

机器学习在新生儿呼吸窘迫综合征早期预警中的研究现状与展望

杨 兰, 郝莉霞

延安大学附属医院儿科, 陕西 延安

收稿日期: 2025年10月21日; 录用日期: 2025年11月14日; 发布日期: 2025年11月24日

摘 要

新生儿呼吸窘迫综合征(NRDS)是威胁早产儿生命健康的重大疾病, 传统诊断方法存在明显局限性。随着医疗信息化的发展, 电子病历系统积累了大量的临床数据, 机器学习技术为NRDS的早期预警提供了新的解决方案。本文系统分析了机器学习在NRDS早期预警中的研究价值及临床应用, 探讨了其在推动精准医学发展和探究发病机制方面的潜力。同时, 深入剖析了当前研究面临的数据质量、模型可解释性、泛化能力及临床整合等关键问题, 并为该领域的未来研究方向提供了建设性展望。

关键词

机器学习, 新生儿呼吸窘迫综合征, 早期预警, 人工智能, 医疗诊断

Research Status and Prospects of Machine Learning in Early Warning of Neonatal Respiratory Distress Syndrome

Lan Yang, Lixia Hao

Department of Pediatrics, Affiliated Hospital of Yan'an University, Yan'an Shaanxi

Received: October 21, 2025; accepted: November 14, 2025; published: November 24, 2025

Abstract

Neonatal Respiratory Distress Syndrome (NRDS) is a major life-threatening condition in preterm infants, with traditional diagnostic methods exhibiting significant limitations. With the advancement of medical informatization, electronic health record systems have accumulated vast amounts of clinical data, enabling machine learning technologies to offer novel solutions for early warning

文章引用: 杨兰, 郝莉霞. 机器学习在新生儿呼吸窘迫综合征早期预警中的研究现状与展望[J]. 临床医学进展, 2025, 15(11): 2226-2231. DOI: 10.12677/acm.2025.15113340

of NRDS. This paper systematically analyzes the research value and clinical applications of machine learning in NRDS early warning, exploring its potential in advancing precision medicine and elucidating disease mechanisms. It also delves into critical challenges facing current research, including data quality, model interpretability, generalization capability, and clinical integration, while offering constructive perspectives for future research directions in this field.

Keywords

Machine Learning, Neonatal Respiratory Distress Syndrome, Early Warning, Artificial Intelligence, Medical Diagnosis

Copyright © 2025 by author(s) and Hans Publishers Inc.

This work is licensed under the Creative Commons Attribution International License (CC BY 4.0).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

1. 引言

1.1. 研究背景

新生儿呼吸窘迫综合征(Neonatal Respiratory Distress Syndrome, NRDS)是早产儿常见且危及生命健康的疾病之一,其主要病因是肺表面活性物质缺乏,导致肺泡气体交换功能障碍。流行病学研究表明,NRDS在早产儿中发病率较高,尤其在胎龄小于 28 周的早产儿中患病风险显著增加[1]。在医疗资源匮乏的地区,NRDS 所致的新生儿死亡率可超过 40%。早产导致胎儿肺部发育不成熟,肺表面活性物质合成不足。此外,高胰岛素血症等因素也可能影响胎儿肺部的发育进程。

目前,NRDS 的诊断主要依赖临床表现观察和胸部 X 线检查,但症状不典型时易与其他呼吸系统疾病混淆,且 X 线检查常在病变典型时才发现,此时已错过最佳治疗窗口。随着医疗信息化的发展,电子病历系统积累了大量的临床数据,为基于机器学习的 NRDS 早期诊断提供了新的机遇[2]。

1.2. 研究意义

机器学习技术在临床中的应用,可整合多源数据(如生命体征监测、基因信息、影像学特征等),构建 NRDS 早期预警模型。在新生儿监护环境中,该类模型能够识别高危患儿,实现早期干预,从而改善预后、降低死亡率。此外,通过融合遗传学与临床特征数据,机器学习有助于动态追踪疾病进展,为揭示 NRDS 的发病机制提供新视角。在社会层面,NRDS 的长期治疗往往给家庭带来沉重的经济负担。而有效的早期预警能通过减少严重并发症、缩短住院时间,直接降低整体医疗成本,缓解家庭压力。机器学习还能基于个体差异(如遗传背景、生理参数)制定个性化诊疗方案,推动新生儿疾病管理向精准医学方向发展。

