華東理工大學学报 自然科学版 Natural Science Edition
JOURNAL OF EAST CHINA UNIVERSITY OF SCIENCE AND TECHNOLOGY

AI驱动的化学品安全性评估：技术范式、应用进展与未来展望

杨晨 张苗 唐赟

AI-Driven Chemical Safety Assessment: Technical Paradigms, Application Progress and Future Prospects

YANG Chen, ZHANG Miao, TANG Yun

在线阅读 View online: http://doi.org/10.14135/j.cnki.1006-3080.20260311002

您可能感兴趣的其他文章

Articles you may be interested in

农业化工的新进展：纳米农用化学品

Development of Agrochemical Engineering: Nano-Agrochemicals

华东理工大学学报（自然科学版）. 2021, 47(5): 519-527 http://doi.org/10.14135/j.cnki.1006-3080.20200929001

氢气传感技术研究进展

Recent Progress in Hydrogen Sensing Technology

华东理工大学学报（自然科学版）. 2025, 51(6): 731-747 http://doi.org/10.14135/j.cnki.1006-3080.20250407001

菲醌染料光点击反应研究进展

Recent Advances in Phenanthrenequinone-Based Photoclick Reactions

华东理工大学学报（自然科学版）. 2026, 52(1): 26-34 http://doi.org/10.14135/j.cnki.1006-3080.20250524001

卤代甲基转移酶的发现与应用研究进展

Discovery and Research Progress of Halide Methyltransferases

华东理工大学学报（自然科学版）. 2024, 50(1): 1-3 http://doi.org/10.14135/j.cnki.1006-3080.20240117001

基于稀疏D-vine Copula的建模方法及其在过程监测中的应用

Sparse D-vine Copula-Based Modeling Approach and Its Application in Process Monitoring

华东理工大学学报（自然科学版）. 2023, 49(3): 391-400 http://doi.org/10.14135/j.cnki.1006-3080.20211231001

基于PCA-ANN的碱性电解水系统气体纯度预测

Gas Purity Prediction of Alkaline Water Electrolysis System Based on PCA-ANN

华东理工大学学报（自然科学版）. 2023, 49(3): 305-314 http://doi.org/10.14135/j.cnki.1006-3080.20220905003

扫码关注公众号，获取更多信息！

文章编号：1006-3080(2026)04-0457-09 DOI：10.14135/j.cnki.1006-3080.20260311002

AI驱动的化学品安全性评估:技术范式、应用进展与未来展望

杨 晨[1]， 张 苗[2]， 唐 赟[1]

(1. 华东理工大学药学院, 上海市新药设计重点实验室, 上海 200237； 2. 山东省淄博海关综合技术服务中心, 山东淄博 255035)

摘要: 人工智能(Artificial Intelligence, AI)技术的快速发展正在深刻重塑化学品安全性评估的传统模式。本文系统综述了AI，尤其是机器学习和深度学习方法在计算毒理学领域的研究进展与应用现状。从数据基础、算法体系与模型可解释性这3个核心维度，梳理了AI驱动安全性评估技术范式的演进路径，并结合药物研发、化妆品安全性评估以及环境新污染物治理等典型应用场景，评述了相关技术从“替代实验手段”向“安全性赋能设计工具”的角色转变；在此基础上，文章进一步分析了国际监管框架对计算毒理学方法的接受与规范化进展，探讨了AI模型在可解释性、复杂体系评估、系统集成以及伦理与标准化等方面面临的关键挑战。总之，本文梳理了技术演进脉络，提出了智能毒理系统的概念框架，旨在为AI驱动化学品安全性评估技术的科学研究与监管应用提供系统参考与前瞻性思考。

关键词: 计算毒理学；安全性评估；新途径方法(NAMs)；不良结局通路(AOP)；监管科学

中图分类号: X9 **文献标志码:** A

化学品已深度融入现代社会生产与生活的各个方面,广泛应用于医药、农药、食品及其添加剂、化妆品、日用化学品和工业材料等领域。据不完全统计,目前全球在用的商业化学品数量已超过15万种[1],且仍以每年数百种的速度持续增长。这些化学品在显著提升社会生产效率和生活质量的同时,也带来了不可忽视的人体健康与生态环境风险。化学品安全性问题一旦失控,可能造成深远的社会和环境影响。

然而,与化学品数量的快速增长形成鲜明对比的是其安全性数据的严重不足。有文章指出,在《化学品注册、评估、许可与限制法规》(REACH)注册档案中,约90%缺乏化学品毒理学特定的关键数据[2],由此形成了显著的“安全数据鸿沟”。传统的化学品安全性评估主要依赖动物实验,虽然在过去数十年中发挥了重要作用,但其在成本、周期、物种外推不确定性以及伦理争议等方面的局限性日益凸显[3]。一项全面的毒理学评估往往需要数年时间和高昂投入,已难以满足当前化学品快速迭代和精细化监管的现实需求。

在此背景下,以人工智能(Artificial Intelligence, AI)为核心的计算毒理学逐渐兴起,并成为推动化学品安全性评估范式转型的重要技术力量。随着机器学习、深度学习及高性能计算技术的成熟,研究者得以基于化学结构、生物活性和毒理数据构建预测模型,对化合物潜在毒性及其作用机制进行快速估算。这一模式推动安全性评估从以动物实验为中心的“事后验证”,逐步转向以预测模型为先导的“事前筛查”。

近年来,AI在科学研究领域的突破性进展进一步加速了这一转型进程。相关成果在多个学科方向获得高度关注,也促使监管机构重新审视新方法在风险评估体系中的定位。在化学品安全性领域,

收稿日期: 2026-03-11

基金项目: 国家重点研发计划课题(2023YFF1204904);海关总署科研项目(2024HK185)

作者简介: 杨 晨(2000—),男,贵州福泉人,博士生,主要研究方向为计算毒理学。E-mail: yc956442829@163.com

通信联系人: 唐 赟, E-mail: ytang234@ecust.edu.cn

引用本文: 杨 晨, 张 苗, 唐 赟. AI驱动的化学品安全性评估:技术范式、应用进展与未来展望[J]. 华东理工大学学报(自然科学版), 2026, 52(4): 457-465.

