

# MIMIC数据库的多维透视：发展历程、数据概况、分析方法及全球影响综述



王晓莉, 毛 智

中国人民解放军总医院第一医学中心重症医学科 (北京 100853)

**【摘要】**在医疗信息化进程不断推进的背景下,麻省理工学院联合以色列女执事医疗中心共同开发了重症监护医疗信息(Medical Information Mart for Intensive Care, MIMIC)数据库,旨在整合重症监护病房临床数据,促进医疗研究与跨学科合作。MIMIC数据库已历经多个版本迭代,各版本在数据量、时间范围、来源等方面持续优化,其包含多种数据结构和类型,支持基础统计、统计推断、机器学习及多模态整合等多样化分析方法。MIMIC数据库具备真实可靠的数据来源和丰富完整的内容,但仍存在数据缺失、时间精度不一致、系统变更兼容性及潜在偏倚等问题,需谨慎对待。其全球影响力体现在医学研究、医疗改进、国际合作、信息化发展及重症医学进步等多方面。未来,数据隐私保护和质量控制仍是MIMIC数据库重要的发展方向。本文通过梳理MIMIC数据库的发展历程与版本迭代,分析其数据质量和潜在偏倚,探讨其全球影响力,为科研人员快速了解MIMIC数据库提供参考。

**【关键词】**MIMIC数据库;发展历程;统计学方法;数据质量;偏倚;全球影响

**【中图分类号】**R 459.7 **【文献标识码】**A

A multidimensional perspective on MIMIC database: a review of development history, data overview, analytical methods, and global influence

WANG Xiaoli, MAO Zhi

Department of Critical Care Medicine, The First Medical Center, Chinese PLA General Hospital, Beijing 100853, China

Corresponding author: MAO Zhi, Email: maozhi@aliyun.com

**【Abstract】**In the context of the continuous progress of medical electronic informatization, Massachusetts Institute of Technology and Beth Israel Deaconess Medical Center jointly developed the Medical Information Mart for Intensive Care (MIMIC) database aiming to integrate intensive care unit data and promote medical research and interdisciplinary cooperation. The MIMIC database has evolved through several versions, and each version has been continuously improved in terms of data volume, time range and sources, etc. It consists of multi-level structure and diverse data types supporting various analytical methods, such as basic statistics, inference, machine learning, and multimodal integration. The MIMIC database offers reliable data sources and comprehensive content, but it still faces issues such as missing data, time precision inconsistencies, compatibility with system changes, and potential biases, which require careful handling. The global influence of the MIMIC database is manifested in multiple aspects including medical research,

DOI: [10.12173/j.issn.1004-4337.202504023](https://doi.org/10.12173/j.issn.1004-4337.202504023)

通信作者: 毛智, 博士, 副主任医师, Email: maozhi@aliyun.com

<https://slyyx.whuznhmedj.com/>

healthcare improvement, international collaboration, informatization development, and advancements in critical care medicine. In the future, data privacy protection and quality control will remain important development directions for the MIMIC database. This paper reviews the development and version iterations of the MIMIC database, analyzes its data quality and potential biases, and explores its global influence, providing a reference for researchers to quickly understand the MIMIC database.

**【Keywords】** MIMIC database; Development history; Statistical methods; Data quality; Bias; Global influence

随着医疗电子信息化的快速发展, 医疗数据的收集、管理和分析利用日趋重要。医疗数据不仅可以用于临床决策支持, 还可以促进医学研究发展、提升医疗质量和效率。其中, 重症监护医疗信息 (Medical Information Mart for Intensive Care, MIMIC) 数据库是这一领域的标志性项目之一, 它为研究人员提供了丰富的临床数据资料, 极大地推动了医学研究的发展。

开展 MIMIC 数据库的相关研究具有重要意义。MIMIC 数据库由麻省理工学院 (Massachusetts Institute of Technology, MIT) 与贝斯以色列女执事医疗中心 (Beth Israel Deaconess Medical Center, BIDMC) 联合开发, 各版本持续改进, 系统梳理其发展历程及其在整合重症监护病房 (intensive care unit, ICU) 数据、推动跨学科合作等方面的作用, 可以为研究者提供更清晰的认知框架。另外, 分析其结构化特性、多样数据类型及独特访问机制有助于医疗从业者更好地利用这些资源, 挖掘其科研价值。本文旨在通过梳理其发展历程, 系统介绍该数据库的结构与特点, 评估其数据质量与潜在偏倚, 并探讨其在全球医疗信息化中的影响与示范作用, 帮助国内研究者深入了解该数据库, 同时为尚未接触该领域的学者快速建立认知框架, 为其开展后续研究提供参考与支持。

