抽取出所需要的属性内容。生理数据包括生命体征(血压、心率、呼吸频率、核心温度)和床边监测设备生成的其他参数(颅内压、脑电图)。若能经各个软件厂商获取设备的数据接口,可通过接口对接方式实现数据采集汇聚。若部分设备不能获取数据接口,可尝试通过获取系统的底层数据交换、客户端和数据库之间的网络流量包,

利用底层输入与输出请求与网络分析等技术,采集目标设备产生的所有数据,将数据转换与重新结构化,输出到新的数据库。

ICU 中的警报,例如心电图(electrocardiogram, ECG)导联、血压袖带从患者身上脱落,输液泵输液完成或管道中出现气泡,机械通气呼吸机出现高气道压力、漏气或呼吸暂停等,也可以归类到生理数据类。这部分数据可尝试通过各种设备的日志源上收集日志的方法来进行采集。连续波形数据由于其连续和高采样率的特点,采集过程则更为复杂。近年来,多项研究采用时序数据库和非结构化数据库,如时间序列数据平台(InfluxDB)、开发者数据平台(MongoDB)等探索了多种连续时序信号的写入、存储、查询等流程,在一定程度上可以解决连续波形数据的存储-传输-交换-利用问题。针对图像数据,鉴于目前大部分图像都存储在影像归档和通信系统(picture archiving and communication system, PACS)系统中,需要明确是从设备上(CT机、超声机等)采集还是通过PACS对接端口的方式采集。

推荐意见 16: 建议建立重症大数据标准体系,规范多中心来源数据,规范约束标准代码、度量单位、字段标准、命名词典,以保障重症大数据的数据使用的同质性、规范性。(共识度 95%)

解读与证据: “信息整合,标准先行”^[102],重症大数据的建设必须遵循相应的规范标准来实施,通过标准规范约束标准代码、度量单位、字段标准、命名词典,来保障后续建模与应用过程中数据处理的一致性,从源头上保障数据的标准化生产,为重症大数据的建设、数据集成、数据交换以及数据共享打下基础。重症大数据为隐私性高和来源多样的多模态数据,具有数据维度多、时效性好、价值密度大和数据质量高的特点。ICU 中的“表型数

表2 标准规范的重症数据集

项目	内容
患者基本信息数据	包含患者入院和出院的时间信息、人口统计信息、入院来源、ICU类别、死亡时间等
患者诊断信息数据	包含患者入住ICU期间所有疾病诊断信息,主要诊断需与次要诊断进行区别
患者监护数据	包含患者入住ICU期间常规生命体征、呼吸机参数信息、血液净化参数信息、主动脉球囊反搏参数信息、患者精神状态、患者评分信息等
患者治疗数据	包含患者入住ICU期间所有药物的给药途径、使用时间以及药物剂量;手术名称、时间及相关信息等;治疗操作名称、时间及相关信息等
患者实验室信息数据	包含患者入住ICU期间的实验室检查信息,包括但不限于血常规、肝肾功能、凝血、心肌标记物、C反应蛋白、血气分析等血压检查,尿常规、粪便常规等常规尿、便检查。如采样时间、标本类型、检测项目、检测结果、正常值参考范围等
患者液体出入量数据	包含患者入住ICU期间所有进入体内和经体内排出的液体数据,包括液体类型、出入途径、时间等
患者影像数据	包含患者入住ICU期间放射成像相关的文字报告
患者病原学数据	包含患者入住ICU期间留取的病原学数据,包括采样时间、标本类型、病原学名称、病原学药敏等



据”和“生理数据”可从类型上分为结构化的离散数据、时序数据以及非结构化的文本数据、图像数据和音视频数据。主要内容如下:(1)离散数据:患者的基本信息和常规体征数据,包括性别、年龄、血型、身高、体重等一系列离散数据,主要为表征数据。这些数据体量较小,并且稳定。(2)时序数据:主要为生理数据,包括血氧、心率、心电等各种生命体征参数的时间序列数据。这些数据与患者的实时症状关系紧密,实时性高,连续性强,数据集大,也是重症大数据区别于其他健康医疗数据的最大特点之一。(3)图像数据:主要为生理数据,包括超声、放射等大量的影像数据。这些图像数据体量大,是诊断和操作的重要参考数据。(4)文本数据:大量关于患者病历和诊断结果的文本数据,主要是为表征数据,包括 EMR、手术记录、检查报告等。在重症的所有数据类型中,时序数据、图像数据和文本数据的信息价值密度较大,对临床诊疗和决策起着重要作用。

由于各个中心信息化水平参差不齐,涉及面广,上述重症数据格式繁杂,难以融合。多个中心经协商后制定重症大数据库统一的数据字段、内容和格式,建立重症大数据标准体系。例如,图像数据的标准化,不同医院影像科水平参差不一,需要多中心对压缩原图上传的图像质量标准进行协商。针对不同类型的数据,为保障重症大数据库的标准化,可制定不同类型数据的治理规则,在数据入库时系统自动进行清洗,必要时辅以人工审核,以保障数据质量。针对入库数据质量评估,可从规范性(数据符合数据标准、数据模型、业务规则、元数据或权威参考数据的程度)、完整性(数据元素按照数据规则要求被赋予数值的程度)、准确性(数据准确表示其所描述的真实实体“实际对象”真实值的程度)、一致性(数据与其他特定上下文中使用的数据无矛盾的程度)、时效性(数据在时间变化中的正确程度)、可访问性(数据能被访问的程度)6个方面来管理和评价。

共识意见 17: 推荐建立数据安全机制,保障数据的存储、处理、共享和使用安全。(共识度 98%)

解读与证据: 我国信息安全保障体系主要包括风险评估与等级保护、监控系统、密码技术与网络信任体系、应急机制、灾备五项技术工作。信息系统安全保护等级分为五级,一级至五级等级逐级增高。围绕《中华人民共和国网络安全法》、《中华人民共和国数据安全法》和《中华人民共和国个人信

息保护法》,国家开展数据分类分级制度的建设,数据安全必须同时注重数据本身的安全(采用现代密码算法对数据进行主动保护)和数据防护的安全(采用现代信息存储手段对数据进行主动防护)。大数据环境下新的安全问题需要解决,其中包括平衡隐私与实用性,对加密数据开展分析和治理,以及核查认证用户和匿名用户。随着重症大数据的应用范围不断扩大,重症大数据内容越来越丰富、价值越来越宝贵,含有大量个人信息敏感数据,应当建立重症大数据的安全体系 and 安全管理责任制度,保障重症大数据在数据存储、数据开放、数据处理过程中的安全。

在数据存储时,应做好分布式环境下(如 Apache Hadoop)系统安全加固及软硬件架构设计,静态数据要设置严谨的细粒度访问权限控制和风险登记管理策略,涉及隐私的数据存储要实现分级隔离、数据加密(如 AES、RSA 和 SHA-256 等加密方法)等安全技术手段,动态数据通过加密和动态审计能力,对重要敏感数据进行分级、标识,使用传输层安全技术进行集群节点间的通讯在传输过程中维持保密性,实现跨平台(端点、移动设备、网络和存储系统)的统一管理。

在数据处理时,应根据数据库体量及访问方式,尤其是面向多中心,设计软件架构及网络配置,根据软件架构设计适宜的硬件架构,同时做好网络安全等策略配置,以保障数据安全。在数据被授权其他方处理后,最重要的问题是处理过程中是否产生滥用和恶意还原敏感数据,是否符合法律法规,是否符合双方或各方同意的隐私条款。在多方计算中,通过数据脱敏、联邦学习^[103-104]等系统策略设计来避免数据外泄。

在数据共享时,应做好数据脱敏、权限管理、日志审计等措施保障数据安全,数据不能无条件向公众或者第三方开放,要考虑单个信息脱敏后经多源碰撞产生敏感信息易还原的安全风险,因此只能做点对点的共享,或者基于某种特殊约束的多边交易,例如共享健康档案、患者用药信息、医疗影像等重症大数据信息。数据共享的正当与否要综合权衡该数据的使用场合及数据主体的知情权。

五、重症医学大数据问题解决途径与方法

(一)数据类型及处理原则

推荐意见 18: 推荐使用滤波器等数字信号处理方法对时序数据进行预处理,使用深度学习的方式处理图像类数据,使用自然语言处理(natural

language processing, NLP) 技术处理非结构化的文本数据。(共识度 93%)

解读与证据:从机器模型构建的角度,重症数据可大致分为数值型时序数据、数值型非时序数据、文本数据、图像数据四类。其中数值型数据可根据采集密度分为两类:(1)时序数据,或称为“流数据”(streaming data),包含如 ECG, 动脉与颅内压,血流动力学监测,呼吸机数据,脑电波等采集频率相对较高的数据^[4];(2)非时序数据,或称为“稀疏数据”(sparse data),包含如血气分析,实验室检查结果,病史等采集频率相对较低的数据。结合不同类别的数据,可以提升 AI 预测模型的准确性^[105],可以在复杂与不确定的诊断条件下提供决策支持^[106],也可更好地适配临床的实时数据环境。

针对时序数据,在进一步通过不同算法进行模式识别或其他处理之前,通常会使用滤波器等数字信号处理方法进行预处理,其主要目的是通过各种数学方法,把信号中不同频率的分量剥离出来,进行针对性的处理。例如在 ECG 数据处理中,常使用五分钟移动平均值进行低通高通滤波^[107], Narula 等^[108]在构建脑电信号模型时,也通过带通滤波器来移除基线漂移与高频干扰。

针对非时序数据,处理技巧主要体现在根据不同的算法模型,解决数据(参数)异常值缺失值、筛选降维等问题。当数据完成对应的预处理后,不管是简单朴素的如线性回归、逻辑回归算法,还是精密复杂的如提升算法、增强学习等^[10, 109],都可以在对应场景中取得很好的效果,故不做特殊推荐。