Citation: YANG Chen, ZHANG Miao, TANG Yun. AI-Driven Chemical Safety Assessment: Technical Paradigms, Application Progress and Future Prospects[J]. Journal of East China University of Science and Technology, 2026, 52(4): 457-465.

AI驱动的计算毒理学不仅被视为动物实验的重要补充,更被赋予在早期风险识别、优先排序和安全性设计中的战略意义。

基于上述背景,本文系统综述了AI驱动化学品安全性评估的技术体系与应用进展,重点分析其在数据、算法与可解释性方面的发展脉络,并结合典型应用场景探讨其在实际监管与产业实践中的价值与挑战,旨在为该领域的持续发展与规范化应用提供参考。

1 AI驱动安全性评估的技术范式:从数据到可解释系统的演进

计算毒理学,亦称预测毒理学,是指基于化学结构、生物活性与毒理学数据,借助AI和计算模型对化学品潜在健康或环境风险进行定量或定性评估的方法体系[4]。随着AI技术的发展,该领域已从早期以单一定量构效关系(Quantitative Structure-Activity Relationship, QSAR)模型为核心的研究范式,逐步演进为一个融合多源数据、复杂算法与监管需求的系统工程。当前,AI驱动的安全性评估不再局限于毒性终点的预测准确性,而是更加关注模型的适用范围、可解释性以及在实际决策场景中的可用性。

1.1 AI驱动化学品安全性评估的基本流程

尽管具体模型和应用场景存在差异,AI毒性预测模型的构建和应用通常遵循相对标准化的技术流程(图1)[5],其中模型构建一般可分为6个步骤:(1)数据收集与处理。系统整合公开数据库、监管提交数据、文献报道及高通量实验结果,涵盖化学结构、理化性质、生物活性及毒理终点等信息。由于多源数据往往存在格式不一、终点定义模糊等问题,必须对其进行严格的清洗与标准化,以消除数据冲突。因此,数据可靠性与标准化正是确保模型可靠性的前提。(2)分子结构表征。将化学结构转化为模型可处理的数学表示形式,包括分子描述符、分子指纹、图结构表示以及三维构象特征等。(3)特征筛选与降维。通过统计分析或机器学习方法去除冗余和噪声特征,提升模型泛化能力并降低过拟合风险。(4)模型构建与训练。采用回归或分类算法建立毒性预测模型,常用方法包括随机森林(Random Forest, RF)、支持向量机、深度神经网络及图神经网络(Graph Neural Network, GNN)等。(5)模型评价与验证。利用交叉验证、外部验证集等手段,从拟合度、稳健性和预测能力等方面对模型性能进行系统评估。(6)应用域界定。应用域不仅是技术评估的最终环节,更是模型获得监管认可的生命线。明确模型适用的化学空间和数据分布范围,避免在超出训练域的条件下误用模型结果。

上述流程和经济合作与发展组织(Organization for Economic Co-operation and Development, OECD)提出的QSAR模型验证原则在终点明确性、算法透明性和应用域定义等方面高度契合,为AI模型在监管场景中的应用奠定了方法学基础。

1.2 数据基础:迈向标准化、多模态与规模化

高质量数据是AI驱动安全性评估的基础条件。当前,计算毒理学正经历从“小规模、单一来源数据”向“大规模、标准化、多模态数据体系”的转

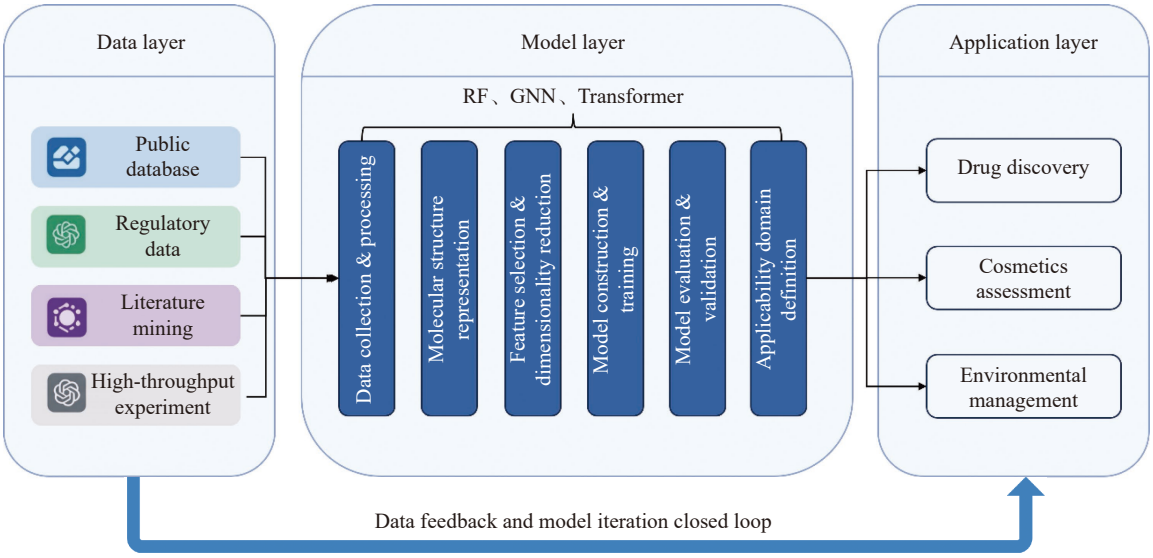


图1 AI驱动化学品安全性评估的技术流程与应用闭环

Fig. 1 Technical workflow and application closed-loop of AI-driven chemical safety assessment