2. 机器学习在新生儿呼吸窘迫综合征早期预警中的国内外研究现状

2.1. 国外研究现状

近年来,机器学习技术在新生儿呼吸窘迫综合征(NRDS)的早期预警研究中取得了显著进展。国外学者从多角度探索机器学习在 NRDS 辅助诊断、风险预测和治疗优化中的应用,其研究路径主要围绕不同的数据来源展开。

在影像学分析方面,肺部超声因其无创、实时等优点,逐渐成为 NRDS 辅助诊断的重要工具。然而,

传统人工图像分析存在主观性强、标准不一等问题。Costa 等人(2024)应用深度学习算法对肺部超声图像进行自动分析,显著提高了 B 线增多等特征的识别精度。尽管其模型在内部数据集上表现出较高的分类准确率,但泛化能力仍需进一步验证[3]。

在遗传学层面, NRDS 的发病机制与特定基因突变的相关性也受到关注。Stone 与 Huynh(2024)报道了一例携带 ABCA3 基因 E292V 纯合突变的 NRDS 患儿,发现该基因突变导致肺表面活性物质代谢障碍[4]。研究者通过整合临床信息与基因组数据,利用机器学习模型识别遗传风险因素,为遗传相关性 NRDS 患儿的精准诊疗提供了新思路。

除了单一的影像和基因数据,利用常规收集的围产期临床数据构建预测模型是实现早期预警更直接的途径。Zeng 等人(2024)探讨了血小板计数与 NRDS 严重程度之间的复杂关系,发现机器学习算法能够有效捕捉其间的非线性交互作用,从而显著提升预测性能[5]。更进一步, Begimai K M(2025)与 Mousavi F S(2025)的研究均表明,应用梯度提升树等集成学习算法,仅利用产妇病史、新生儿基本信息等常规数据,即可高效预测早产儿 NRDS 风险,模型区分度(AUC)可达 0.9 以上,显著优于传统逻辑回归模型[6][7]。这表明,常规健康数据结合机器学习算法,有潜力辅助临床医生在出生后不久快速识别高危新生儿。

在生物标志物研究方面, Zhang 等人(2022)发现血浆中血管内皮生长因子(VEGF)和 apelin 水平在 NRDS 患儿中显著降低,基于这些标志物构建的支持向量机模型在独立验证集中表现出较高的特异性[8]。Chan 等人(2022)则结合 IL-27 和 IL-6 检测数据,构建了用于区分急性呼吸窘迫综合征(ARDS)和 NRDS 的分类模型,有效降低了临床表现相近的两种疾病的误诊率[9]。

此外,还有研究致力于构建更复杂的综合预测模型。一项发表于 2020 年的研究[10]利用极低出生体重早产儿的多维度数据,对比多种算法后提出一个 5 层深度神经网络模型。该模型经集成方法优化,在预测 NRDS 时表现出高灵敏度和特异性,并基于此开发了便于临床使用的网页应用,展示了机器学习向临床工具转化的早期尝试。

综上所述,国外研究在 NRDS 早期预警的模型构建上已取得显著进展,覆盖了从影像、基因到临床指标等多种数据类型。然而,这些研究的临床转化仍面临严峻挑战。首先,现有研究多局限于单一中心的小样本,缺乏大规模、多中心的验证,且研究方法各异,结论难以通用。其次,在技术层面,研究虽已呈现出利用多模态数据探索疾病机制的趋势,但诸如深度学习等复杂模型固有的“黑箱”特性,使得医护人员难以理解和信任其输出,这构成了模型落地临床的主要障碍。