针对图像数据,例如 CT、病理切片、超声图像等,大多采用深度学习(如 CNN)的方式进行处理,挖掘图像中的信息,完成诸如 CT 图像分类,细胞核识别、标注、分割等任务^[110-113]。特别的, Walsh 等^[114]认为,深度学习的方法能够直接从图像中提取到重要特征,有助于产生新型的生物标志物和更准确的图像分析工具。

对于非结构化的文本数据,例如 EMR 中的叙述性文字,以及放射学、病理学报告等,通过 NLP 技术,能够对其内容进行挖掘与处理,以得到诸如病理信息、社会环境信息等内容,与现有的专家知识库(如统一医学语言系统 UMLS 等)结合作为补充,可以提升相关预测模型的准确率,且展现出超过人工处理的速度与精度^[115-117]。特别的,针对中文的 NLP, ICTCLAS 系统, THULAC 工具包等,都是不错的辅助工具,但 UMLS 等知识库的中文化(或者其

他中文医学知识库)有待文献论证。

(二)数据预处理

推荐意见 19:建议采用重采样方法处理非平衡数据集。(共识度 78%)

解读与证据:在重症医学数据集中,非平衡数据十分常见,非平衡数据是指在分类任务中各类别间的样本数量分布不均匀,会存在差距特别大的情况,这会极大地影响预测模型的最终性能。例如,在重症医学数据集中少数的死亡样本会携带关于死亡率预测的重要信息,但由于模型对于数据不平衡现象不敏感而被忽视。针对数据不平衡现象,专家组推荐采用重采样方法对不平衡数据进行处理,主要分为欠采样、过采样和合成过采样技术 3 种。欠采样就是从大多数类随机抽取较少的样本,使数据趋于平衡。编辑最近邻(edited nearest neighbors, ENN)是欠采样方法中最典型的方法。过采样就是按照样本标签较少的样本规则生成更多的标签样本,使数据趋于平衡。合成少数类过采样技术(synthetic minority over-sampling technique, SMOTE)是一种过采样技术,可以为少数类生成合成样本。为了减少过采样和欠采样带来的拟合问题,在此基础上,又引申出了过采样和欠采样相结合的方法用于处理数据不平衡。Tiwari 等^[118]在使用机器学习对房颤进行预测的研究中,使用多种采样方法来处理对照组数据远多于实验组数据的不平衡问题,并对不同采样方式下的数据进行了建模比较,最终根据分类器效果择优选择了随机过采样方式。Papp 等^[119]应对类别不平衡问题,使用 SMOTE 采样方式对少数类的样本进行合成,并通过交叉验证和混淆矩阵对合成的新数据结果进行了分析。

推荐意见 20:建议通过独热编码、顺序编码等方式将原始类别变量和数值变量转化为机器学习算法能够直接处理的变量。(共识度 83%)

解读与证据:变量类别转化的作用是将包含上述信息的原始重症医学数据类别转换为适于数据挖掘、便于模型理解的形式。变量类别转化通过编码等操作使得原始数据变得更加整洁和一致,建议采用独热编码、顺序编码等方法实现。独热编码是无序类别变量的常见数值化处理方式,用“1”表示属于该类别,用“0”表示不属于该类别,独热编码会在原始变量中增加新的变量,新增变量的数量即为类别的数量。顺序编码是有序类别变量的常见数值化处理方式,这种编码根据有序变量所代表的不同

同程度进行有序的数字编码,例如从 0 到 5 对患者的健康状态进行评分等。

推荐意见 21: 建议使用主成分分析法 (principal component analysis, PCA) 等降维方法对重症医学数据集中的高维特征进行变量筛选。(共识度 90%)

解读与证据: 在大多数重症医学大数据研究问题中,采用的数据集通常具有高维特征变量,容易导致过拟合的问题,同时增加训练成本。因此需要通过变量筛选来提取重要特征,达到数据降维的目的。专家建议采用 PCA、方差选择法、单变量特征选择、正则化模型、基于机器学习模型的特征排序、递归特征消除方法。

PCA 是比较受欢迎的一种通用特征降维方法,可以用于对数值、文本以及图像等多种类型数据进行降维,本质上是将多个变量合成为少数几个相互独立的成分,且每个成分都可以反映原始变量的信息,可以提升学习速度并且降低训练成本。方差选择法是一种简单的特征选择方法,通过移除低方差的特征对特征进行筛选。单变量特征选择通常采用 χ^2 检验、 F 检验等统计检验方法或皮尔逊相关系数、距离相关系数等衡量标准判断变量之间的关系。正则化模型主要分为 L1 正则化及 L2 正则化,通过把额外的约束或者惩罚项加到已有模型的损失函数上,以防止过拟合并提高模型的泛化能力。L2 正则化相对 L1 正则化更加稳定,并且对特征的理解更加有利,正则化模型在医学数据的特征选择中经常用到, Liang 等^[120] 在对新型冠状病毒肺炎危重症患者进行早期分诊的研究中,通过 Lasso 方法选择了十个具有统计学意义的变量作为预测因子。很多机器学习方法都可以实现对特征进行打分,例如通过衡量特征重要性得到特征排序。因此建议利用选择的机器学习模型完成特征选择,包括 SVM、随机森林、决策树、XGBoost、LGBM 等模型,通过调整特征重要性的计算参数可以得到不同方法的特征排序,这种方法便捷有效,易于理解模型与特征之间的关系,但需要以交叉验证等方式对模型拟合效果进行验证。另外,可以考虑递归特征消除方法对重症医学数据的特征进行筛选。需要说明的是,PCA 等降维方法均会影响到特征的解释性,并存在丢失特征信息的风险,在使用过程中需加以考虑。

(三)模型构建

推荐意见 22: 建议根据不同场景和不同数据

类型选择监督学习、非监督学习、半监督学习和强化学习模型进行重症疾病预测与识别。(共识度 97%)

解读与证据: 重症监护病房监测系统收集了大量患者的呼吸、血流动力学、神经和临床数据,其 EMR 系统也详细记录了患者各项临床治疗和用药信息,数据类型包括文本型、数字型和图像型。通过机器学习算法对大数据的处理和分析能力,可以挖掘数据的关键特征,辅助诊断决策。根据数据集是否有标签,机器学习算法可分为有监督学习、无监督学习、半监督学习和强化学习。其中,监督学习能够对数据集进行学习和模式归纳,常见的算法包括决策树、SVM、随机森林、朴素贝叶斯模型、ANN 络等;无监督学习模型不需要人工标注就能发现隐藏的模式或数据分组,能够发现数据中潜在的相似性和差异性,常见的算法有 K-means、PCA、层次聚类等;半监督学习通过使用有标签和无标签数据来共同训练模型;强化学习则能够从经验中学习应该采取的最佳行为或模式,模型类型应该根据数据类型和医学任务进行针对性选择。其中,对于数字型数据和临床预测类问题,可采用有监督学习模型;对于文本型数据,可采用 NLP 模型和无监督学习模型;对于图像型数据,可采用半监督学习、CNN、循环神经网络 (recurrent neural network, RNN) 进行医学图像的识别与分割;对于临床辅助决策类任务,可采用强化学习模型。根据文献调研,三种不同学习方式的使用场景包括:(1) 监督学习:预后预测、表型划分、用药策略选择、死亡风险预测、疾病严重程度预测、ICU 住院时长预测等;(2) 非监督学习:基于电子健康记录的疾病模式挖掘与表征;(3) 半监督学习:医学图像分割;(4) 强化学习:治疗方案决策、补液量推荐、机器人辅助手术等。

具体来说,三种学习的使用情况和适应证举例如下:(1) 监督学习:肝素用药患者预后预测和剂量推荐^[82];局部枸橼酸抗凝的监测和调整^[80];预测重症患者院内死亡风险^[115, 121];预测念珠菌血症患者死亡风险^[116];预测肺部超声对 ICU 患者严重程度^[117]等。(2) 无监督学习:机械通气患者表型划分与镇静策略选择^[122];重症患者体温模式识别^[67]、血压模式识别^[123];从电子健康病历数据中提取疾病亚型^[124-125]。(3) 半监督学习:基于半监督学习的三维超声医疗器械分割^[126];识别增强 CT 中的肿瘤区域^[127]等。(4) 强化学习:为 ICU 中的患者动态提供



最佳治疗方案及选择静脉输液和血管升压药物剂量^[128]。

推荐意见 23: 推荐使用因果推断模型进行重症领域因果关系的探索与发现。(共识度 89%)

解读与证据: 因果推断的模型体系建立在因果启发学习推理的基础上,对相关数据进行深入挖掘从而提取因果结构,并进行因果启发估计,研究干预变量对预后的影响程度,得到预后评估的关键指标。涉及的方向包括因果发现、因果结构学习、因果推断、因果深度学习等。针对 ICU 预后治疗差或患者生存率低,以及准确判断患者预后的影响因素的需求,基于重症疾病多中心和专病专科数据库,使用 DoWhy、CDT 和 CausalML 框架,建立因果启发的学习推理和决策评估体系。

首先,实施大数据驱动的因果结构辨识,挖掘因果关系,进行特征分析、效果分析和可解释性分析。Richens 等^[129]提出了预期失效和预期充分性的反事实的诊断策略,打破传统的根据症状诊断疾病的诊断方式,通过使用反事实问题缩小患者可能出现状况的范围。Wei 等^[130]从因果推断的角度描述了推荐系统中一些变量间的因果关系,并从事实推理出发来解决流行度偏差对于模型的影响。Goudet 等^[131]利用深度学习方法,提出因果生成神经网络,利用条件独立性和分布不对称性来发现双变量和多变量的因果结构,从观测数据中学习函数因果模型,得出临床病理特征之间的因果路线图。