变。国家级和行业级数据库的建设成为重要发展趋势[6-7], 其目标在于整合分散于不同机构和研究中的毒理数据, 缓解长期存在的数据孤岛问题。

除传统结构化数据外, 非结构化数据的价值正逐步被挖掘。借助自然语言处理和图像识别等技术, 研究人员可从文献、实验报告及历史档案中自动提取毒性相关信息[8-9], 实现数据的快速结构化。这一过程不仅显著提升了数据获取效率, 也为模型覆盖更多毒性终点提供了可能。

与此同时, 数据异质性、标签噪声以及实验条件差异等问题仍然制约模型性能。如何在保证数据规模的同时提升数据一致性和可信度, 是当前数据驱动型安全评估亟待解决的核心问题之一。

1.3　核心算法: 从二维描述符到三维感知与预训练模型

算法体系的演进是推动预测性能提升的关键动力。早期 QSAR 模型主要依赖二维分子描述符和指纹信息, 虽具有实现简单、计算效率高的优势, 但在刻画立体构象和复杂相互作用方面存在明显局限。

近年来, 引入三维结构信息的模型逐渐受到关注。等变神经网络等方法能够在原子层面学习分子的空间构象特征[10], 从而更准确地表征分子与生物靶标之间的相互作用关系。这类模型在区分立体异构体毒性差异等问题上展现出明显优势。

另一方面, 基于自监督学习的预训练模型为毒性预测提供了新的技术路径[11-12]。通过在千万级规模未标注分子数据上学习通用化学表示, 再针对具体毒性终点进行微调, 模型可在有限标注数据条件下保持较好的预测性能。这一策略为慢性毒性、发育毒性等数据稀缺终点的建模提供了重要补充。

1.4　模型可解释性: 从“结果可信”走向“机制可理解”

在安全性评估领域, 模型的可解释性与预测精度同等重要。高性能但不可解释的“黑箱模型”难以获得监管机构和领域专家的充分信任, 也限制了其在机制研究中的应用。

当前, 可解释 AI 方法已被广泛引入计算毒理学研究中。通过特征归因分析、注意力机制可视化等手段[13], 研究人员可识别对预测结果贡献较大的分子结构特征或子结构片段。这不仅有助于增强模型透明度, 也为识别潜在毒性警示结构提供了技术支撑。

更进一步的研究开始尝试将毒理学知识和不良结局通路(Adverse Outcome Pathway, AOP)信息嵌入模型构建过程[14], 使可解释性由主要依赖数据特征的“事后相关性解释”向“机制约束下的因果推断”演进。这一方向被认为是推动 AI 模型从相关性预测迈向机制理解的重要途径。

1.5　在线预测工具与平台化实践

随着计算毒理学方法的逐步成熟, AI 驱动的安全性评估正以在线平台和软件工具的形式向实际应用场景拓展。相较于单一模型研究, 平台化系统通过集成多毒性终点预测模型、数据管理模块和结果可视化功能, 有效降低了技术使用门槛, 成为连接方法研究与产业实践的重要载体。

目前, 国际上已开发了多种面向化学品安全性评估的在线预测平台, 用于支持早期风险筛查和优先排序。在此基础上, 国内研究团队也持续推进具有自主知识产权的预测系统建设。例如, 作者团队自 2012 年起开发并持续更新的 admetSAR 在线系统(https://lmmd.ecust.edu.cn/admetsar3/index.php)[15-17], 已集成多种健康毒性与生态毒性终点预测模型, 并在后续版本中扩展了性质优化和多终点评估功能。相关系统以网络服务和单机软件等多种形式提供, 已在科研和实际评估工作中得到应用。

需要指出的是, 平台化应用的关键不在于具体工具本身, 而在于其背后的模型更新机制、应用域控制和结果可追溯性设计。随着监管对模型透明度和稳定性要求的提升, 如何在平台层面实现模型版本管理、性能监控及结果解释, 将成为未来平台化预测系统持续发展的重要方向。

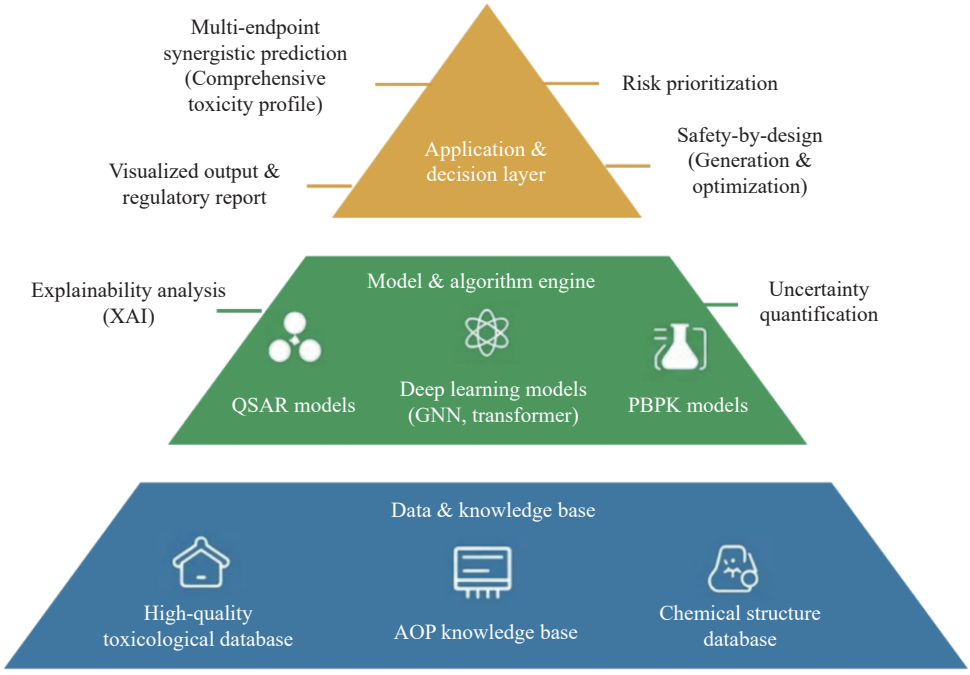
1.6　前沿趋势: 从预测模型走向“智能毒理系统”