2.2. 国内研究现状

国内学者在 NRDS 的机器学习研究方面也取得了重要进展。在数据资源构建方面,陈琦(2024)基于多中心 NICU 数据,收集了数千例 NRDS 患儿的临床资料,构建了预测支气管肺发育不良的模型[11]。该研究采用 Nomogram 方法,结合出生体重等变量进行风险评估,为后续研究提供了重要数据基础。

在影像特征量化方面,高敏团队(2022)通过研究大量肺部超声图像,验证了高频超声评分对 NRDS 预测及预后评估的价值[12]。在临床预测模型方面,张珂(2024)通过机器学习方法筛选出与俯卧位通气响应相关的临床参数[13]。这些参数被纳入模型训练中,显著提高了对治疗反应的预测能力。柴风云等人(2024)对雾化吸入治疗过程中的药物剂量等变量进行建模,为个性化治疗推荐系统的构建提供了理论基础[14]。在生物标志物研究方面,刘伟华(2024)发现某些血清蛋白在 NRDS 发病前数小时至数天即出现异常波动,提示其作为早期预警信号的潜力[15]。这些标志物已被纳入多个预测模型的特征集中,提升了模型对疾病进展的敏感性。

在区域适应性研究方面,韩同英等人(2022)发现高海拔地区 NRDS 患儿的临床特征与平原地区存在显著差异[16]。研究表明,未来构建预测模型时需考虑环境因素和地理特征,以增强模型的泛化能力。

在模型验证与临床转化方面, 丁冉等(2021)开发的早产儿预后评估工具在多中心验证中表现出性能差异, 提示模型的外部适应性仍需提升[17]。何明嫒等人(2021)指出, 机器学习模型需要与传统诊疗路径深度融合, 才能提高其临床实用性[18]。

国内研究在 NRDS 早期预警方面进展显著, 其突出特点是紧密结合临床实践, 在治疗优化和区域适应性研究上形成了优势。然而, 模型的临床转化仍受限于数据与验证的瓶颈, 例如, 陆俊秀(2019)与王美琪(2018)等研究先后指出, 跨中心数据标准不一会导致模型稳定性差[19], 而外部验证的不足则使其泛化能力受限[20]。这些挑战要求我们必须将未来的工作重心放在推动数据标准化、鼓励模型方法创新, 以及推行系统性的多中心临床验证, 从而提升研究成果的普适性与可靠性。

2.3. 国内外研究对比与展望

总体而言, 国外研究更注重机制探索和多中心验证, 在遗传学研究和生物标志物发现方面较为深入; 而国内研究则更侧重于临床数据建模和治疗策略优化, 尤其在区域适应性方面具有特色。然而, 国内外研究共同面临数据标准化不足、模型可解释性差和泛化能力有限等挑战。

未来研究应重点关注以下几个方向: 首先, 建立多中心、标准化的数据共享平台是基础。其次, 亟需开发可解释性强、泛化能力好的混合模型。再次, 推动机器学习模型与临床工作流程的深度融合是关键。最后, 加强跨学科合作与国际交流将共同推动该领域的发展。

3. 存在的问题

尽管机器学习在 NRDS 早期预警研究中展现出巨大潜力, 但其临床转化仍面临多重挑战。首先, 数据质量与标准化问题尤为突出: 研究表明, 电子病历数据常存在质量问题[21], 多模态数据(如生理信号、影像学资料和遗传信息)的整合与标准化程度不足, 加之 NRDS 病例在整体新生儿数据中占比有限, 导致样本不平衡问题[22], 直接影响模型的训练效果和预测准确性。对于临床医护人员而言, 深度学习模型的高精度优势常被其“黑箱”决策所掩盖, 导致难以建立信任。现有的模型解释方法, 如 LIME (局部可解释模型)和 SHAP (沙普利加和解释), 在应用于 NRDS 模型时各有优劣。LIME 通过局部窥探单个预测周围的模型行为来生成解释, 解释更直观迅速但可能因探查方式不同而导致结果不稳定; SHAP 则基于严谨的博弈论, 为每个特征计算出其对预测结果的精确贡献值, 解释更具一致性, 且能展现全局特征重要性, 但计算成本更高。在 NRDS 临床应用中, 前者可能更适用于床旁决策辅助, 后者则更适用于模型的评估与改进[23]。但它们在真实医疗环境下的解释是否可靠、是否易于理解并转化为临床行动, 仍是巨大挑战, 这严重限制了模型的落地推广。