## 1 MIMIC数据库发展历程

### 1.1 项目起源

20 世纪 90 年代末期, 随着信息技术的飞速发展, 医疗数据的收集和管理逐渐成为医疗领域的研究热点。ICU 数据库长期以来都是探究危重疾病风险因素、自然病史及不同治疗策略效果的重要基础。但受限于数据质量及技术条件, 以及对患者隐私保护的必要考量, 这些因素对医学研究的发展产生了一定制约。

面对这一挑战, MIT 研究团队意识到, 集中管理高质量的医疗数据对于推动医学研究和临床决策支持具有重要价值。为此, 他们联合 BIDMC 开发了一个专门的重症监护数据库, 即 MIMIC 数据库。该项目的目标为提供免费开放的数据库, 以支持对重症监护患者的深入研究和分析, 促进医学研究和临床决策支持的发展。

### 1.2 早期发展及版本更新

在项目初期, 研究团队面临诸多挑战, 从不同医院信息系统中提取数据是首要难题, 不同系统的数据格式、存储方式千差万别。研究团队需要开发专门的提取、转换和加载工具, 以确保数据能够准确、完整地迁移到 MIMIC 数据库中。同时, 伦理问题是另一个重大挑战, 患者数据涉及个人隐私, 必须在合法合规的前提下使用。因此, 项目团队与伦理委员会合作, 制定了严格的数据脱敏方案和访问权限控制机制, 确保患者信息的保密性。MIMIC 数据库构建示意图见图 1。

自 MIMIC 数据库建立以来, 已经发布了多个版本, 包括 MIMIC-II、MIMIC-III 和 MIMIC-IV。每个新版本都在前代版本基础上进行了数据更新和扩展, 为研究者提供了更全面、更丰富的数据集。2010 年最早发布的 MIMIC-II 是 MIMIC 数据库项目的一个重要里程碑, 数据主要从 CareVue 监视器获取 BIDMC 患者的资料信息, 收集时间为 2001—2007 年, 涵盖约 26 000 名患者的临床数据。每条记录包含了详细的临床信息, 包括实验室数据、治疗干预情况、护理记录、出院小结、放射学报告、医嘱单、ICD-9 编码, 以及部分患者的高分辨率生命体征趋势和波形数据, 数据在收集时自动进行了去标识化处理<sup>[1]</sup>。尽管该版本为研究人员提供了更多的研究机会, 但仍存在数据陈旧、隐私保护机制不健全等问题, 目前已不再公开获取使用。