此外,在因果关系的基础上进一步进行因果效应的估计,利用广义随机森林(generalized random forest, GRF)^[132]等机器学习方法计算条件平均处理效应(conditional average treatment effect, CATE)和异质处理效应(heterogenous treatment effects, HTE),对不同 ICU 干预手段下预后的差异进行预测,研究干预变量对预后的影响程度。Tan 等^[133]使用类似对抗训练的方法,给出了对于推荐系统的可解释方法。这些方法的优势在于能够用数据推理异质性的来源特征以进行一系列估计量的估计,支持高维数据和缺失数据,且具有良好的解释性。通过基于因果关系发现和因果效应估计的技术,学习最具判别力的表征,发现诊断依据和关键特征指标,准确判断患者预后,为临床治疗提供有效的干预。

(四)模型的验证

推荐意见 24: 推荐在内部模型验证的基础上增加外部验证。(共识度 94%)

解读与证据: 模型验证是在模型构建后评价模

型预测性能的过程。模型验证的重要性体现在:衡量预测模型的预测准确度,并反馈模型构建过程,必要时调整模型构建思路。模型验证思路目前较为成熟,有较为一致的方法共识。在实践中,模型验证主要分为内部验证与外部验证。专家组认为,对于模型验证过程,可采用以下方法进行评价。

内部验证: 一般情况下,需要先基于自身数据进行验证(内部验证)。即随机选取部分数据(如总体的 80%)作为训练集用于构建预测模型,其余数据作为测试集用于评价模型性能。为了验证模型在临床新产生的数据上具有良好的性能,可在随机划分的基础上增加“时空划分”,即专门划分最新一段时间的数据作为独立验证集^[134]。为了提高评价指标的估计稳健性,可采用 K 折交叉验证的方式^[135]。即将数据集等分为 K 份(如 10 份),使用 K-1 份数据构建预测模型,余下的一份数据进行验证,重复 K 次,将 K 次的模型预测评价指标取均值作为最终的模型准确度指标。内部验证实施较为简单,但由于训练集与测试集均来源于同一份数据,模型外推能力(即“泛化”能力)相对较弱。

外部验证: 不同区域、不同医院可能会由于人群、发病特点、诊疗习惯等不同导致数据分布存在差异,为了验证模型具有良好的外推性,推荐在不同区域、不同医院的多中心数据上进行外部验证。

推荐意见 25: 推荐使用灵敏度、特异度、F1 评分及 AUROC 等指标评价分类模型性能,使用决定系数(coefficient of determination, R^2)、均方误差(mean squared error, MSE)、均方根误差(root mean squared error, $RMSE$)以及平均绝对误差(mean absolute error, MAE)等指标评价回归模型性能。(共识度 91%)

解读与证据: 在模型验证时,需采用一系列评价指标衡量模型性能(即预测效果)。针对分类模型与回归模型,采用不同指标进行评价。

分类模型性能评价指标: 对于分类模型(即预测值为分类变量的模型),一般使用灵敏度(也称为召回率)、特异度、F1 评分、精准率以及 AUROC 等指标评价性能。其中, F1 评分是灵敏度与阳性精准率的调和值,取值越大模型性能越好。AUROC 是由“1-特异度”与“灵敏度”绘制的 ROC 曲线下的面积,取值越大模型性能越好。当样本类别不均衡时,推荐使用 PR 曲线下面积 AUPRC 评价模型性能。

回归模型性能评价指标: 对于回归模型(即预



测值为连续变量的模型),一般使用 R^2 、 MSE 、 $RMSE$ 、 MAE 等指标评价性能。决定系数 R^2 越接近 1,模型性能越好。 MSE 、 $RMSE$ 及 MAE 越接近 0,模型性能越好。

(五)模型的可解释性

推荐意见 26: 推荐对模型的可解释性进行探索,以利于复杂机器学习模型的临床转化,推荐的模型解释方法包括特征重要性、局部可解释不可知模型(local interpretable model-agnostic explanations, LIME)、Shapley 值。(共识度 91%)

解读与证据: 基于重症医学大数据训练的 AI 模型往往较为复杂,其复杂程度主要体现在参数多以及各种参数之间函数关系复杂。这样的复杂模型往往不利于临床医师去裂解其中的病理生理学机制,对于变量之间的因果关系难以确定,进而严重阻碍了 AI 研究成果的临床转化。模型的可解释性被认为是解决上述问题的有效方法。理解指标特征、分类、预测,进而理解为什么一个机器学习模型会做出这样的决定、什么特征在决定中起最重要作用,能让我们判断模型是否符合常理。例如,一个强化学习模型训练的 AI 医生用于治疗脓毒性休克^[10],AI 提示需要增加去甲肾上腺素,而适当限制补液,了解这样的算法背后的机制,对于模型的可信赖性非常重要。如果通过算法告诉你,该患者要增加去甲肾上腺素剂量,是因为该患者主要矛盾是外周血管扩张,而不是液体不足,则能大大增强医师对该模型使用的信心,因为 AI 做出的诊疗决策是符合临床病理生理学变化的。

此外,另外几种方法也用于模型的可解释性探索。可考虑使用特征重要性的统计量^[136]。其主要工作原理是改变数据表格中某一系列数据的排列,保持其余特征不动,评估其对预测精度的影响有多大。LIME 是一种算法^[137],它通过可解释和可信任的方式解释任何预测模型的结果。它的工作原理是围绕想要解释的预测在本地训练可解释的模型。通俗来讲,选择一个样本以及样本附近的点,然后训练一个简单模型来拟合,虽然简单模型不能在完整数据集上有效,但至少在这个点附近都是有效的,这个简单模型的特征是人类可解析的,而训练出的权重也可以表示特征重要性。Shapley 值由美国洛杉矶加州大学教授洛伊德·夏普利(Lloyd Shapley)提出,用于解决合作博弈的贡献和收益分配问题。多人合作中,单个成员的贡献不一样,收益分配也应该不一样。理想的分配方式是:贡献等

于收益,而通过 Shapley 方法就实现了贡献和收益分配的量化。一个特征的 Shapley 值是该特征在所有的特征序列中的平均边际贡献。

六、重症大数据的临床应用

推荐意见 27: 推荐将脓毒症早期预警、AKI 早期预警、病情恶化早期预警等相对成熟的预警系统进行转化及推广。(共识度 91%)

解读与证据: 预警工具的构建能够针对 ICU 各类不良事件的发生风险进行早期预测,从而有助于临床工作者及时采取措施以防患于未然,有效降低患者不良事件发生率,同时提升早期应对能力。当前,虽然国内外已对多种危重症疾病的发生及预后进行了预警模型的构建及验证,但真正开展大规模临床试验评估其应用价值的研究尚不充分,仅在脓毒症、AKI 以及病情恶化方面实施了早期预警工具的大数据临床对照试验(big-data clinical trial, BCT)。然而,不同预警工具在不同应用场景中的预测效能尚存在差异,仍需进一步推广验证。截至目前,国内外尚无成熟的疾病靶向预警工具上市。专家组认为目前能够借助 AI 技术针对 ICU 各类不良事件进行早期预警,同时需要开展 BCT 研究进一步验证预警工具的临床实用价值,从而实现疾病早发现、早诊断、早治疗。

针对脓毒症早期预警,2017 年 Shimabukuro 等^[138]开展 BCT 研究发现,使用脓毒症早期预警工具的患者住院天数显著缩短(10.3 比 13.0 d, $P=0.042$)、住院病死率显著降低(8.96% 比 21.30%, $P=0.018$)。然而, Semler 等^[139]开展的单中心 BCT 研究发现,应用脓毒症电子预警系统既不能提升脓毒症 6 h 集束化治疗完成度($P=0.159$),也不能改善患者临床结局(包括 ICU 病死率、住 ICU 天数、血管活性药物使用天数)。

针对 AKI 早期预警,2021 年美国一项大型多中心 BCT 研究发现,AKI 预警系统不能改善患者疾病进展($P=0.67$)^[140]。然而, Selby 等^[141]研究发现,AKI 预警系统无法改善患者病死率。但 Wu 等^[142]发现,面向普通病房高危患者的基于机器学习预测 AKI 的系统,可以更好地提高 AKI 的早期诊断率(预警组 7.9% 比非预警组 2.7%, $P=0.001$)。另外一项来自美国的 BCT 研究发现,AKI 电子自动警报系统并不能改善患者 7 d 内复合结局(肌酐变化的最大值以及需要接受透析或死亡)($P=0.88$)^[143]。

针对病情恶化早期预警,2020 年美国开展了一项共计纳入 43 949 人次的多中心 BCT 研究(其中



干预组 15 487 例,对照组 28 462 例),结果发现病情恶化早期预警工具能够显著降低患者病死率(校正后 $RR=0.84, 95\%CI: 0.78\sim 0.90, P<0.001$)^[144]。

推荐意见 28:建议以重症信息系统为载体,接入实时数据、输出建议决策。(共识度 91%)

解读与证据:重症患者病情复杂、变化快,ICU 设备仪器信息量大,因此数据维度高、更新频率快,应用载体应和医院信息系统有效结合,能够实时获取高维度信息,并可配属预测模型。Xia 等^[30]基于 Hadoop 分布式处理技术,设计了一个重症医疗大数据分析系统,并通过“血必净治疗 AKI 相关脓毒症的效果研究”进行了性能测试。重症医学大数据信息系统可以整合 ICU 高维度信息、实时获取分析数据,并作为载体搭载运行预测模型、评分等重症大数据成果^[145]。Boss 等^[26]开发了一项可用于收集多模式波形数据、基于 AI 计算疾病建模的在线实时 ICU 决策支持平台,称其为“ICU Cockpit”;自 2016 年以来,在该 12 张 ICU 床位的神经危重症监护室收治的 979 例患者队列中,“ICU Cockpit”平台处理和存储的数据点总数约为 889 亿个。张素珍等^[146]基于重症信息系统,使用 XGBoost 模型集成相关参数后进行机器学习,预测脓毒性休克患者 AKI 发生风险,其预测结果灵敏度为 73.3%,特异度为 71.7%,准确度为 72.5%,预测能力较传统评分明显提升。