此外, 模型泛化能力有限也是一个关键问题。这在对比国内外研究时尤为明显。例如, 国外研究多在拥有标准化诊疗规程的大型医学中心开展, 其模型所学的“NRDS 模式”高度依赖于当地人群特征和设备标准。而国内研究则揭示了显著的地区差异, 如韩同英等[16]指出高海拔地区 NRDS 患儿的临床特征与平原地区存在显著不同。一个基于平原地区数据训练的模型, 在应用到高海拔地区时, 可能因未能纳入“血氧饱和度”等关键的环境适应性特征而导致性能显著下降。这凸显了单一来源数据训练的模型难以适应不同地域、不同等级医院在患者群体、诊疗规范与资源条件上的差异。最后, 临床整合困难同样不容忽视。现有预警系统与医院信息系统(HIS)和电子病历(EMR)系统的集成度较低, 医护人员需要手动录入数据或切换不同平台, 增加了工作负担; 同时, 模型输出的风险概率与具体临床干预措施之间缺乏明确的对应指南, 导致预警信息难以转化为实际的诊疗行动。

4. 总结与展望

机器学习在 NRDS 早期预警中的应用具有重要的临床价值, 能够通过多源数据整合和智能分析实现

高危患儿的早期识别, 为及时干预和改善预后提供技术支持。然而, 要实现其全临床应用, 仍需从数据、技术、临床整合及伦理规范等多个层面进行完善。

在数据层面, 应建立多中心数据共享平台, 制定统一的数据采集和标注标准, 并通过生成对抗网络 (GAN) 等数据增强技术解决样本不平衡问题[24], 同时加强对多模态数据(包括临床指标、影像学特征、遗传信息和环境因素等)的融合利用。在技术层面, 未来的模型开发应更好地贴合临床实际。一个重要的方向是构建多阶段动态预警模型。例如, 可以设计一个系统: 首先, 在产前利用产妇病史、胎龄等数据完成首次风险评估; 接着, 在出生后立即纳入生命体征、Apgar 评分等指标进行动态更新; 最后, 在入院初期融合肺部超声影像等特征, 实现风险的精准分层与早期预警。这种分阶段的动态预测模式更符合临床诊疗的时间线, 也能为医生提供连续性的决策支持。为实现上述多中心数据的有效利用, 联邦学习提供了一条可行的技术路径[25] [26], 其具体设想是: 由一个协调中心向多家合作的医院下发统一的初始模型; 各医院在本地服务器上使用自己的数据训练模型, 只将模型参数的更新加密后传回中心进行聚合, 生成一个更强健的全局模型后再下发给各医院。该技术既能保护患者隐私和安全, 又能利用更多数据来提升模型的泛化能力, 为打破数据孤岛、实现有效的跨医院合作提供了新的解决方案。然而, 实施路径上也面临真实挑战, 例如, 各家医院数据标准不一会影响模型效果, 网络通信和安全保障成本较高, 还需要建立跨机构的协作与信任机制。此外, 应探索多任务学习框架, 同时预测 NRDS 风险、严重程度和最佳治疗方案。

临床整合层面需要推动预警系统与现有医疗信息系统的深度融合, 实现数据的自动提取和风险预警的实时推送, 并建立模型输出与临床干预措施的明确对应关系, 制定基于机器学习预测结果的临床决策指南。同时, 应加强对医护人员的培训, 提高其对人工智能辅助决策的接受度和使用能力。伦理与规范层面则需建立完善的质量评估体系和伦理审查机制, 确保模型的可靠性、安全性和公平性, 特别要关注数据隐私保护和算法偏差问题, 避免因数据或算法问题导致的不公平决策。