随着 AI 方法在计算毒理学中的持续演进, 化学品安全性评估的技术形态正发生结构性变化。当前研究表明, 该领域正在由以单一预测模型为核心的分析模式, 逐步迈向融合多模型、多数据源与决策逻辑的系统化评估范式[18], 即“智能毒理系统”。

首先, 在评估对象层面, 安全性判断正由单一毒性终点预测转向多终点协同刻画。通过并行整合健康毒性、生态毒性及理化性质等多个维度的信息, 模型能够构建更接近真实风险情景的综合毒性表征, 为后续风险分级和优先排序提供基础。其次, 在模型组织形式上, 预测任务逐渐从“孤立建模”转向“模型集成”。针对不同毒性终点、数据类型和适用范围构建的模型, 通过集成学习或权重证据(Weight-of-Evidence)策略进行联合应用, 有助于提升整体评估结果的稳健性, 并降低单一模型偏差带来的不确定性。再次, 在知识利用方式上, 计算毒理学正从事后解释逐步过渡到机制信息的前置嵌入。通过将 AOP、代谢转化规则及既有毒理学认知引入模型的构建与结果整合过程, 预测结果的生物学合理性和

可解释性得到增强，使模型更贴近毒理学分析逻辑。此外，平台化趋势日益明显。预测模型正以在线系统或软件平台的形式被系统整合，与数据管理、结果可视化和应用域控制等模块协同运行，从而支持非专业用户在规范框架下开展安全性评估。

总体而言，AI 驱动的化学品安全性评估正由单点预测工具，演进为具备多终点协同、模型集成和知识约束特征的智能毒理系统(图 2)。这一转变为其在后续应用场景和监管决策中的推广奠定了技术基础。



XAI—Explainable artificial intelligence; PBPK—Physiologically based pharmacokinetic

图 2 面向未来的智能毒理系统概念框架

Fig. 2 A conceptual framework for the future intelligent toxicology system

2 应用进展评述:从药物研发到多元复杂场景

2.1 药物研发:早期安全风险识别与“失败前移”策略

在药物研发领域，安全性问题是导致候选化合物在临床前和临床阶段失败的主要原因之一。统计数据显示，毒性相关问题约占药物研发失败原因的 1/3 以上[19-20]，且往往在投入数年时间和高昂成本后才被发现。因此，如何在药物研发早期实现对潜在毒性风险的快速识别，是提升研发效率、降低整体成本的关键科学问题。

AI 驱动的安全性评估模型在该阶段发挥着“失败前移”的核心作用。通过整合分子结构信息、高通量体外实验数据以及既有毒理学知识，机器学习模型能够对药物诱导肝毒性、心脏毒性(如人类 Ether-a-go-go 相关基因(hERG)通道抑制)、致突变性等关键安全终点进行并行预测，从而在化合物筛选和先导优化阶段提前排除高风险分子。与传统基于单一实验或经验规则的判断方式相比，该策略在覆盖终点数量和筛选效率方面具有明显优势。

近年来，随着多组学技术和高内涵实验手段的发展，部分研究开始尝试将转录组、代谢组及细胞表型数据引入 AI 模型，构建多尺度毒性预测框架。这类方法不仅关注化合物是否具有毒性，还尝试刻画其可能影响的生物通路和作用机制，为后续机制研究和风险管理提供线索。然而，此类模型在数据一致性、跨实验平台泛化能力以及结果可解释性方面仍面临挑战，其在监管申报中的应用仍以辅助证据为主。

总体来看，AI 驱动的安全性评估已成为现代药物研发流程中不可或缺的组成部分，其价值正从“简单筛选工具”向“机制感知型决策支持手段”转变。未来，通过进一步提升模型的可解释性和与监管框架的适配度，AI 驱动的安全性评估在新药安全性评价体系中的权重有望持续提升。

2.2 化妆品安全评估:动物替代驱动下的 AI 应用高地

化妆品领域是当前 AI 驱动安全性评估技术落地最早、监管接受度最高的应用场景之一，其核心驱动力来自全球范围内对化妆品动物实验的严格限制乃至全面禁止，使得传统毒理学测试路径在该领域面临根本性挑战。在此背景下，以计算毒理学为