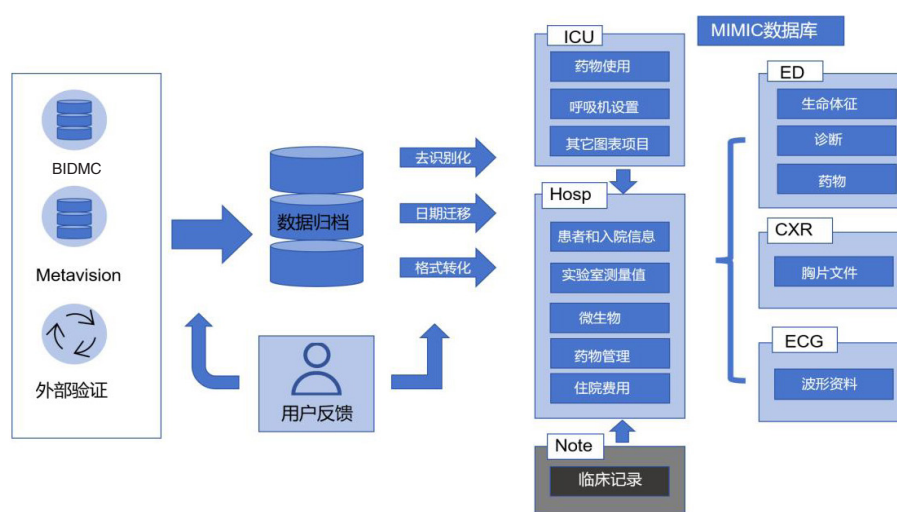


图1 MIMIC数据库构建示意图

Figure 1. Schematic diagram of the construction of the MIMIC database

注：图中MIMIC数据库的几个模块中，蓝色背景表示该模块资源目前可获取，灰色背景表示该模块目前尚未公开。

MIMIC-III<sup>[2]</sup> 的发布标志着 MIMIC 数据库的发展进入了一个新阶段。MIMIC-III 包含了 2001—2012 年的数据，收集了 Metavision 和 CareVue 两个床旁监视器的数据，数据规模进一步扩大，涵盖超过 40 000 名患者的数据，数据时间段也进一步延长。经过严格的质量控制和验证，MIMIC-III 的数据质量进一步提升，成为研究人员的重要资源，许多重要的研究成果都是基于这一版本的数据而来。目前 MIMIC-III 最新版本为 v1.4，该版本于 2016 年发布，也是其最终版本。

2020 年 MIMIC-IV 首次发布，该版本在 MIMIC-III 的基础上进行了多项改进，包括数据更新、部分表格重构及新增医学影像数据等。MIMIC-IV 数据时间段进一步延长，收集了 2008—2022 年的数据资料。相较于 MIMIC-III，MIMIC-IV 数据来源更加广泛，不仅包括 BIDMC 的数据，还整合了其他多家医院的 ICU 数据，数据量达到了数十万条。MIMIC-IV 进一步扩展了数据的广度和深度，在多个方面对日益增多的可公开获取的重症监护数据集进行了补充；还纳入了新的精确数字信息来源，如电子用药记录；此外，MIMIC-IV 构建了组成数据的模块化组织，使得该数据库能够与外部科室及不同类型的数据相链接<sup>[3]</sup>。这一版本的数据为跨学科研究提供了更多支持。MIMIC-IV 的数据结构复杂化使得研究人员可以从多个角度和层面研究患者的病情和治疗过程。

## 2 MIMIC数据库概况

### 2.1 数据库结构

MIMIC 数据库是一个结构化的医疗信息数据库，包含了多个表格。这些表格通过严格的集成和标准化过程形成，确保了数据的准确性和一致性。其主要表格包括患者信息表、入院记录表、诊断信息表、治疗信息表、实验室结果表等。各表格都包含了丰富的数据字段和变量，为研究者提供了更多样的数据分析可能。

### 2.2 数据类型与特点

MIMIC 数据库涵盖多模态数据体系，主要包括时间序列数据、非结构化文本数据及医学影像数据等<sup>[3]</sup>。具体特征如下：①时间序列数据。源自患者的生理参数和治疗过程中的监测数据（如心率、血压、血氧饱和度、机械通气参数等），以时间戳记录患者生理状态及治疗干预的动态演变过程，支持基于时序依赖关系的临床状态分析与预测建模。②非结构化文本数据。包含电子病历中的病程记录、医嘱文本、护理评估、出院小结等自然语言数据，可以通过自然语言处理技术提取隐含的临床语义，为疾病表型分析及诊疗路径挖掘提供文本证据。③医学影像数据。部分版本（如 MIMIC-IV）纳入胸部 X 射线等影像数据，虽需结合计算机视觉算法进行量化分析，但可直观呈现解剖结构特征，辅助验证临床诊断（如肺部感染、气胸等）。

## 2.3 数据访问与使用

MIMIC 数据库是一个开放访问的医疗信息数据库，但为限制访问流量，用户访问前需签署数据使用协议，并完成相应培训<sup>[1]</sup>。研究者可以通过 <https://physionet.org/> 网站申请访问权限。申请过程中需要提交详细的申请表格和研究计划，并接受审查以确保数据的合理使用。获得访问权限后，研究者就可以使用 MIMIC 数据库所提供的多种工具和资源来开展数据分析工作。

## 3 MIMIC数据库相关统计学方法

MIMIC 数据库的多层次数据结构为多元化分析方法提供了有力支撑。研究者可以根据数据模态特征，选择适配的统计与算法工具，构建“数据特征 - 分析目标 - 方法选择”的系统性分析范式。