当无重症医疗信息系统时,也可为重症大数据配属网页在线预测工具、应用程序、小程序或者床旁表单等载体。Flechet 等^[147]开发了一项 AKI 预测模型 AKI Predictor,并进行了多中心前瞻性队列研究验证,比较临床医生和 AKI Predictor 预测的效果,二者在 ICU 入院时的表现分别为:AUROC 为 0.80(95%CI: 0.69~0.92)和 0.75(95%CI: 0.62~0.88)($n=120, P=0.25$),净受益范围为 0~26% 和 0~74%;基于机器学习的 AKI Predictor 在预测 AKI-2、3 期方面取得了与医生相似的鉴别性能,总体净效益更高,因为医生高估了 AKI 的风险。这表明 AKI Predictor 对医生的预测有附加价值。该研究还配属了在线版预测模型(<https://www.akipredictor.com/en/>)。

推荐意见 29:建议当前重症诊疗实践仍应由临床医师主导,利用大数据技术协同,以提高医疗效率、保障医疗质量与安全。(共识度 98%)

解读与证据:重症大数据应用发展近年来突飞猛进,大量文章发表,包括预测疾病、预警风险、实时指导临床用药等,在可预见的将来大数据应用可

辅助 ICU 临床进行诊疗活动。但与此同时,大数据应用目前仍存在与临床结合尚浅、缺乏高质量验证、可解释性差、应用场景少及涉及伦理等问题。故本共识认为以目前大数据应用发展水平,仍应以现有的循证证据和临床经验为主导,借助大数据的技术辅助临床诊疗,提高医疗质量和效率。

大数据模型通过复杂计算产生看似准确的结果,但往往无法向最终用户提供背后的逻辑原理。AI 在确定因果关系方面表现乏力,至少其可解释性不能满足目前临床需求。基于重症大数据开发的模型若使用来自相同人群的数据进行验证时,其预测结果往往较为准确,但在外部人群测试时可能出现结果不可靠的情况^[148]。在临床实践中,诊疗过程往往存在高度主观性,尤其是复杂的重症患者,其诊疗方案也存在较大的个体异质性^[149],导致 ICU 的模型可靠性低。综上,目前大部分研究仍处于开发阶段,缺乏有效外部验证。因此,在没有科学证据支持的情况下,进行不必要的干预或改变治疗策略,可能出现过度医疗或治疗失败等医疗安全问题。

当这些算法被开发成智能辅助系统部署为警报工具时,则需足够简洁和准确,以防止警报疲劳,从而避免临床决策延迟^[150]。考虑到科学严谨性,AI 驱动模型的成熟度和稳定性对于临床实践更缺乏一定程度的说服力,不加选择地开发和使用的数据模型可能会导致过度诊断和资源浪费。此外,重症大数据在临床上应用还面临着伦理问题。目前大数据应用在患者隐私、安全责任方面的隐患不可忽视。首先数据库的建立难免涉及患者隐私数据,保护患者隐私成为重症大数据发展必须解决的问题。在未能保证隐私安全前,不应全速发展建立医疗数据库。其次应用安全性方面,在大数据辅助临床诊疗实践的过程中,如果出现医疗安全事故,以目前伦理学及 AI 的发展水平,计算机算法无法为临床决策负责,为避免大数据诊疗系统的误用滥用,临床医生必须作为临床决策的负责人,为大数据应用“兜底”。

重症大数据应用中国专家共识撰写组成员

项目主持人:隆云(中国医学科学院北京协和医院重症医学科)

执笔统筹:苏龙翔(中国医学科学院北京协和医院重症医学科);刘圣均(中国医学科学院北京协和医院重症医学科)

执笔者(按姓氏汉语拼音排序):陈朝东(北京朝阳医院重症医学科);陈凯(福建省立医院重症医学科);陈鸣(南京鼓楼医院重症医学科);程毅松(四川大学华西医院重症医学



科);崔雅婷(解放军总医院重症医学科);丁琪(北京朝阳医院重症医学科);高涛(南京鼓楼医院重症医学科);顾小花(江苏省苏北人民医院重症医学科);何怀武(中国医学科学院北京协和医院重症医学科);贺宏丽(电子科技大学附属医院四川省人民医院重症医学科);贺家伟(北京友谊医院重症医学科);胡畅(武汉大学中南医院重症医学科);黄睿(哈尔滨医科大学附属第二医院重症医学科);姜会珍(中国医学科学院北京协和医院信息中心);蒋静(重庆市人民医院重症医学科);潘盼(解放军总医院呼吸与危重症学部);兰蕴平(电子科技大学附属医院四川省人民医院重症医学科);李俊(福建省省立医院重症医学科);李璐(武汉大学中南医院重症医学科);林瑾(北京友谊医院重症医学科);罗旭飞(兰州大学循证医学中心);刘圣均(中国医学科学院北京协和医院重症医学科);毛智(解放军总医院重症医学科);苗鹤(中国医科大学附属第一医院重症医学科);申玉文(山东大学齐鲁医院心脏外科重症监护室);苏龙翔(中国医学科学院北京协和医院重症医学科);孙启航(英国华人医疗信息协会);孙为军(广东工业大学自动化学院);唐之韵(浙江省人民医院重症医学科);王海军(中国医学科学院肿瘤医院肿瘤研究所重症医学科);王陆豪(中山大学附属第一医院重症医学科);王斯聪(哈尔滨医科大学附属第二医院重症医学科);王占稳(中南大学湘雅医院重症医学科);王仲(中国医科大学附属第一医院重症医学科);尉东(山东大学齐鲁医院心脏外科重症监护室);吴骏(四川大学华西医院重症医学科);余愿(华中科技大学同济医学院附属协和医院重症医学科);於江泉(江苏省苏北人民医院重症医学科);原皓(中山大学附属第一医院重症医学科);张昊(中国医学科学院肿瘤医院肿瘤研究所重症医学科);张建成(华中科技大学同济医学院附属协和医院重症医学科);张萌(重庆市人民医院重症医学科);赵春光(中南大学湘雅医院重症医学科);钟磊(浙江省人民医院重症医学科)

共识制定专家组(按姓氏汉语拼音排序):陈耀龙(中国医学科学院循证评价与指南研究创新单元主任);翟茜(山东大学齐鲁医院心脏外科重症监护室);丁仁彧(中国医科大学附属第一医院重症医学科);段美丽(北京友谊医院重症医学科);胡波(武汉大学中南医院重症医学科);黄晓波(电子科技大学附属医院四川省人民医院重症医学科);李林峰(重庆医科大学医学数据研究院);李文雄(北京朝阳医院重症医学科);李永在(山东大学齐鲁医院信息中心);刘娇(上海交通大学医学院附属瑞金医院北院重症医学科);隆云(中国医学科学院北京协和医院重症医学科);吕丰(中南大学计算机学院);尚小溥(北京交通大学经济管理学院信息管理系);尚秀玲(福建省省立医院重症医学科);尚游(华中科技大学同济医学院附属协和医院重症医学科);史颖欢(南京大学健康医疗大数据国家研究院);王波(四川大学华西医院重症医学科);王洪亮(哈尔滨医科大学附属第二医院重症医学科);王丽(中国医学科学院基础医学

研究所流行病学与卫生统计学系);吴健锋(中山大学附属第一医院重症医学科);邢学忠(中国医学科学院肿瘤医院肿瘤研究所重症医学科);杨缙(重庆市人民医院重症医学科);杨向红(浙江省人民医院重症医学科);虞文魁(南京鼓楼医院重症医学科);张丽娜(中南大学湘雅医院重症医学科);章仲恒(浙江大学医学院附属邵逸夫医院急诊科);郑瑞强(江苏省苏北人民医院重症医学科);周飞虎(解放军总医院重症医学科);朱卫国(中国医学科学院北京协和医院全科医学科)