代表的新途径方法(New Approach Methodologies, NAMs)逐步从“辅助工具”演变为化妆品安全评估体系中的关键技术支柱。

欧盟自 2013 年起全面禁止化妆品及其原料的动物实验[21], 并明确要求企业采用包括 QSAR、Read-Across 和权重证据(Weight-of-Evidence)在内的替代方法开展安全性评估。这一政策直接推动了 AI 预测模型在皮肤刺激性、皮肤致敏性、眼刺激性、光毒性及内分泌干扰性等关键毒性终点上的快速发展[22-23]。相关研究表明, 基于机器学习的模型在多个终点上的预测性能已与传统体外实验相当, 且在高通量筛查与早期风险排除方面展现出显著优势。

我国化妆品监管体系近年来也持续推进与国际接轨。国家药品监督管理局发布的《化妆品安全评估技术导则(2021 年版)》首次系统性地认可了计算毒理学方法在化妆品安全评估中的应用价值, 明确提出可在满足科学合理性和适用性的前提下, 采用预测模型对部分毒理学终点进行评估。这一政策导向为 AI 模型在我国化妆品行业的规范化应用奠定了制度基础, 也显著提升了企业对相关技术的接受度和应用意愿。

从技术实现角度看, 化妆品原料结构类型相对集中、暴露途径较为明确(以皮肤接触为主), 这为构建高质量预测模型提供了有利条件。当前主流方法通常整合分子结构特征、理化性质参数及既有毒理数据, 建立多终点并行预测模型, 用于对新原料或配方成分进行快速风险筛查。同时, 结合可解释 AI 技术, 对模型给出的风险判断进行结构层面的溯源分析, 有助于识别潜在的“高风险结构警示片段”, 为配方优化和原料替代提供科学依据。

值得关注的是, 随着监管要求从“是否安全”逐步转向“安全性证据链的完整性”, 化妆品安全评估正从单一模型预测向集成化决策支持系统演进。未来的发展趋势将更加注重多模型结果的综合权衡、不确定性量化以及与体外实验和文献证据的协同使用, 从而构建透明、可追溯、可监管的 AI 辅助安全评估流程。这一转变不仅提升了评估结论的可信度, 也为 AI 技术在其他高监管要求领域的推广应用提供了可借鉴的实践范式。

2.3　环境污染物与新兴风险物质的智能筛查与优先管控

在环境领域, 化学品种类繁多、来源复杂, 新污染物不断涌现, 传统依赖实验的风险评估模式在时间、成本和覆盖范围上均难以满足实际监管需求。特别是对于环境中广泛存在、但毒性数据极为有限的新兴污染物, 如何实现快速筛查和风险优先排序, 已成为环境管理领域的突出挑战[24-25]。

AI 驱动的计算毒理学为解决这一问题提供了有效技术路径。通过整合化学结构特征、环境行为参数及有限的毒理数据, 预测模型可在大规模化学品数据库中识别潜在高风险物质, 并结合暴露情景对其生态和健康风险进行分级评估。这种“预测-排序-聚焦”的评估模式显著提高了监管资源配置效率, 为精准管控提供了科学依据。

近年来, 我国围绕新污染物治理持续推进相关平台和技术体系建设, 逐步形成了以计算毒理学为核心的风险筛查技术框架。相关实践表明, 基于 AI 模型的初筛结果与后续重点实验验证之间具有较好的一致性, 能够有效缩短评估周期并降低总体成本。同时, 结合区域环境数据开展本土化模型构建, 有助于提升预测结果在具体监管场景中的适用性。

需要指出的是, 环境风险评估通常涉及多介质迁移、多物种效应及长期低剂量暴露等复杂因素, 这对 AI 模型的泛化能力和系统集成水平提出了更高要求。未来的发展方向在于将毒性预测与环境暴露建模、生态效应评估和不确定性分析深度耦合, 逐步构建面向监管决策的智能风险评估系统。

2.4　极端条件与复杂体系: 传统实验难以覆盖场景的计算探索

除常规应用场景外, AI 驱动的安全性评估在传统实验难以开展或风险极高的特殊体系中展现出独特价值。这类场景通常包括高毒性物质、极端反应条件以及实验成本或伦理风险显著高于常规研究对象的体系。

在相关研究中, 学者通过结合量子化学计算、分子动力学模拟与机器学习方法, 对高危化学品在不同环境条件下的稳定性、降解路径及潜在毒性产物进行预测分析[26]。这种完全基于计算的方法避免了直接实验带来的安全风险, 为理解物质环境行为和制定风险应对策略提供了重要信息。此外, 通过分子生成和虚拟筛选手段, 设计结构和理化性质相近但毒性显著降低的模拟物, 也为后续实验研究提供了可行替代方案。

类似的计算范式也正在被探索用于纳米材料、微塑料以及化学品混合物等复杂体系的安全性评估[27]。这些对象往往具有高度异质性, 其毒性效应不仅受粒径、形貌、组成比例及暴露方式等多重因素影响, 更关键的是, 混合物体系中的各组分之间会存在互相影响, 使得总毒性并非简单的各组分毒性相加, 即非加和性相互作用(如协同作用或拮抗作用), 该作用非常复杂, 且相关实验数据严重匮乏, 这使得传

统 QSAR 模型难以直接适用。近年来的研究尝试通过多模态数据融合和深度学习方法刻画此类复杂关系,尽管仍处于探索阶段,但已显示出一定潜力。

总体而言,AI 在极端与复杂体系中的应用尚未形成成熟范式,但其在突破实验边界、拓展安全评估适用范围方面具有不可替代的作用。随着数据积累和模型方法的持续进步,该方向有望成为计算毒理学未来的重要增长点之一。

3 全球监管演进与科学挑战:从方法接受到体系共建

随着 AI 和计算毒理学方法的快速发展,全球化学品监管体系正经历从“谨慎观望”到“主动引导”的转变。相较于早期仅将计算模型作为辅助分析工具,当前监管机构更加关注其在风险评估流程中的系统性角色,以及模型结果在决策中的可重复性、透明性和可解释性。这一转变标志着计算毒理学正逐步从研究工具走向监管科学的重要组成部分。