展望未来, 随着技术的不断发展和临床应用的深入, 机器学习有望在 NRDS 早期预警和个性化治疗中发挥更为重要的作用。通过多学科合作和国际交流, 共同推动数据标准化、算法创新和临床验证工作, 最终实现新生儿呼吸系统疾病的精准防控和优化管理, 为提升新生儿健康水平提供有力支持。

参考文献

- [1] Stoll, B.J., Hansen, N.I., Bell, E.F., Walsh, M.C., Carlo, W.A., Shankaran, S., *et al.* (2015) Trends in Care Practices, Morbidity, and Mortality of Extremely Preterm Neonates, 1993-2012. *JAMA*, **314**, 1039-1051. <https://doi.org/10.1001/jama.2015.10244>
- [2] Rumsfeld, J.S., Joynt, K.E. and Maddox, T.M. (2016) Big Data Analytics to Improve Cardiovascular Care: Promise and Challenges. *Nature Reviews Cardiology*, **13**, 350-359. <https://doi.org/10.1038/nrcardio.2016.42>
- [3] Costa, F., Titolo, A., Ferrocino, M., Biagi, E., Dell'Orto, V., Perrone, S., *et al.* (2024) Lung Ultrasound in Neonatal Respiratory Distress Syndrome: A Narrative Review of the Last 10 Years. *Diagnostics*, **14**, Article No. 2793. <https://doi.org/10.3390/diagnostics14242793>
- [4] Stone, A. and Huynh, T. (2024) Neonatal Respiratory Distress Syndrome in E292V Homozygous ABCA3. *BMJ Case Reports*, **17**, e261347. <https://doi.org/10.1136/bcr-2024-261347>
- [5] Zeng, Y. (2024) Relationship between Platelet Count and Severity of Neonatal Respiratory Distress Syndrome. *Italian Journal of Pediatrics*, **50**, 208.
- [6] Kyzy Begimai, M., Mounika, S., Manoj, J., Priyanka, S. and Das, A. (2025) Challenges of Neonatal Respiratory Distress Syndrome. *International Journal of Advanced Research*, **13**, 245-250. <https://doi.org/10.21474/ijar01/20557>
- [7] Mousavi, S.F., Bakht, R., Tapak, L., Goli, M.A. and Refaei, M. (2025) Supporting Mothers of Discharged Premature Infants: A Clinical Trial on Tele Counseling for Self-Efficacy and Anxiety Management. *BMC Pregnancy and Childbirth*, **25**, Article No. 794. <https://doi.org/10.1186/s12884-025-07910-8>
- [8] Zhang, Y., Shao, S., Mu, Q., Feng, J., Liu, J., Zeng, C., *et al.* (2022) Plasma Apelin and Vascular Endothelial Growth