利益冲突 所有作者均声明不存在利益冲突

参 考 文 献

- [1] Sanchez-Pinto LN, Luo Y, Churpek MM. Big data and data science in critical Care[J]. Chest, 2018, 154(5):1239-1248. DOI: 10.1016/j.chest.2018.04.037.
- [2] Rumsfeld JS, Joynt KE, Maddox TM. Big data analytics to improve cardiovascular care: promise and challenges[J]. Nat Rev Cardiol, 2016, 13(6): 350-359. DOI: 10.1038/nrcardio.2016.42.
- [3] Yang S, Stansbury LG, Rock P, et al. Linking big data and prediction strategies: tools, pitfalls, and lessons learned [J]. Crit Care Med, 2019, 47(6):840-848. DOI: 10.1097/CCM.0000000000003739.
- [4] Carra G, Salluh J, da Silva Ramos FJ, et al. Data-driven ICU management: using big data and algorithms to improve outcomes[J]. J Crit Care, 2020, 60:300-304. DOI: 10.1016/j.jcrrc.2020.09.002.
- [5] Le Roux P, Menon DK, Citerio G, et al. Consensus summary statement of the International Multidisciplinary Consensus Conference on Multimodality Monitoring in Neurocritical Care: a statement for healthcare professionals from the Neurocritical Care Society and the European Society of Intensive Care Medicine[J]. Neurocrit Care, 2014, 21 Suppl 2: S1-S26. DOI: 10.1007/s12028-014-0041-5.
- [6] Schmidt JM, De Georgia M. Multimodality monitoring: informatics, integration data display and analysis[J]. Neurocrit Care, 2014, 21 Suppl 2: S229-S238. DOI: 10.1007/s12028-014-0037-1.
- [7] Citerio G, Park S, Schmidt JM, et al. Data collection and interpretation[J]. Neurocrit Care, 2015, 22(3): 360-368. DOI: 10.1007/s12028-015-0139-4.
- [8] Docherty AB, Lone NI. Exploiting big data for critical care research[J]. Curr Opin Crit Care, 2015, 21(5): 467-472. DOI: 10.1097/MCC.0000000000000228.
- [9] Zampieri FG, Soares M. Reply to: the Epimed Monitor ICU Database®: a cloud-based national registry for adult intensive care unit patients in Brazil[J]. Rev Bras Ter Intensiva, 2018, 30(3): 398. DOI: 10.5935/0103-507x.20180048.
- [10] Komorowski M, Celi LA, Badawi O, et al. The artificial intelligence clinician learns optimal treatment strategies for sepsis in intensive care[J]. Nat Med, 2018, 24(11): 1716-1720. DOI: 10.1038/s41591-018-0213-5.
- [11] Nemati S, Holder A, Razmi F, et al. An interpretable machine learning model for accurate prediction of sepsis in the ICU[J]. Crit Care Med, 2018, 46(4):547-553. DOI:

- 10.1097/CCM.0000000000002936.
- [12] Celi LA, Mark RG, Stone DJ, et al. "Big data" in the intensive care unit. Closing the data loop[J]. *Am J Respir Crit Care Med*, 2013, 187(11):1157-1160. DOI: 10.1164/rccm.201212-2311ED.
- [13] Goldberger AL, Amaral LA, Glass L, et al. PhysioBank, PhysioToolkit, and PhysioNet: components of a new research resource for complex physiologic signals[J]. *Circulation*, 2000, 101(23):E215-E220. DOI: 10.1161/01.cir.101.23.e215.
- [14] Pollard TJ, Johnson A, Raffa JD, et al. The eICU Collaborative Research Database, a freely available multi-center database for critical care research[J]. *Sci Data*, 2018, 5:180178. DOI: 10.1038/sdata.2018.178.
- [15] Hyland SL, Faltys M, Hüser M, et al. Early prediction of circulatory failure in the intensive care unit using machine learning[J]. *Nat Med*, 2020, 26(3):364-373. DOI: 10.1038/s41591-020-0789-4.
- [16] Thorat PJ, Peppink JM, Driessen RH, et al. Sharing ICU Patient Data Responsibly Under the Society of Critical Care Medicine/European Society of Intensive Care Medicine Joint Data Science Collaboration: the Amsterdam University Medical Centers Database (AmsterdamUMCdb) Example[J]. *Crit Care Med*, 2021, 49(6):e563-e577. DOI: 10.1097/CCM.0000000000004916.
- [17] Qi S, Mao Z, Hu X, et al. Introduction of critical care database based on specialized information systems: a model of critical care medicine database in large Level III Grade A hospital[J]. *Zhonghua Wei Zhong Bing Ji Jiu Yi Xue*, 2020, 32(6): 743-749. DOI: 10.3760/cma.j.cn121430-20200520-00393.
- [18] Xu P, Chen L, Zhu Y, et al. Critical care database comprising patients with infection[J]. *Front Public Health*, 2022, 10:852410. DOI: 10.3389/fpubh.2022.852410.
- [19] Zeng X, Yu G, Lu Y, et al. PIC, a paediatric-specific intensive care database[J]. *Sci Data*, 2020, 7(1):14. DOI: 10.1038/s41597-020-0355-4.
- [20] Zhang Z, Cao L, Chen R, et al. Electronic healthcare records and external outcome data for hospitalized patients with heart failure[J]. *Sci Data*, 2021, 8(1):46. DOI: 10.1038/s41597-021-00835-9.
- [21] 许杰, 周瑜, 夏星球, 等. 重症医学专科大数据平台的建设及应用[J]. *中华急诊医学杂志*, 2022, 31(1):129-132. DOI: 10.3760/cma.j.issn.1671-0282.2022.01.028.
- [22] Morrison JL, Cai Q, Davis N, et al. Clinical and economic outcomes of the electronic intensive care unit: results from two community hospitals[J]. *Crit Care Med*, 2010, 38(1):2-8. DOI: 10.1097/CCM.0b013e3181b78fa8.
- [23] Elbers PW, Girbes A, Malbrain ML, et al. Right dose, right now: using big data to optimize antibiotic dosing in the critically ill[J]. *Anaesthesiol Intensive Ther*, 2015, 47(5): 457-463. DOI: 10.5603/AITa2015.0061.
- [24] Kindle RD, Badawi O, Celi LA, et al. Intensive care unit telemedicine in the era of big data, artificial intelligence, and computer clinical decision support systems[J]. *Crit Care Clin*, 2019, 35(3): 483-495. DOI: 10.1016/j.ccc.2019.02.005.
- [25] Noshad M, Rose CC, Chen JH. Signal from the noise: a mixed graphical and quantitative process mining approach to evaluate care pathways applied to emergency stroke care[J]. *J Biomed Inform*, 2022, 127: 104004. DOI: 10.1016/j.jbi.2022.104004.
- [26] Boss JM, Narula G, Straessle C, et al. ICU Cockpit: a platform for collecting multimodal waveform data, AI-based computational disease modeling and real-time decision support in the intensive care unit[J]. *J Am Med Inform Assoc*, 2022, 29(7): 1286-1291. DOI: 10.1093/jamia/ocac064.
- [27] Taglang G, Jackson DB. Use of "big data" in drug discovery and clinical trials[J]. *Gynecol Oncol*, 2016, 141(1):17-23. DOI: 10.1016/j.ygyno.2016.02.022.
- [28] Zhu Y, Yin H, Zhang R, et al. The effect of dobutamine vs milrinone in sepsis: a big data, real-world study[J]. *Int J Clin Pract*, 2021, 75(11): e14689. DOI: 10.1111/ijcp.14689.
- [29] Huang X, Shan S, Khan YA, et al. Risk assessment of ICU patients through deep learning technique: a big data approach[J]. *J Glob Health*, 2022, 12:04044. DOI: 10.7189/jogh.12.04044.
- [30] Xia Y, Wang X, Wu W, et al. Rehabilitation of sepsis patients with acute kidney injury based on intelligent medical big data[J]. *J Healthc Eng*, 2022, 2022:8414135. DOI: 10.1155/2022/8414135.
- [31] Vergetis V, Skaltsas D, Gorgoulis VG, et al. Assessing Drug Development Risk Using Big Data and Machine Learning [J]. *Cancer Res*, 2021, 81(4): 816-819. DOI: 10.1158/0008-5472.CAN-20-0866.
- [32] Fleuren LM, Klausch T, Zwager CL, et al. Machine learning for the prediction of sepsis: a systematic review and meta-analysis of diagnostic test accuracy[J]. *Intensive Care Med*, 2020, 46(3): 383-400. DOI: 10.1007/s00134-019-05872-y.
- [33] Delahanty RJ, Alvarez J, Flynn LM, et al. Development and evaluation of a machine learning model for the early identification of patients at risk for sepsis[J]. *Ann Emerg Med*, 2019, 73(4): 334-344. DOI: 10.1016/j.annemergmed.2018.11.036.
- [34] Goh KH, Wang L, Yeow A, et al. Artificial intelligence in sepsis early prediction and diagnosis using unstructured data in healthcare[J]. *Nat Commun*, 2021, 12(1):711. DOI: 10.1038/s41467-021-20910-4.
- [35] Wardi G, Brice J, Correia M, et al. Demystifying lactate in the emergency department[J]. *Ann Emerg Med*, 2020, 75(2): 287-298. DOI: 10.1016/j.annemergmed. 2019. 06.027.
- [36] Ozrazgat-Baslanti T, Loftus TJ, Ren Y, et al. Advances in artificial intelligence and deep learning systems in ICU-related acute kidney injury[J]. *Curr Opin Crit Care*, 2021, 27(6): 560-572. DOI: 10.1097/MCC.0000000000000887.
- [37] Xiao Z, Huang Q, Yang Y, et al. Emerging early diagnostic methods for acute kidney injury[J]. *Theranostics*, 2022, 12(6):2963-2986. DOI: 10.7150/thno.71064.
- [38] Churpek MM, Carey KA, Edelson DP, et al. Internal and external validation of a machine learning risk score for acute kidney injury[J]. *JAMA Netw Open*, 2020, 3(8): e2012892. DOI: 10.1001/jamanetworkopen.2020.12892.
- [39] Kate RJ, Pearce N, Mazumdar D, et al. A continual prediction model for inpatient acute kidney injury[J]. *Comput Biol Med*, 2020, 116: 103580. DOI: 10.1016/j.combiomed.2019.103580.
- [40] Le S, Pellegrini E, Green-Saxena A, et al. Supervised