3.1 国际监管框架对计算毒理学方法的逐步接受

在国际层面,计算毒理学方法的监管接受度持续提升。《全球化学品统一分类和标签制度(GHS)》已明确提出,在化学品危害分类中可使用 QSAR 等非实验方法生成证据。OECD 自 2004 年起系统推进 QSAR 计划,并提出了广泛采用的模型验证原则,对毒性终点、算法透明性、应用域及模型性能评价等提出明确要求。2023 年发布的《(Q)SAR Assessment Framework》[28] 进一步为监管人员评估模型科学合理性提供了操作性指南,标志着 QSAR 方法在监管应用中的规范化程度持续提高。

欧盟在计算毒理学监管实践方面处于国际领先地位。《REACH》[29] 明确规定,在满足特定条件下,可使用 QSAR、Read-Across 等新途径方法替代动物实验获取物质固有属性信息。尤其是在化妆品领域,欧盟通过全面禁止动物实验的立法安排,倒逼非动物替代方法的快速发展,使 AI 驱动的预测模型在安全评估实践中获得了广泛应用。

美国在计算毒理学领域起步较早,并形成了以高通量筛查和预测模型为核心的技术体系。美国环保署通过国家计算毒理学中心持续推进 ToxCast、Tox21 等计划,构建了覆盖大量化学品的毒性预测和数据平台。2022 年通过的《FDA 现代化法案 2.0》[30] 取消了新药审批中对动物实验的强制要求,进一步为计算模型和 AOP 其他非动物方法在监管中的应用创造了制度空间。目前,美国食品药品监督管理局正通过多项试点计划探索 AI 模型在毒理学评估中的规范使用路径。

3.2 我国计算毒理学监管实践的加速推进

近年来,我国在化学品安全监管领域对计算毒理学方法的重视程度显著提升。2020 年修订发布的《新化学物质环境管理登记指南》[31] 首次明确提出,在新化学物质申报过程中可采用预测模型估算部分毒理学参数,这为计算毒理学方法的制度化应用奠定了基础。随后,国家药品监督管理局发布的《化妆品安全评估技术导则(2021 年版)》[32] 进一步明确了预测模型在化妆品安全评估中的应用原则。

在组织层面,生态环境部相继成立化学物质环境风险评估中心及预测模型工具学组,推动计算毒理学方法在新污染物治理和环境风险评估中的实际应用。这些举措表明,我国监管体系正在从“原则性认可”迈向“规范化实施”,计算毒理学正逐步成为支撑精细化化学品管理的重要技术手段。

3.3 从技术工具到监管科学:面临的核心挑战

尽管全球范围内对计算毒理学方法的接受度不断提高,但其在监管决策中的广泛应用仍面临一系列科学与技术挑战。首先,不同模型在数据来源、算法结构和应用域方面存在显著差异,预测结果的可重复性和跨平台一致性仍有待加强。其次,多数高性能模型依赖复杂的深度学习结构,其决策过程难以完全解释,增加了监管人员理解和采信结果的难度。

此外,监管应用往往需要对预测不确定性进行量化评估,而当前多数模型仍以点预测为主,难以系统反映预测结果的可信区间。这一问题在涉及低剂量长期暴露、混合物效应以及敏感人群风险评估时尤为突出。为满足监管决策对量化不确定性的需求,未来亟须引入深度集成等不确定性量化技术,将预测结果转化为附带置信区间的概率分布,从而为复杂场景提供更可靠的评估依据。

从更长远的视角看,计算毒理学面临的核心问题并非单一模型性能的提升,而是如何构建可验证、可追溯、可监管的模型应用体系。这不仅需要技术层面的持续创新,也依赖于监管机构、学术界和产业界之间的协同合作,共同推动模型标准、数据规范和评估流程的完善。

4 结束语

尽管 AI 驱动的化学品安全性评估在方法和应用层面已取得显著进展,但其进一步发展仍面临一

系列关键科学问题和技术挑战。未来研究有望在以下几个方向持续深化: (1)从静态预测走向毒性动态过程建模。现有模型多基于单一剂量或静态实验条件进行毒性预测, 难以充分反映剂量-时间-效应之间的动态关系。未来, 融合毒代动力学模型[33]、生理药代动力学(Physiologically Based Pharmacokinetic, PBPK)模型、生物通路信息和时间序列数据的 AI 方法, 有望实现对毒性发生与演化过程的更精细刻画, 为慢性毒性和低剂量长期暴露风险评估提供技术支撑。(2)从相关性预测迈向机制约束与因果推断。单纯依赖数据驱动的相关性模型在复杂场景下存在解释能力不足的问题。将 AOP、毒理学机制知识和因果推断方法[34]引入模型构建过程, 有望提升预测结果的生物学合理性和可解释性, 推动计算毒理学从“经验拟合”向“机制驱动”转变。(3)面向复杂体系的专用 AI 评估框架。化学品混合物、新材料、纳米材料及微塑料等对象的安全性评估仍是当前方法体系中的薄弱环节。这类体系具有高度异质性和多因素耦合特征, 亟须发展融合多模态数据和多尺度信息的专用 AI 模型, 以突破传统 QSAR 方法的适用边界。(4)从单一模型到系统集成与闭环优化。未来的安全性评估不再局限于孤立的预测模型, 而将向集成化、自动化系统演进。通过将 AI 模型与高通量体外实验平台、器官芯片及自动化合成技术相结合, 可构建“设计-评估-验证-反馈”闭环体系, 从源头上实现化学品安全性的主动设计。(5)模型治理、标准化与伦理问题[35]的系统推进。随着 AI 模型在监管决策中的作用不断增强, 其算法偏差、结果可追溯性以及责任界定等问题日益凸显。其中, 监管要求的结果可追溯性与数据隐私保密之间存在冲突问题, 为实现二者的平衡, 未来亟须引入联邦学习等技术, 在数据可用不可见的前提下进行联合建模与验证。在标准化体系建设方面, 可借助模型卡与数据清单等标准化文档工具, 通过规范化记录模型的适用域、性能基准、底层数据分布特征及潜在局限性, 从文档层面切实增强 AI 模型在监管应用中的透明度和可追溯性。此外, 推动建立统一的模型验证标准、性能基准和质量控制流程, 并同步开展算法伦理与治理研究, 是确保 AI 技术安全、可靠应用的重要前提。