### 3.1 基础统计

针对数据库中的结构化数据（如人口学特征、实验室检查结果等），描述性统计可作为初始分析步骤，量化呈现队列特征。例如，采用描述性统计分析 ICU 患者的年龄、生命体征变化，从而识别出可能影响预后的关键变量，并通过可视化工具进一步探索这些变量之间的关系<sup>[4-5]</sup>。

### 3.2 统计推断

基于数据库中关联表的结构化设计，推断性统计学方法可用于分析变量间的相互关系。例如，对两组患者的连续变量（如 APACHE II 评分）采用独立样本  $t$  检验，对分类变量（如机械通气与否）采用  $\chi^2$  检验，探究治疗措施与预后结局的关联性；构建 Cox 比例风险模型，纳入相关变量，评估多因素对患者生存时间的独立影响<sup>[6]</sup>。

### 3.3 机器学习

MIMIC 数据库提供了大量高频时间序列数据，使得分析方法从传统模式逐步向机器学习与深度学习转变。例如，时序数据建模：利用长短期记忆网络（long short-term memory, LSTM）处理连续 24 小时的 APACHE II、SOFA 和 SAPS II 评分系统的时间序列变量，开发了基于 LSTM 的医院 ICU 死亡率预测模型，显著提高 ICU 死亡率预测的准确性<sup>[7]</sup>；无监督学习应用：对患者的人口统计（如年龄等）、实验室指标（如白细胞计数等）进行聚类分析，识别创伤性脑损伤的不同表型，为精准治疗提供依据<sup>[8]</sup>。

## 3.4 多模态整合

随着深度学习技术的快速发展，多模态数据融合与分析技术在基于 MIMIC 数据库的科研工作中愈发重要，正逐步占据该领域的主导地位。例如，通过卷积神经网络对 MIMIC-IV 的胸部 X 射线影像进行特征编码，结合临床数据构建多模态模型，可准确预测患者的住院时间和死亡率<sup>[9]</sup>。

总之，MIMIC 数据库的多层次数据结构和丰富的数据类型为研究者提供了多样化的分析工具和方法支持。从传统的统计分析到机器学习再到跨模态整合，研究者可以根据具体需求选择合适的统计分析方法，从而实现从数据到科研实践的转化。

## 4 MIMIC数据库的数据质量

MIMIC 数据库在数据质量方面具有多面性，首先，其数据来源真实可靠、数据内容丰富且经过整理，为研究提供了极大便利，但同时也存在数据缺失、时间信息处理特殊，及系统变更带来的兼容性问题，在使用时需要谨慎对待并加以处理。

数据质量优势<sup>[10]</sup>：①数据来源与整理可靠。MIMIC 数据来源于 BIDMC 的真实医疗记录，由 MIT 计算生理学实验室进行整理，去除了个人可识别信息，保证数据适合外部研究使用。其遵循美国健康保险流通与责任法案（Health Insurance Portability and Accountability Act, HIPAA）建议进行匿名化处理，虽对部分数据分析有影响，但保留了时间相关信息，为后续分析提供了基础。②数据丰富性与完整性较好。包含数万名患者的医疗记录，时间跨度达几十年，数据类型多样，如生命体征、用药、检验结果、护理记录等，以及患者的行政活动数据，能全面反映患者的治疗过程。数据库中含时间戳信息，可用于构建事件日志，且部分关键表（如 admissions、icustays）的核心标识（如 subject\_id、hadm\_id、icustay\_id）相对完整，为科学研究提供了必要的数据库。

数据质量存在的问题<sup>[10]</sup>：①部分数据缺失。部分表格存在数据缺失的情况，如 admissions 表中部分时间戳缺失，labevents 表中有 hadm\_id 缺失等，不同表格缺失的元素和比例有所不同，可能影响数据分析的完整性和准确性。②时间信息问题。时间戳存在多种记录方式，部分表格时间精