- Factor Levels in Preterm Infants: Relationship to Neonatal Respiratory Distress Syndrome. *The Journal of Maternal-Fetal & Neonatal Medicine*, 35, 10064-10071. <https://doi.org/10.1080/14767058.2022.2089554>
- [9] Liu, C., He, Y., Ai, Q., et al. (2022) A Pilot Study of Plasma Interleukin-6 and Interleukin-27 in Differential Diagnosis of Acute Respiratory Distress Syndrome and Neonatal Respiratory Distress Syndrome in Preterm Infants. *Chinese Journal of Contemporary Pediatrics*, 24, 428-432.
 - [10] Sidey-Gibbons, J.A.M. and Sidey-Gibbons, C.J. (2019) Machine Learning in Medicine: A Practical Introduction. *BMC Medical Research Methodology*, 19, Article No. 64. <https://doi.org/10.1186/s12874-019-0681-4>
 - [11] 陈琦, 徐仕霞. 床旁肺超声评分与新生儿呼吸窘迫综合征患儿病情严重程度的关系[J]. 中国超声医学杂志, 2025, 41(5): 505-508.
 - [12] 高敏, 张楠楠, 荣辉, 等. 肺部高频超声评分对新生儿呼吸窘迫综合征预测及预后评估的价值[J]. 中国现代医学杂志, 2022, 32(17): 81-87.
 - [13] 张珂, 饶兴愉. 新生儿呼吸窘迫综合征患儿发生支气管肺发育不良的 Nomogram 预测模型的建立与评估[J]. 重庆医科大学学报, 2024, 49(10): 1110-1118.
 - [14] 柴凤云, 全实, 韩梅, 等. 俯卧位在新生儿呼吸窘迫综合征有创呼吸支持中的临床研究[J]. 中国当代儿科杂志, 2024, 26(6): 619-624.
 - [15] 刘伟华, 解谦, 周雅玲. 新生儿呼吸窘迫综合征雾化吸入疗法的定量分析[J]. 哈尔滨理工大学学报, 2024, 29(1): 124-132.
 - [16] 韩同英, 叶琼波, 德吉玉珍, 等. 高海拔地区新生儿呼吸窘迫综合征初始呼吸支持策略的影响因素及早期结局分析[J]. 中国全科医学, 2022, 25(27): 3384-3389.
 - [17] 丁冉, 陈强, 张倩薇, 等. 不同分期组织学绒毛膜羊膜炎与胎龄小于 32 周早产儿呼吸窘迫综合征关系的研究[J]. 中国当代儿科杂志, 2021, 23(3): 248-253.
 - [18] 何明媛. 早产儿呼吸窘迫综合征高频振荡通气后两种撤机方式的安全性研究: 前瞻性随机病例对照试验[D]: [硕士学位论文]. 厦门: 厦门大学, 2021.
 - [19] 陆俊秀, 赖春华, 杨冰岩, 等. 血清氨基末端脑钠肽前体及白细胞介素-6 水平在新生儿呼吸窘迫综合征早期诊断和严重程度评估中的应用[J]. 中国医学科学院学报, 2019, 41(1): 80-85.
 - [20] 王美琪, 梅花, 刘春枝. 新生儿呼吸窘迫综合征及支气管肺发育不良防治进展[J]. 临床儿科杂志, 2018, 36(9): 702-706.
 - [21] Kahn, M.G., Callahan, T.J., Barnard, J., Bauck, A.E., Brown, J., Davidson, B.N., et al. (2016) A Harmonized Data Quality Assessment Terminology and Framework for the Secondary Use of Electronic Health Record Data. *eGEMs*, 4, Article No. 1244. <https://doi.org/10.13063/2327-9214.1244>
 - [22] Johnson, J.M. and Khoshgoftaar, T.M. (2019) Survey on Deep Learning with Class Imbalance. *Journal of Big Data*, 6, Article No. 27. <https://doi.org/10.1186/s40537-019-0192-5>
 - [23] Lundberg, S.M. and Lee, S.I. (2017) A Unified Approach to Interpreting Model Predictions. *Annual Conference on Neural Information Processing Systems*, Long Beach, 4-9 December 2017, 4768-4777.
 - [24] Frid-Adar, M., Diamant, I., Klang, E., Amitai, M., Goldberger, J. and Greenspan, H. (2018) GAN-Based Synthetic Medical Image Augmentation for Increased CNN Performance in Liver Lesion Classification. *Neurocomputing*, 321, 321-331. <https://doi.org/10.1016/j.neucom.2018.09.013>
 - [25] Rieke, N., Hancox, J., Li, W., Milletari, F., Roth, H.R., Albarqouni, S., et al. (2020) The Future of Digital Health with Federated Learning. *NPJ Digital Medicine*, 3, Article No. 119. <https://doi.org/10.1038/s41746-020-00323-1>
 - [26] Sheller, M.J., Edwards, B., Reina, G.A., Martin, J., Pati, S., Kotrotsou, A., et al. (2020) Federated Learning in Medicine: Facilitating Multi-Institutional Collaborations without Sharing Patient Data. *Scientific Reports*, 10, Article No. 12598. <https://doi.org/10.1038/s41598-020-69250-1>