由此可见, AI 技术正在推动化学品安全性评估发生深刻的范式转型。通过整合多源数据、先进算法与可解释性分析, 计算毒理学正从以动物实验为核心的后验性评估模式, 逐步演进为以预测模型为先导的前瞻性风险管理工具, 其应用范围已从早期的毒性筛查, 拓展至药物研发、化妆品安全评估、环境新污染物治理以及极端和复杂体系的风险分析等多个领域。

在监管层面, 全球主要国家和地区正逐步将计算毒理学方法纳入化学品管理和风险评估框架, 推动其从技术探索走向规范化应用。我国在相关制度建设和技术实践方面也取得了积极进展, 显示出将 AI 驱动安全性评估作为支撑精细化监管和新污染物治理的重要技术方向的明确态度。

总体而言, 尽管在模型泛化能力、复杂体系评估和标准化等方面仍面临挑战, AI 驱动的化学品安全性评估已展现出不可替代的科学价值和应用潜力。随着技术进步、数据积累及监管协同的持续推进, 该领域有望在保障化学品全生命周期安全、促进绿色与可持续化学发展中发挥更加关键的作用。

参考文献:

[1] WANG Z, WALKER G W, MUIR D C G, *et al*. Toward a global understanding of chemical pollution: A first comprehensive analysis of national and regional chemical inventories[J]. Environmental Science & Technology, 2020, 54(5): 2575-2584.

[2] CORIA J. Policy monitor—The Economics of toxic substance control and the REACH directive[J]. Review of Environmental Economics and Policy, 2018, 12(2): 342-358.

[3] KREWSKI D, ANDERSEN M E, TYSHENKO M G, *et al*. Toxicity testing in the 21st century: Progress in the past decade and future perspectives[J]. Archives of Toxicology, 2020, 94(1): 1-58.

[4] RAIES A B, BAJIC V B. *In silico* toxicology: Computational methods for the prediction of chemical toxicity[J]. Wiley Interdisciplinary Reviews: Computational Molecular Science, 2016, 6(2): 147-172.

[5] YANG H, SUN L, LI W, *et al*. *In silico* prediction of chemical toxicity for drug design using machine learning methods and structural alerts[J]. Frontiers in Chemistry, 2018, 6: 30.

[6] WU L, YAN B, HAN J, *et al*. TOXRIC: A comprehensive database of toxicological data and benchmarks[J]. Nucleic Acids Research, 2023, 51(D1): D1432-D1445.

[7] WEN Y, YU Y, SUN Q, *et al*. Study of construction of health risk and toxicity database of pollutants in China[J]. Research of Environmental Sciences, 2022, 35(11): 2608-2617.

[8] BAUER C, DANG D L T, VAN DEN BEUCKEN T,

et al. Systematic analysis of hepatotoxicity: Combining literature mining and AI language models[J]. Frontiers in Artificial Intelligence, 2025, 8: 1561292.

[9] KLEINSTREUER N, HARTUNG T. Artificial intelligence (AI)—It's the end of the tox as we know it (and I feel fine)[J]. Archives of Toxicology, 2024, 98(3): 735-754.

[10] WU L, WANG F, ZHANG Y, *et al*. Multiscale-information-embedded universal toxicity prediction framework[J]. Environmental Science & Technology, 2025, 59(41): 21830-21847.

[11] CHITHRANANDA S, GRAND G, RAMSUNDAR B. ChemBERTa: Large-scale self-supervised pretraining for molecular property prediction[EB/OL]. (2020-10-23). https://arxiv.org/abs/2010.09885.

[12] ROSS J, BELGODERE B, CHENTHAMARAKSHAN V, *et al*. Large-scale chemical language representations capture molecular structure and properties[J]. Nature Machine Intelligence, 2022, 4(12): 1256-1267.

[13] HA S, BANG D, KIM S. Fate-tox: Fragment attention transformer for E(3)-equivariant multi-organ toxicity prediction[J]. Journal of Cheminformatics, 2025, 17(1): 74.

[14] LIN Z, CHOU W C. Machine learning and artificial intelligence in toxicological sciences[J]. Toxicological Sciences, 2022, 189(1): 7-19.

[15] CHENG F, LI W, ZHOU Y, *et al*. admetSAR: A comprehensive source and free tool for assessment of chemical ADMET properties[J]. Journal of Chemical Information and Modeling, 2012, 52(11): 3099-3105.

[16] YANG H, LOU C, SUN L, *et al*. admetSAR 2.0: Web-service for prediction and optimization of chemical ADMET properties[J]. Bioinformatics, 2019, 35(6): 1067-1069.