度不一致,在进行跨表分析时容易导致时间顺序错误。MIMIC 通过日期偏移实现匿名化,将所有真实日期移至未来,这使得一些与真实日期相关的分析无法进行。③数据一致性问题。不同表格间存在数据不一致的情况,在 ICD 编码方面, `d_icd_diagnoses` 表有缺失编码,可能影响数据的准确性和分析结果的可靠性。④系统变更带来的问题。2008 年医院电子健康记录系统从 Philips CareVue (CV) 更换为 iMDSoft MetaVision (MV),两个系统的数据格式和 item id 存在差异,尽管可以通过特定方法标记和区分,但仍给数据的处理和分析带来挑战,影响了流程挖掘结果和数据质量评估。

## 5 MIMIC数据的潜在偏倚

MIMIC 数据库虽为医学研究提供了丰富的数据,但也存在多种偏倚。首先是选择偏倚,由于数据源于特定医院 ICU 患者,无法代表所有患者群体,不同地区、不同级别医院患者的疾病特征、治疗方式等存在差异,这就导致样本缺乏广泛代表性,从而影响研究结果的外推;在抽样偏倚上,数据收集可能倾向特定时间段、科室或病情的患者,部分病情极端患者未被充分纳入,使样本难以代表目标总体,从而降低了统计模型的有效性与可泛化性<sup>[11]</sup>;测量偏倚也较为突出,ICU 患者病情复杂,数据采集可能不及时、不全面,且人工录入易出错,医护人员记录标准不一致,致使数据准确性受损,进而影响对患者病情的评估及研究的可靠性<sup>[12]</sup>。此外,由于数据库存在时间跨度,医疗技术、诊断标准等随时间变化,早期数据无法反映当前医疗实践,若研究未合理调整时间因素,也会导致偏差。这些偏倚在不同程度上对研究的准确性与可靠性产生影响,研究人员应充分认识并采取相应措施加以控制,以提升研究质量。

## 6 MIMIC数据库的全球影响

### 6.1 推动医学研究与发展

MIMIC 数据库的开发和开放访问为医学研究提供了丰富的数据资源,研究者可以利用 MIMIC 数据库开展医学研究,以推动医学知识的进步和临床实践的改进。近年来,越来越多的研究者开始利用 MIMIC 数据库开展数据分析和研究。这些研究主要涉及但不限于重症监护治疗效果评估、疾病预后预测等医学领域。如在重症监护治

疗效果评估方面, Hu 等利用 MIMIC 数据库评估重症心房颤动患者的甘油三酯-葡萄糖体重指数 (TyG-BMI) 与全因死亡率之间的相关性,结果发现, TyG-BMI 水平较低时,患者在 30 天、90 天、180 天及 365 天的全因死亡率会显著升高,表明 TyG-BMI 指数可作为重症心房颤动患者分级和治疗的有效指标<sup>[13]</sup>。在疾病预后预测方面,有研究团队利用 MIMIC 数据库预测 ICU 患者的住院和再住院风险,该研究基于机器学习算法构建预测模型,结果显示,该模型对需要再入院 ICU 的脓毒症患者显示出良好的预后预测能力,为临床决策提供了有力支持<sup>[14]</sup>。

### 6.2 助力数据驱动的医疗改进

MIMIC 数据库不仅为医学研究提供了数据资源,还为数据驱动的医疗质量改进提供了可能。通过分析 MIMIC 数据库,研究者可以发现医疗实践中的问题和不足,并提出改进措施和建议。有研究基于 MIMIC 数据库分析发现白蛋白与晶体液给药的早期联合可能有益于脓毒症患者的生存<sup>[15]</sup>;早期预防性使用肝素可能显著降低急性呼吸窘迫综合征患者的死亡率<sup>[16]</sup>等。这些研究结果为临床用药提供了重要的参考,有助于优化药物治疗方案并改善患者预后。还有研究利用 MIMIC 数据库评估医疗质量和安全性,通过分析患者的并发症和不良事件等数据,发现医疗实践中的潜在风险和问题,并提出改进措施和建议,以提升医疗质量和安全性<sup>[17-18]</sup>。