- machine learning for the early prediction of acute respiratory distress syndrome (ARDS) [J]. *J Crit Care*, 2020, 60:96-102. DOI: 10.1016/j.jcrc.2020.07.019.
- [41] Mayampurath A, Churpek MM, Su X, et al. External validation of an acute respiratory distress syndrome prediction model using radiology reports[J]. *Crit Care Med*, 2020, 48(9): e791-e798. DOI: 10.1097/CCM.0000000000004468.
- [42] Sjoding MW, Taylor D, Motyka J, et al. Deep learning to detect acute respiratory distress syndrome on chest radiographs: a retrospective study with external validation[J]. *Lancet Digit Health*, 2021, 3(6):e340-e348. DOI: 10.1016/S2589-7500(21)00056-X.
- [43] Linnen DT, Escobar GJ, Hu X, et al. Statistical modeling and aggregate-weighted scoring systems in prediction of mortality and ICU transfer: a systematic review[J]. *J Hosp Med*, 2019, 14(3):161-169. DOI: 10.12788/jhm.3151.
- [44] Muralitharan S, Nelson W, Di S, et al. Machine learning-based early warning systems for clinical deterioration: systematic scoping review[J]. *J Med Internet Res*, 2021, 23(2):e25187. DOI: 10.2196/25187.
- [45] Tisdale JE, Jaynes HA, Overholser BR, et al. Enhanced response to drug-induced QT interval lengthening in patients with heart failure with preserved ejection fraction[J]. *J Card Fail*, 2020, 26(9): 781-785. DOI: 10.1016/j.cardfail.2020.06.008.
- [46] Broch Porcar MJ, Rodríguez Cubillo B, Domínguez-Roldán JM, et al. Practical document on the management of hyponatremia in critically ill patients[J]. *Med Intensiva (Engl Ed)*, 2019, 43(5): 302-316. DOI: 10.1016/j.medint.2018.12.002.
- [47] Levi R, Carli F, Arévalo AR, et al. Artificial intelligence-based prediction of transfusion in the intensive care unit in patients with gastrointestinal bleeding[J]. *BMJ Health Care Inform*, 2021, 28(1)DOI: 10.1136/bmjhci-2020-100245.
- [48] Ryan L, Mataraso S, Siefkas A, et al. A machine learning approach to predict deep venous thrombosis among hospitalized patients[J]. *Clin Appl Thromb Hemost*, 2021, 27: 1076029621991185. DOI: 10.1177/1076029621991185.
- [49] Yuan KC, Tsai LW, Lee KH, et al. The development an artificial intelligence algorithm for early sepsis diagnosis in the intensive care unit[J]. *Int J Med Inform*, 2020, 141: 104176. DOI: 10.1016/j.ijmedinf.2020.104176.
- [50] Burdick H, Pino E, Gabel-Comeau D, et al. Effect of a sepsis prediction algorithm on patient mortality, length of stay and readmission: a prospective multicentre clinical outcomes evaluation of real-world patient data from US hospitals[J]. *BMJ Health Care Inform*, 2020, 27(1)DOI: 10.1136/bmjhci-2019-100109.
- [51] Bose SN, Greenstein JL, Fackler JC, et al. Early prediction of multiple organ dysfunction in the pediatric intensive care unit[J]. *Front Pediatr*, 2021, 9:711104. DOI: 10.3389/fped.2021.711104.
- [52] Romero-Brufau S, Whitford D, Johnson MG, et al. Using machine learning to improve the accuracy of patient deterioration predictions: Mayo Clinic Early Warning Score (MC-EWS) [J]. *J Am Med Inform Assoc*, 2021, 28(6): 1207-1215. DOI: 10.1093/jamia/ocaa347.
- [53] Bedoya AD, Clement ME, Phelan M, et al. Minimal impact of implemented early warning score and best practice alert for patient deterioration[J]. *Crit Care Med*, 2019, 47(1):49-55. DOI: 10.1097/CCM.0000000000003439.
- [54] Ardila D, Kiraly AP, Bharadwaj S, et al. Author correction: end-to-end lung cancer screening with three-dimensional deep learning on low-dose chest computed tomography [J]. *Nat Med*, 2019, 25(8): 1319. DOI: 10.1038/s41591-019-0536-x.
- [55] Qin ZZ, Sander MS, Rai B, et al. Using artificial intelligence to read chest radiographs for tuberculosis detection: a multi-site evaluation of the diagnostic accuracy of three deep learning systems[J]. *Sci Rep*, 2019, 9(1):15000. DOI: 10.1038/s41598-019-51503-3.
- [56] Reamaroon N, Sjoding MW, Gryak J, et al. Automated detection of acute respiratory distress syndrome from chest X-Rays using Directionality Measure and deep learning features[J]. *Comput Biol Med*, 2021, 134:104463. DOI: 10.1016/j.combiomed.2021.104463.
- [57] Rueckel J, Kunz WG, Hoppe BF, et al. Artificial intelligence algorithm detecting lung infection in supine chest radiographs of critically ill patients with a diagnostic accuracy similar to board-certified radiologists[J]. *Crit Care Med*, 2020, 48(7): e574-e583. DOI: 10.1097/CCM.0000000000004397.
- [58] Zhang K, Liu X, Shen J, et al. Clinically applicable ai system for accurate diagnosis, quantitative measurements, and prognosis of COVID-19 pneumonia using computed tomography[J]. *Cell*, 2020, 182(5): 1360. DOI: 10.1016/j.cell.2020.08.029.
- [59] Farzaneh N, Williamson CA, Jiang C, et al. Automated segmentation and severity analysis of subdural hematoma for patients with traumatic brain injuries[J]. *Diagnostics (Basel)*, 2020, 10(10)DOI: 10.3390/diagnostics10100773.
- [60] van Sloun R, Demi L. Localizing B-lines in lung ultrasonography by weakly supervised deep learning, in-vivo results[J]. *IEEE J Biomed Health Inform*, 2020, 24(4):957-964. DOI: 10.1109/JBHI.2019.2936151.
- [61] Silva S, Ait Aissa D, Cocquet P, et al. Combined thoracic ultrasound assessment during a successful weaning trial predicts postextubation distress[J]. *Anesthesiology*, 2017, 127(4):666-674. DOI: 10.1097/ALN.0000000000001773.
- [62] Lv Y, Huang Z. Account of deep learning-based ultrasonic image feature in the diagnosis of severe sepsis complicated with acute kidney injury[J]. *Comput Math Methods Med*, 2022, 2022:8158634. DOI: 10.1155/2022/8158634.
- [63] Ying F, Chen S, Pan G, et al. Artificial intelligence pulse coupled neural network algorithm in the diagnosis and treatment of severe sepsis complicated with acute kidney injury under ultrasound image[J]. *J Healthc Eng*, 2021, 2021:6761364. DOI: 10.1155/2021/6761364.
- [64] Strodthoff N, Strodthoff C, Becher T, et al. Inferring respiratory and circulatory parameters from electrical impedance tomography with deep recurrent models[J]. *IEEE J Biomed Health Inform*, 2021, 25(8): 3105-3111. DOI: 10.1109/JBHI.2021.3059016.
- [65] Bos LD, Schouten LR, van Vught LA, et al. Identification and validation of distinct biological phenotypes in patients with acute respiratory distress syndrome by cluster analysis[J]. *Thorax*, 2017, 72(10): 876-883. DOI:

- 10.1136/thoraxjnl-2016-209719.
- [66] Chaudhary K, Vaid A, Duffy Á, et al. Utilization of deep learning for subphenotype identification in sepsis-associated acute kidney injury[J]. *Clin J Am Soc Nephrol*, 2020, 15(11): 1557-1565. DOI: 10.2215/CJN.09330819.
- [67] Bhavani SV, Carey KA, Gilbert ER, et al. Identifying novel sepsis subphenotypes using temperature trajectories[J]. *Am J Respir Crit Care Med*, 2019, 200(3): 327-335. DOI: 10.1164/rccm.201806-11970C.
- [68] Sinha P, Churpek MM, Calfee CS. Machine learning classifier models can identify acute respiratory distress syndrome phenotypes using readily available clinical data [J]. *Am J Respir Crit Care Med*, 2020, 202(7): 996-1004. DOI: 10.1164/rccm.202002-03470C.
- [69] Kudo D, Goto T, Uchimido R, et al. Coagulation phenotypes in sepsis and effects of recombinant human thrombomodulin: an analysis of three multicentre observational studies[J]. *Crit Care*, 2021, 25(1): 114. DOI: 10.1186/s13054-021-03541-5.
- [70] Seymour CW, Kennedy JN, Wang S, et al. Derivation, validation, and potential treatment implications of novel clinical phenotypes for Sepsis[J]. *JAMA*, 2019, 321(20): 2003-2017. DOI: 10.1001/jama.2019.5791.
- [71] Wiersema R, Jukarainen S, Vaara ST, et al. Two subphenotypes of septic acute kidney injury are associated with different 90-day mortality and renal recovery[J]. *Crit Care*, 2020, 24(1): 150. DOI: 10.1186/s13054-020-02866-x.
- [72] Scicluna BP, van Vught LA, Zwinderman AH, et al. Classification of patients with sepsis according to blood genomic endotype: a prospective cohort study[J]. *Lancet Respir Med*, 2017, 5(10): 816-826. DOI: 10.1016/S2213-2600(17)30294-1.
- [73] Calfee CS, Delucchi K, Parsons PE, et al. Subphenotypes in acute respiratory distress syndrome: latent class analysis of data from two randomised controlled trials[J]. *Lancet Respir Med*, 2014, 2(8): 611-620. DOI: 10.1016/S2213-2600(14)70097-9.
- [74] Famous KR, Delucchi K, Ware LB, et al. Acute respiratory distress syndrome subphenotypes respond differently to randomized fluid management strategy[J]. *Am J Respir Crit Care Med*, 2017, 195(3): 331-338. DOI: 10.1164/rccm.201603-06450C.
- [75] Calfee CS, Delucchi KL, Sinha P, et al. Acute respiratory distress syndrome subphenotypes and differential response to simvastatin: secondary analysis of a randomised controlled trial[J]. *Lancet Respir Med*, 2018, 6(9): 691-698. DOI: 10.1016/S2213-2600(18)30177-2.
- [76] Zhang Z, Ho KM, Hong Y. Machine learning for the prediction of volume responsiveness in patients with oliguric acute kidney injury in critical care[J]. *Crit Care*, 2019, 23(1): 112. DOI: 10.1186/s13054-019-2411-z.
- [77] Parreco J, Hidalgo A, Parks JJ, et al. Using artificial intelligence to predict prolonged mechanical ventilation and tracheostomy placement[J]. *J Surg Res*, 2018, 228: 179-187. DOI: 10.1016/j.jss.2018.03.028.
- [78] Fabregat A, Magret M, Ferré JA, et al. A machine learning decision-making tool for extubation in intensive care unit patients[J]. *Comput Methods Programs Biomed*, 2021, 200: 105869. DOI: 10.1016/j.cmpb.2020.105869.
- [79] Hur S, Min JY, Yoo J, et al. Development and validation of unplanned extubation prediction models using intensive care unit data: retrospective, comparative, machine learning study[J]. *J Med Internet Res*, 2021, 23(8): e23508. DOI: 10.2196/23508.
- [80] Chen H, Ma Y, Hong N, et al. Early warning of citric acid overdose and timely adjustment of regional citrate anticoagulation based on machine learning methods[J]. *BMC Med Inform Decis Mak*, 2021, 21(Suppl 2): 126. DOI: 10.1186/s12911-021-01489-8.
- [81] Su L, Liu C, Li D, et al. Toward optimal heparin dosing by comparing multiple machine learning methods: retrospective study[J]. *JMIR Med Inform*, 2020, 8(6): e17648. DOI: 10.2196/17648.
- [82] Li D, Gao J, Hong N, et al. A clinical prediction model to predict heparin treatment outcomes and provide dosage recommendations: development and validation study[J]. *J Med Internet Res*, 2021, 23(5): e27118. DOI: 10.2196/27118.
- [83] Maviglia R, Michi T, Passaro D, et al. Machine learning and antibiotic management[J]. *Antibiotics (Basel)*, 2022, 11(3) DOI: 10.3390/antibiotics11030304.
- [84] Wang Y, Zhang H, Fan Y, et al. Propofol anesthesia depth monitoring based on self-attention and residual structure convolutional neural network[J]. *Comput Math Methods Med*, 2022, 2022: 8501948. DOI: 10.1155/2022/8501948.
- [85] van de Sande D, van Genderen ME, Huiskens J, et al. Moving from bytes to bedside: a systematic review on the use of artificial intelligence in the intensive care unit[J]. *Intensive Care Med*, 2021, 47(7): 750-760. DOI: 10.1007/s00134-021-06446-7.
- [86] Barchitta M, Maugeri A, Favara G, et al. Early prediction of seven-day mortality in intensive care unit using a machine learning model: results from the SPIN-UTI project[J]. *J Clin Med*, 2021, 10(5) DOI: 10.3390/jcm10050992.
- [87] Li K, Shi Q, Liu S, et al. Predicting in-hospital mortality in ICU patients with sepsis using gradient boosting decision tree[J]. *Medicine (Baltimore)*, 2021, 100(19): e25813. DOI: 10.1097/MD.00000000000025813.
- [88] García-Gallo JE, Fonseca-Ruiz NJ, Celi LA, et al. A machine learning-based model for 1-year mortality prediction in patients admitted to an intensive care unit with a diagnosis of sepsis[J]. *Med Intensiva (Engl Ed)*, 2020, 44(3): 160-170. DOI: 10.1016/j.medint.2018.07.016.
- [89] Nie X, Cai Y, Liu J, et al. Mortality prediction in cerebral hemorrhage patients using machine learning algorithms in intensive care units[J]. *Front Neurol*, 2020, 11: 610531. DOI: 10.3389/fneur.2020.610531.
- [90] Halonen KI, Leppäniemi AK, Lundin JE, et al. Predicting fatal outcome in the early phase of severe acute pancreatitis by using novel prognostic models[J]. *Pancreatol*, 2003, 3(4): 309-315. DOI: 10.1159/000071769.
- [91] Ding N, Guo C, Li C, et al. An artificial neural networks model for early predicting in-hospital mortality in acute pancreatitis in MIMIC-III [J]. *Biomed Res Int*, 2021, 2021: 6638919. DOI: 10.1155/2021/6638919.
- [92] Weiss CH. Why do we fail to deliver evidence-based practice in critical care medicine? [J]. *Curr Opin Crit Care*, 2017, 23(5): 400-405. DOI: 10.1097/mcc.0000000000000436.



- [93] Rosa RG, Teixeira C, Sjoding M. Novel approaches to facilitate the implementation of guidelines in the ICU[J]. *J Crit Care*, 2020, 60:1-5. DOI: 10.1016/j.jcrc.2020.07.014.
- [94] Giuliano KK, Lecardo M, Staul L. Impact of protocol watch on compliance with the surviving sepsis campaign[J]. *Am J Crit Care*, 2011, 20(4): 313-321. DOI: 10.4037/ajcc2011421.
- [95] Liu VX, Morehouse JW, Marelich GP, et al. Multicenter implementation of a treatment bundle for patients with sepsis and intermediate lactate values[J]. *Am J Respir Crit Care Med*, 2016, 193(11): 1264-1270. DOI: 10.1164/rccm.201507-14890C.
- [96] Eslami S, Abu-Hanna A, Schultz MJ, et al. Evaluation of consulting and critiquing decision support systems: effect on adherence to a lower tidal volume mechanical ventilation strategy[J]. *J Crit Care*, 2012, 27(4):425.e1-e8. DOI: 10.1016/j.jcrc.2011.07.082.
- [97] Trogrlić Z, van der Jagt M, Lingsma H, et al. Improved guideline adherence and reduced brain dysfunction after a multicenter multifaceted implementation of ICU delirium guidelines in 3, 930 patients[J]. *Crit Care Med*, 2019, 47(3): 419-427. DOI: 10.1097/CCM.0000000000003596.
- [98] Bourdeaux C, Ghosh E, Atallah L, et al. Impact of a computerized decision support tool deployed in two intensive care units on acute kidney injury progression and guideline compliance: a prospective observational study[J]. *Crit Care*, 2020, 24(1): 656. DOI: 10.1186/s13054-020-03343-1.
- [99] American College of Chest Physicians/Society of Critical Care Medicine Consensus Conference: definitions for sepsis and organ failure and guidelines for the use of innovative therapies in sepsis[J]. *Crit Care Med*, 1992, 20(6):864-874.
- [100] Levy MM, Fink MP, Marshall JC, et al. 2001 SCCM/ESICM/ACCP/ATS/SIS International Sepsis Definitions Conference[J]. *Crit Care Med*, 2003, 31(4): 1250-1256. DOI: 10.1097/01.Ccm.0000050454.01978.3b.
- [101] 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.
- [102] 朱志勇, 陈一昕, 李建功. 大数据技术在医疗急重症领域的应用[J]. *邮电设计技术*, 2016, (8):28-32. DOI: 10.16463/j.cnki.issn1007-3043.2016.08.006.
- [103] 刘再毅, 石镇维, 梁长虹. 推进联邦学习技术在医学影像人工智能中的应用[J]. *中华医学杂志*, 2022, 102(5): 318-320. DOI: 10.3760/cmaj.cn112137-20210619-01389.
- [104] Rieke N, Hancox J, Li W, et al. The future of digital health with federated learning[J]. *NPJ Digit Med*, 2020, 3: 119. DOI: 10.1038/s41746-020-00323-1.
- [105] Mahmoudi E, Kamdar N, Kim N, et al. Use of electronic medical records in development and validation of risk prediction models of hospital readmission: systematic review[J]. *BMJ*, 2020, 369:m958. DOI: 10.1136/bmj.m958.
- [106] Liang H, Tsui BY, Ni H, et al. Evaluation and accurate diagnoses of pediatric diseases using artificial intelligence [J]. *Nat Med*, 2019, 25(3): 433-438. DOI: 10.1038/s41591-018-0335-9.
- [107] Rusin CG, Acosta SI, Vu EL, et al. Automated prediction of cardiorespiratory deterioration in patients with single ventricle[J]. *J Am Coll Cardiol*, 2021, 77(25):3184-3192. DOI: 10.1016/j.jacc.2021.04.072.
- [108] Narula G, Haeberlin M, Balsiger J, et al. Detection of EEG burst-suppression in neurocritical care patients using an unsupervised machine learning algorithm[J]. *Clin Neurophysiol*, 2021, 132(10):2485-2492. DOI: 10.1016/j.clinph.2021.07.018.
- [109] Maddali MV, Churpek M, Pham T, et al. Validation and utility of ARDS subphenotypes identified by machine-learning models using clinical data: an observational, multicohort, retrospective analysis[J]. *Lancet Respir Med*, 2022, 10(4):367-377. DOI: 10.1016/S2213-2600(21)00461-6.
- [110] Niazi M, Parwani AV, Gurcan MN. Digital pathology and artificial intelligence[J]. *Lancet Oncol*, 2019, 20(5): e253-e261. DOI: 10.1016/S1470-2045(19)30154-8.
- [111] Massion PP, Antic S, Ather S, et al. Assessing the accuracy of a deep learning method to risk stratify indeterminate pulmonary nodules[J]. *Am J Respir Crit Care Med*, 2020, 202(2):241-249. DOI: 10.1164/rccm.201903-05050C.
- [112] Cheng N, Ren Y, Zhou J, et al. Deep learning-based classification of hepatocellular nodular lesions on whole-slide histopathologic images[J]. *Gastroenterology*, 2022, 162(7): 1948-1961.e7. DOI: 10.1053/j.gastro.2022.02.025.
- [113] Walsh S, Humphries SM, Wells AU, et al. Imaging research in fibrotic lung disease; applying deep learning to unsolved problems[J]. *Lancet Respir Med*, 2020, 8(11): 1144-1153. DOI: 10.1016/S2213-2600(20)30003-5.
- [114] Walsh S, Calandriello L, Silva M, et al. Deep learning for classifying fibrotic lung disease on high-resolution computed tomography: a case-cohort study[J]. *Lancet Respir Med*, 2018, 6(11): 837-845. DOI: 10.1016/S2213-2600(18)30286-8.
- [115] Winslow CJ, Edelson DP, Churpek MM, et al. The impact of a machine learning early warning score on hospital mortality: a multicenter clinical intervention trial[J]. *Crit Care Med*, 2022, 50(9): 1339-1347. DOI: 10.1097/CCM.0000000000005492.
- [116] Yuan S, Sun Y, Xiao X, et al. Using machine learning algorithms to predict candidaemia in ICU patients with new-onset systemic inflammatory response syndrome[J]. *Front Med (Lausanne)*, 2021, 8: 720926. DOI: 10.3389/fmed.2021.720926.
- [117] Dastider AG, Sadik F, Fattah SA. An integrated autoencoder-based hybrid CNN-LSTM model for COVID-19 severity prediction from lung ultrasound[J]. *Comput Biol Med*, 2021, 132: 104296. DOI: 10.1016/j.combiomed.2021.104296.
- [118] Tiwari P, Colborn KL, Smith DE, et al. Assessment of a machine learning model applied to harmonized electronic health record data for the prediction of incident atrial fibrillation[J]. *JAMA Network Open*, 2020, 3(1): e1919396.
- [119] Papp L, Spielvogel CP, Grubmüller B, et al. Supervised machine learning enables non-invasive lesion characterization in primary prostate cancer with [(68)Ga] Ga-PSMA-11 PET/MRI[J]. *Eur J Nucl Med Mol Imaging*, 2021, 48(6): 1795-1805. DOI: 10.1007/s00259-020-05140-y.
- [120] Liang W, Yao J, Chen A, et al. Early triage of critically ill COVID-19 patients using deep learning[J]. *Nat Commun*, 2020, 11(1):3543. DOI: 10.1038/s41467-020-17280-8.