[17] GU Y, YU Z, WANG Y, *et al*. admetSAR3.0: A comprehensive platform for exploration, prediction and optimization of chemical ADMET properties[J]. Nucleic Acids Research, 2024, 52(W1): W432-W438.

[18] LUECHTEFELD T, HARTUNG T. Navigating the AI frontier in toxicology: Trends, trust, and transformation[J]. Current Environmental Health Reports, 2025, 12: 51-69.

[19] GUENGERICH F P. Mechanisms of drug toxicity and relevance to pharmaceutical development[J]. Drug Metabolism and Pharmacokinetics, 2011, 26(1): 3-14.

[20] CARRACEDO-REBOREDO P, LIÑARES-BLANCO J, RODRÍGUEZ-FERNÁNDEZ N, *et al*. A review on machine learning approaches and trends in drug discovery[J]. Computational and Structural Biotechnology Journal, 2021, 19: 4538-4558.

[21] European Parliament Council of the EU. Regulation (EC) No 1223/2009. on cosmetic products (recast) [R/OL]. https://eur-lex.europa.eu/eli/reg/2009/1223/oj/eng.

[22] OECD. Guideline No. 497: Defined approaches on skin sensitisation [S/OL]. https://www.oecd.org/en/publications/guideline-no-497-defined-approaches-on-skin-sensitisation_b92879a4-en.html.

[23] 所雅琼, 罗飞亚, 张凤兰, 等. 皮肤致敏替代方法和整合策略研究进展及思考 [J]. 香料香精化妆品, 2021(5): 96-102.

[24] TARIQ F, AHRENS L, ALYGIZAKIS N A, *et al*. Computational tools to facilitate early warning of new emerging risk chemicals[J]. Toxics, 2024, 12(10): 736.

[25] DULIO V, ALYGIZAKIS N, NG K, *et al*. Beyond target chemicals: Updating the NORMAN prioritisation scheme to support the EU chemicals strategy with semi-quantitative suspect/non-target screening data[J]. Environmental Sciences Europe, 2024, 36: 113.

[26] 施荣鑫, 张林, 曹泽星. 机器学习辅助预测潜在 A 类神经毒剂及其降解性质 [J]. 环境化学, 2025, 44(7): 2727-2738.

[27] ZHANG F, WANG Z, PEIJNENBURG W J G M, *et al*. Machine learning-driven QSAR models for predicting the mixture toxicity of nanoparticles[J]. Environment International, 2023, 177: 108025.

[28] OECD. (Q)SAR Assessment framework: Guidance for the regulatory assessment of (Quantitative) structure activity relationship models, predictions, and results based on multiple predictions[R]. Paris: OECD Publishing, 2023.

[29] ECHA. How to use read-across and (Q)SARs under REACH[R]. Helsinki: European Chemicals Agency, 2022.

[30] U.S. Senate. S. 5002: FDA Modernization Act 2.0 (117 th Congress) [R/OL]. https://www.congress.gov/bill/117th-congress/senate-bill/5002.

[31] 中华人民共和国生态环境部. 新化学物质环境管理登记指南 [R]. 北京: 生态环境部, 2020.

[32] 国家药品监督管理局. 国家药品监督管理局公告2021 年第 51 号: 化妆品安全评估技术导则(2021 年版)[EB/OL]. (2021-04-19). https://www.nmpa.gov.cn/xxgk/ggtg/hzhpggtg/jmhzhptg/20210419163037171.html.

[33] HAN P, LI X, YANG J, *et al*. Advancing toxicity predictions: A review on in vitro to *in vivo* extrapolation in next-generation risk assessment[J]. Environment & Health, 2024, 2(7): 499-513.

[34] HEMMERICH J, ECKER G F. *In silico* toxicology: From structure-activity relationships towards deep learning and adverse outcome pathways[J]. Wiley Interdisciplinary Reviews: Computational Molecular Science, 2020, 10(4): e1475.

[35] SADEK M, KALLINA E, BOHNÉ T, *et al*. Challenges of responsible AI in practice: Scoping review and recommended actions[J]. AI & Society, 2025, 40(1): 199-215.

AI-Driven Chemical Safety Assessment: Technical Paradigms, Application Progress and Future Prospects

YANG Chen[1], ZHANG Miao[2], TANG Yun[1]

(1. Shanghai Key Laboratory of New Drug Design, School of Pharmacy, East China University of Science and Technology, Shanghai 200237, China; 2. Zibo Customs Comprehensive Technical Service Center, Zibo 255035, Shandong, China)

Abstract: The rapid development of artificial intelligence (AI) is profoundly reshaping the traditional paradigm of chemical safety assessment. This paper systematically reviews the research progress and application status of AI, particularly machine learning and deep learning methods, in the field of computational toxicology. From three core dimensions—data foundation, algorithm system and model interpretability, the evolutionary path of the AI-driven safety assessment technical paradigm is summarized. Combined with typical application scenarios such as drug research and development, cosmetic safety evaluation, and the control of emerging environmental pollutants, the role transformation of related technologies from "alternative experimental methods" to "safety-enabled design tools" is discussed. On this basis, the paper further analyzes the acceptance and standardization progress of computational toxicology methods in international regulatory frameworks, and explores the key challenges faced by AI models in terms of interpretability, complex system assessment, system integration, ethics, and standardization. In summary, this paper sorts out the context of technological evolution and proposes a conceptual framework for intelligent toxicology systems, aiming to provide a systematic reference and forward-looking perspective for the scientific research and regulatory application of AI-driven chemical safety assessment technologies.

Key words: computational toxicology; safety assessment; new approach methodologies (NAMs); adverse outcome pathway (AOP); regulatory science

（责任编辑：李娟）