### 6.3 促进国际合作与交流

MIMIC 数据库的开放共享还促进了国际合作与交流。通过国际合作与交流,研究者可以共享数据资源和分析工具,共同推动科学研究的发展。同时,国际合作还可以促进不同国家和地区之间的医疗经验和技术交流,有助于提升全球医疗水平和服务质量。例如,一个由中国与美国学者组成的研究团队基于 MIMIC 数据库探究了术后早期给予对乙酰氨基酚与心脏术后发生严重急性肾损伤之间的关系,其研究成果不仅为临床决策提供了有力依据,还极大地促进了国际间的合作与交流<sup>[19]</sup>。

### 6.4 引领医学信息化与智能化发展

MIMIC 数据库还引领了医学信息化与智能化发展。随着大数据和人工智能技术的不断发展,越来越多的研究者开始利用这些技术来分析和处

理研究数据。MIMIC 数据库作为一个典型的开放访问医疗数据库，为大数据和人工智能技术在医学领域的应用提供了丰富的数据资源和研究平台。利用 MIMIC 数据库，研究者可以训练和验证各种机器学习算法和模型，以实现对医学数据的智能化分析和处理。有研究利用 MIMIC 数据库训练心电图深度学习（ECG-AI）模型来预测患者新发心房颤动，结果发现，ECG-AI 提供了一种准确的工具预测新发心房颤动，且优于临床和多基因评分<sup>[20]</sup>。

### 6.5 推动重症医学的进步

MIMIC 数据库对推进重症医学发展有重大意义。在深化疾病认知上，它整合了海量多模态重症患者临床数据，助力科研人员从多层次探究重症疾病机制，完善知识体系，如为脓毒症分型及早期预警提供新思路<sup>[21]</sup>。在临床诊疗方面，支持开发的智能决策系统能实时分析患者数据，预测并发症及预后，为医生提供个性化治疗建议，如对 ARDS 患者的死亡风险进行预测，从而进行早期干预，提高生存率<sup>[22]</sup>。在人才培养方面，其为医学教育提供了丰富且真实的教学资源，有助于培养医学生的临床思维与实践能力，为学科发展提供了人才保障<sup>[23]</sup>。

## 7 小结

MIMIC 数据库的发展历程反映了医疗数据从分散到集中整合的趋势。其诞生背景与医疗行业对数据驱动决策的需求及信息技术的发展紧密相关。如今，MIMIC 数据库在全球范围内展现出了巨大的影响力，从促进国际研究合作、变革医疗实践到推动医学教育进步，已成为现代医学发展中不可或缺的一部分。未来，随着数据处理技术的不断发展与应用的日益广泛，MIMIC 数据库将继续发挥重要作用，为全球范围内的医学研究与临床实践提供更为详尽且深入的数据支撑。

MIMIC 数据库未来的发展还需重点关注数据隐私保护、质量控制、拓展应用及国际协作与标准化等方面。升级加密技术，加强隐私保护机制，并强化基于角色的访问控制，以确保患者数据安全；建立自动化监测系统，优化清洗流程，引入外部验证，以提高数据的可靠性与一致性；在拓展应用领域方面，推动精准医学研究、医疗物联网设备数据融合，并促进人工智能创新；继续加

强国际协作与标准化，推动跨机构的数据共享，促进通用数据标准的采用。这些举措将进一步提升 MIMIC 数据库的科研价值与临床转化潜力，为全球医学发展提供更强大的支持。