- [121] Su L, Xu Z, Chang F, et al. Early prediction of mortality, severity, and length of stay in the intensive care unit of sepsis patients based on sepsis 3.0 by machine learning models[J]. *Front Med (Lausanne)*, 2021, 8:664966. DOI: 10.3389/fmed.2021.664966.
- [122] Su L, Zhang Z, Zheng F, et al. Five novel clinical phenotypes for critically ill patients with mechanical ventilation in intensive care units: a retrospective and multi database study[J]. *Respir Res*, 2020, 21(1):325. DOI: 10.1186/s12931-020-01588-6.
- [123] Liu S, Su L, Liu X, et al. Recognizing blood pressure patterns in sedated critically ill patients on mechanical ventilation by spectral clustering[J]. *Ann Transl Med*, 2021, 9(18):1404. DOI: 10.21037/atm-21-2806.
- [124] Landi I, Glicksberg BS, Lee HC, et al. Deep representation learning of electronic health records to unlock patient stratification at scale[J]. *NPJ Digit Med*, 2020, 3:96. DOI: 10.1038/s41746-020-0301-z.
- [125] Hyun S, Kaewprag P, Cooper C, et al. Exploration of critical care data by using unsupervised machine learning[J]. *Comput Methods Programs Biomed*, 2020, 194:105507. DOI: 10.1016/j.cmpb.2020.105507.
- [126] Yang H, Shan C, Bouwman A, et al. Medical instrument segmentation in 3D US by hybrid constrained semi-supervised learning[J]. *IEEE J Biomed Health Inform*, 2022, 26(2): 762-773. DOI: 10.1109/JBHI.2021.3101872.
- [127] Heller N, Isensee F, Maier-Hein KH, et al. The state of the art in kidney and kidney tumor segmentation in contrast-enhanced CT imaging: results of the KiTS19 challenge[J]. *Med Image Anal*, 2021, 67: 101821. DOI: 10.1016/j.media.2020.101821.
- [128] Greco M, Caruso PF, Cecconi M. Artificial intelligence in the intensive care unit[J]. *Semin Respir Crit Care Med*, 2021, 42(1):2-9. DOI: 10.1055/s-0040-1719037.
- [129] Richens JG, Lee CM, Johri S. Improving the accuracy of medical diagnosis with causal machine learning[J]. *Nat Commun*, 2020, 11(1): 3923. DOI: 10.1038/s41467-020-17419-7.
- [130] Wei T, Feng F, Chen J, et al. Model-agnostic counterfactual reasoning for eliminating popularity bias in recommender system[EB/OL]. [2022-02-01]. <https://arxiv.org/abs/2010.15363>.
- [131] Goudet O, Kalainathan D, Caillou P, et al. Learning functional causal models with generative neural networks [EB/OL]. [2022-02-01]. <https://arxiv.org/abs/1709.05321>.
- [132] Athey S, Tibshirani J, Wager S. Generalized random forests [J]. *Ann Stat*, 2019, 47(2): 1148-1178. DOI: 10.1214/18-AOS1709.
- [133] Tan J, Xu S, Ge Y, et al. Counterfactual explainable recommendation[EB/OL]. [2022-02-01]. <https://arxiv.org/abs/2108.10539>.
- [134] Tharwat A. Classification assessment methods[J]. *Appl Comput Inform*, 2021, 17(1): 168-192. DOI: 10.1016/j.aci.2018.08.003.
- [135] Jung Y, Hu J. A K-fold averaging cross-validation procedure [J]. *J Nonparametr Stat*, 2015, 27(2): 167-179. DOI: 10.1080/10485252.2015.1010532.
- [136] Zhang Z, Chen L, Xu P, et al. Predictive analytics with ensemble modeling in laparoscopic surgery: a technical note[J]. *Laparosc Endosc Robot Surg*, 2022, 5(1): 25-34. DOI: 10.1016/j.lers.2021.12.003.
- [137] Zhang Z, Beck MW, Winkler DA, et al. Opening the black box of neural networks: methods for interpreting neural network models in clinical applications[J]. *Ann Transl Med*, 2018, 6(11):216. DOI: 10.21037/atm.2018.05.32.
- [138] Shimabukuro DW, Barton CW, Feldman MD, et al. Effect of a machine learning-based severe sepsis prediction algorithm on patient survival and hospital length of stay: a randomised clinical trial[J]. *BMJ Open Respir Res*, 2017, 4(1):e000234. DOI: 10.1136/bmjresp-2017-000234.
- [139] Semler MW, Weavind L, Hooper MH, et al. An electronic tool for the evaluation and treatment of sepsis in the ICU: a randomized controlled trial[J]. *Crit Care Med*, 2015, 43(8): 1595-1602. DOI: 10.1097/CCM.0000000000001020.
- [140] Wilson FP, Martin M, Yamamoto Y, et al. Electronic health record alerts for acute kidney injury: multicenter, randomized clinical trial[J]. *BMJ*, 2021, 372:m4786. DOI: 10.1136/bmj.m4786.
- [141] Selby NM, Casula A, Lamming L, et al. An organizational-level program of intervention for AKI: a pragmatic stepped wedge cluster randomized trial[J]. *J Am Soc Nephrol*, 2019, 30(3): 505-515. DOI: 10.1681/ASN.2018090886.
- [142] Wu Y, Chen Y, Li S, et al. Value of electronic alerts for acute kidney injury in high-risk wards: a pilot randomized controlled trial[J]. *Int Urol Nephrol*, 2018, 50(8): 1483-1488. DOI: 10.1007/s11255-018-1836-7.
- [143] Wilson FP, Shashaty M, Testani J, et al. Automated, electronic alerts for acute kidney injury: a single-blind, parallel-group, randomised controlled trial[J]. *Lancet*, 2015, 385(9981): 1966-1974. DOI: 10.1016/S0140-6736(15)60266-5.
- [144] Escobar GJ, Liu VX, Kipnis P. Automated identification of adults at risk for in-hospital clinical deterioration. reply [J]. *N Engl J Med*, 2021, 384(5): 486. DOI: 10.1056/NEJMc2034836.
- [145] 齐霜, 毛智, 胡新, 等. 基于专科信息系统建立的重症医学数据库:大型三甲医院重症医学数据库的模式[J]. *中华危重病急救医学* 2020, 32(6):743-749
- [146] 张素珍, 唐素娟, 戎珊, 等. 基于机器学习的重症监护病房脓毒性休克患者早期发生急性肾损伤风险的预测模型构建[J]. *中华危重病急救医学*, 2022, 34(3):255-259. DOI: 10.3760/cma.j.cn121430-20211126-01790.
- [147] Flechet M, Falini S, Bonetti C, et al. Machine learning versus physicians' prediction of acute kidney injury in critically ill adults: a prospective evaluation of the AKI predictor[J]. *Crit Care*, 2019, 23(1): 282. DOI: 10.1186/s13054-019-2563-x.
- [148] Sjoding MW, Hofer TP, Co I, et al. Interobserver reliability of the Berlin ARDS definition and strategies to improve the reliability of ARDS diagnosis[J]. *Chest*, 2018, 153(2): 361-367. DOI: 10.1016/j.chest.2017.11.037.
- [149] Fleuren LM, Thorat P, Shillan Det al. Machine learning in intensive care medicine: ready for take-off? [J]. *Intensive Care Med*, 2020, 46(7): 1486-1488. DOI: 10.1007/s00134-020-06045-y.
- [150] Komorowski M, Celi LA. Will artificial intelligence contribute to overuse in healthcare? [J]. *Crit Care Med*, 2017, 45(5): 912-913. DOI: 10.1097/ccm.0000000000002351.