### 参考文献

- 1 Lee J, Scott DJ, Villarroel M, et al. Open-access MIMIC-II database for intensive care research[J]. *Annu Int Conf IEEE Eng Med Biol Soc*, 2011, 2011: 8315–8318. DOI: [10.1109/IEMBS.2011.6092050](https://doi.org/10.1109/IEMBS.2011.6092050).
- 2 Johnson AE, Pollard TJ, Shen L, et al. MIMIC-III, a freely accessible critical care database[J]. *Sci Data*, 2016, 3: 160035. DOI: [10.1038/sdata.2016.35](https://doi.org/10.1038/sdata.2016.35).
- 3 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](https://doi.org/10.1038/s41597-022-01899-x).
- 4 Shi Y, Shi Y, Liu Y, et al. Association between neutrophil percentage to serum albumin ratio and in-hospital mortality of patients with chronic obstructive pulmonary disease in intensive care unit: a retrospective cohort study[J]. *Int J Chron Obstruct Pulmon Dis*, 2025, 20: 1227–1237. DOI: [10.2147/COPD.S508964](https://doi.org/10.2147/COPD.S508964).
- 5 范勇, 赵宇卓, 李沛尧, 等. 危急重症监护数据库 MIMIC-III 疾病谱分析 [J]. *中华危重病急救医学*, 2018, 30(6): 531–537. [Fan Y, Zhao YZ, Li PY, et al. Analysis of diseases distribution in Medical Information Mart for Intensive Care III database[J]. *Chinese Critical Care Medicine*, 2018, 30(6): 531–537.] DOI: [10.3760/cma.j.issn.2095-4352.2018.06.006](https://doi.org/10.3760/cma.j.issn.2095-4352.2018.06.006).
- 6 Zhou Y, Hu J. The association between pan-immune-inflammation value with mortality in critically ill patients with sepsis-associated acute kidney injury[J]. *BMC Infect Dis*, 2025, 25(1): 486. DOI: [10.1186/s12879-025-10880-z](https://doi.org/10.1186/s12879-025-10880-z).
- 7 Deng Y, Li S, Li J, et al. Enhancing mortality prediction in intensive care units: improving APACHE II, SOFA, and SAPS II scoring systems using long short-term memory[J]. *Intern Emerg Med*, 2025. DOI: [10.1007/s11739-025-03896-5](https://doi.org/10.1007/s11739-025-03896-5).
- 8 Ghaderi H, Foreman B, Reddy CK, et al. Discovery of generalizable TBI phenotypes using multivariate time-series clustering[J]. *Comput Biol Med*, 2024, 180: 108997. DOI: [10.1016/j.combiomed.2024.108997](https://doi.org/10.1016/j.combiomed.2024.108997).
- 9 Chen J, Li Q, Liu F, et al. M3T-LM: a multi-modal multi-task learning model for jointly predicting patient length of stay and mortality[J]. *Comput Biol Med*, 2024, 183: 109237. DOI: [10.1016/j.combiomed.2024.109237](https://doi.org/10.1016/j.combiomed.2024.109237).
- 10 Kurniati AP, Rojas E, Hogg D, et al. The assessment of data quality issues for process mining in healthcare using Medical Information Mart for Intensive Care III, a freely available e-health record database[J]. *Health Informatics J*, 2019, 25(4): 1878–1893. DOI: [10.1177/1460458218810760](https://doi.org/10.1177/1460458218810760).
- 11 Cui Z, Dong Y, Yang H, et al. Machine learning prediction models for multidrug-resistant organism infections in ICU ventilator-associated pneumonia patients: analysis using the MIMIC-IV

- database[J]. *Comput Biol Med*, 2025, 190: 110028. DOI: [10.1016/j.compbimed.2025.110028](https://doi.org/10.1016/j.compbimed.2025.110028).
- 12 Lei J, Zheng LZ, Chen KY, et al. Independent effect of influenza vaccination on all-cause mortality in critically ill patients with atrial fibrillation: a retrospective study from the MIMIC-IV database[J]. *Int J Cardiol*, 2025, 433: 133246. DOI: [10.1016/j.ijcard.2025.133246](https://doi.org/10.1016/j.ijcard.2025.133246).
- 13 Hu Y, Zhao Y, Zhang J, et al. The association between triglyceride glucose-body mass index and all-cause mortality in critically ill patients with atrial fibrillation: a retrospective study from MIMIC-IV database[J]. *Cardiovasc Diabetol*, 2024, 23(1): 64. DOI: [10.1186/s12933-024-02153-x](https://doi.org/10.1186/s12933-024-02153-x).
- 14 Hu C, Li L, Li Y, et al. Explainable machine-learning model for prediction of in-hospital mortality in septic patients requiring intensive care unit readmission[J]. *Infect Dis Ther*, 2022, 11(4): 1695-1713. DOI: [10.1007/s40121-022-00671-3](https://doi.org/10.1007/s40121-022-00671-3).
- 15 Zhou S, Zeng Z, Wei H, et al. Early combination of albumin with crystalloids administration might be beneficial for the survival of septic patients: a retrospective analysis from MIMIC-IV database[J]. *Ann Intensive Care*, 2021, 11(1): 42. DOI: [10.1186/s13613-021-00830-8](https://doi.org/10.1186/s13613-021-00830-8).
- 16 Xiao LX, Zhu DL, Chen J, et al. Exploring the therapeutic role of early heparin administration in ARDS management: a MIMIC-IV database analysis[J]. *J Intensive Care*, 2024, 12(1): 9. DOI: [10.1186/s40560-024-00723-5](https://doi.org/10.1186/s40560-024-00723-5).
- 17 Guo Y, Qiu Y, Xue T, et al. Association between glycemic variability and short-term mortality in patients with acute kidney injury: a retrospective cohort study of the MIMIC-IV database[J]. *Sci Rep*, 2024, 14(1): 5945. DOI: [10.1038/s41598-024-56564-7](https://doi.org/10.1038/s41598-024-56564-7).
- 18 Tomazini BM, Besen BAMP, Taniguchi LU, et al. Association between piperacillin/tazobactam use and acute kidney injury in critically ill patients: a retrospective multicentre cohort study[J]. *J Antimicrob Chemother*, 2024, 79(3): 552-558. DOI: [10.1093/jac/dkae001](https://doi.org/10.1093/jac/dkae001).
- 19 Meng C, Trinh L, Xu N, et al. Interpretability and fairness evaluation of deep learning models on MIMIC-IV dataset[J]. *Sci Rep*, 2022, 12(1): 7166. DOI: [10.1038/s41598-022-11012-2](https://doi.org/10.1038/s41598-022-11012-2).
- 20 Jabbour G, Nolin-Lapalme A, Tastet O, et al. Prediction of incident atrial fibrillation using deep learning, clinical models, and polygenic scores[J]. *Eur Heart J*, 2024, 45(46): 4920-4934. DOI: [10.1093/eurheartj/ehae595](https://doi.org/10.1093/eurheartj/ehae595).
- 21 杨美程, 李润发, 陈辉, 等. 人工智能辅助诊疗在脓毒症管理中的应用进展[J]. *药学进展*, 2023, 47(10): 758-768. [Yang MC, Li RF, Chen H, et al. Advances in the Application of Artificial Intelligence-Assisted Diagnosis and Treatment in Sepsis Management[J]. *Progress in Pharmaceutical Sciences*, 2023, 47(10): 758-768.] DOI: [10.20053/j.issn1001-5094.2023.10.006](https://doi.org/10.20053/j.issn1001-5094.2023.10.006).
- 22 梁瀚文. 合并严重高碳酸血症的 ARDS 患者呼吸频率与预后的关系研究及预后预测模型的建立[D]. 广州: 广州医科大学, 2023. [Liang HW. Analyzing the association between respiratory rate and outcomes of invasively ventilated acute respiratory distress syndrome patients with severe hypercapnia and constructing a predictive model of hospital mortality[D]. Guangzhou: Guangzhou Medical University, 2023.] DOI: [10.27043/d.cnki.ggzyc.2023.000770](https://doi.org/10.27043/d.cnki.ggzyc.2023.000770).
- 23 郑晓博, 张中伟. 人工智能在重症医学教学中的应用展望[J]. *西南医科大学学报*, 2023, 46(6): 468-472. [Zheng XB, Zhang ZW. Application prospect of artificial intelligence in critical care medicine teaching[J]. *Journal of Southwest Medical University*, 2023, 46(6): 468-472.] DOI: [10.3969/j.issn.2096-3351.2023.06.002](https://doi.org/10.3969/j.issn.2096-3351.2023.06.002).

收稿日期: 2025 年 04 月 06 日 修回日期: 2025 年 05 月 27 日  
本文编辑: 张苗 黄笛

引用本文: 王晓莉, 毛智. MIMIC数据库的多维透视: 发展历程、数据概况、分析方法及全球影响综述[J]. *数理医药学杂志*, 2025, 38(7): 484-490. DOI: [10.12173/j.issn.1004-4337.202504023](https://doi.org/10.12173/j.issn.1004-4337.202504023).

Wang XL, Mao Z. A multidimensional perspective on MIMIC database: a review of development history, data overview, analytical methods, and global influence[J]. *Journal of Mathematical Medicine*, 2025, 38(7): 484-490. DOI: [10.12173/j.issn.1004-4337.202504023](https://doi.org/10.12173/j.issn.1004-4337.202504023).