药学管理和信息

人工智能应用于制药领域的监管对策探讨

Discussion on Regulatory Strategies of Application of Artificial Intelligence in the Pharmaceutical Industry

李　帅，董正龙，吴　浩，卓　阳，陈桂良

（上海药品审评核查中心，上海 201203）

LI Shuai, DONG Zhenglong, WU Hao, ZHUO Yang, CHEN Guiliang

(*Shanghai Center for Drug Evaluation and Inspection, Shanghai 201203*)

摘要：目前人工智能（AI）在制药领域的应用日益广泛，涵盖了药物研发、临床试验、药物生产和上市后等全生命周期。面对 AI 应用存在的一些风险和挑战，药品监管机构应对 AI 制药领域的监管予以重视并采取相应的策略。文章通过收集和分析相关文献、政策文件和监管机构声明，总结了各国药品监管机构在 AI 制药领域的监管活动和实践，提出建立指导机构和合作机制、风险评估、关于数据和模型的监管考虑、建立药品全生命周期监管措施 5 个方面的监管对策，以促进 AI 技术的发展并保护公众健康。

关键词：人工智能；制药；应用；监管

中图分类号：R95；TP399　　**文献标志码：**A　　**文章编号：**1001-8255(2024)03-0406-06

DOI：10.16522/j.cnki.cjph.2024.03.013

人工智能（artificial intelligence，AI）是利用算法或模型执行任务并展示学习、决策和预测等行为的统称。其中，机器学习（machine learning，ML）是 AI 的子集，是利用数据和算法模拟人类学习。通过处理日益增多且复杂的数据，并结合先进的计算和方法论，AI 和 ML 能帮助医药领域中的利益相关者改进疗法开发、生产、使用和评估的方式。

2021 年，提交给美国 FDA 的药品或生物制品注册申请中有超过 100 个案例包含了 AI 和 ML 组成部分[1]。但是，AI 在医药领域的应用仍存在风险和挑战，如数据偏见、数据泄露、数据代表性有限、网络安全漏洞、算法偏见、模型错误输出、大型语言模型的记忆风险以及模型过度拟合等[1—2]。为此，欧盟 EMA 和美国 FDA 分别出台了一系列相关指南或与行业讨论并征求意见，以期规范和指导 AI 在医药领域的应用，以确保维护患者的权益，促进医药行业的创新和健康发展。

1　EMA 和 FDA 发布的 AI 监管文件

截至 2023 年 7 月，EMA 和 FDA 分别发布以下关于 AI 行业评估以及监管方面的文件，具体如表 1 所示。

表 1　EMA 和 FDA 发布的 AI 监管文件

Tab.1　Documents for the AI Regulation Issued by EMA and FDA

发布机构	文件名称	发布时间
EMA	《人工智能天际线扫描评估报告》（Horizon Scanning Assessment Report—Artificial Intelligence）[3]	2021-08-06
	《人工智能在医药产品生命周期中的应用思考》［Reflection Paper on the Use of Artificial Intelligence（AI）in the Medicinal Product Lifecycle］[2]	2023-07-13
FDA	《人工智能和机器学习在药品和生物制品开发中的应用》（Using Artificial Intelligence and Machine Learning in the Development of Drug and Biological Products）[4]	2023-05-16
	《人工智能在药品生产中的应用》（Artificial Intelligence in Drug Manufacturing）[5]	2023-05-16

收稿日期：2024-01-29

作者简介：李　帅（1983—），男，药品审评员，从事药品审评、核查以及监管科学研究工作。

E-mail：lishuai@smda.sh.cn

2 国外药品监管机构在 AI 应用方面的监管活动

EMA 的监管活动主要有以下方面。①建立数字创新实验室；②提高网络监管能力，加强网络验证 AI 算法的能力；③对于数据科学、数字技术、AI 相关解决方案及其在监管系统中的应用，提出课程开发和知识共享倡议；④开发大数据培训课程，建立和维护健康数据科学和 AI 论坛；⑤建立制定指南和建议的框架，与其他监管机构（例如 FDA）实现更大范围的全球协调 [3]。

FDA 的监管活动主要有以下方面。①成立 AI 指导委员会，协调跨领域的 AI 和 ML 应用；②为涉及 AI/ML 的药品申报提供咨询；②制定 AI 和 ML 的医疗器械框架，组织数字健康技术以及 AI 和 ML 相关研讨会；③评估新兴 AI 和 ML 产品的安全性和有效性；④开展监管科学研究，开发工具和模型；⑤推出新药创新科学和技术方法试点项目，建立模型指导的药物开发试点项目；⑥利用 AI 和 ML 改进临床试验、优化剂量选择，增强预测性或机制性的安全评估；⑦上市后安全监测，如不良事件报告和质量监控 [4]。

3 AI 在药品全生命周期中的应用风险及监管

AI 在药品全生命周期中发挥着重要作用，涵盖药物发现、非临床研究、临床研究、药品生产和上市后阶段等多个方面。然而，AI 应用也存在一系列潜在风险。这些风险包括：非多数基因型和表型的伦理问题、偏差风险和其他歧视来源、数据性能和稳健性、过度拟合和数据泄露风险、生产商与第三方的质量协议、第三方云数据和模型的复杂性、潜在的网络安全漏洞、数据管理和安全性、开发和验证模型的风险。在确定适应证和剂量等高风险应用中，使用 AI 和 ML 技术来决策治疗分配、给药及参考药品特性信息时应尤其审慎。针对不同阶段应用的监管考虑如下。

3.1 药物发现阶段

申办方应审查所有模型和数据集。

3.2 非临床研究阶段

扩展标准操作规程（standard operating procedure，SOP）到所有临床前的 AI 和 ML 应用。在数据挖掘前，应根据预先指定的分析计划，对任何可能与药品收益 - 风险平衡评估相关的临床前数据进行分析。

3.3 临床研究阶段

在上市许可或临床试验申请时应全面评估模型架构、建模、验证和测试日志，并描述训练数据和数据处理管线。当将 AI 和 ML 模型用于药品临床试验中的数据转换或分析时，应遵循适用的临床试验数据管理和统计分析的原则，并应包括对下游统计推断影响的分析。在后期临床开发中，需要在分析计划中详细描述数据管理管线和用于推断的完全锁定的模型集。

3.4 早期临床试验

应采取措施确保后续临床试验的估计在统计上的稳健性，探索性分析应说明与多样性相关。当早期临床试验的数据可能产生重大的监管影响时，应与监管机构讨论。

3.5 关键临床试验

在模型部署前，应使用先前生成的数据测试性能，确保数据可代表预期使用环境或人群。试验期间任何对模型的修改都需要与监管机构沟通，不接受增量学习方法，并应修改统计分析计划。在解锁假设检验的任何数据集之前，数据预处理管线以及所有模型都应以可追溯的方式锁定，并记录在统计分析计划中。一旦解锁数据集，对于数据处理或模型的任何未经预先指定的修改，导致分析结果发生改变都将被视为事后分析，不适用于生成验证性证据。

3.6 药品生产阶段

①第三方数据管理系统与 AI 分析可为过程监测及控制模型提供支持，并必须相应地更新质量协议； 物联网技术在制药生产中的应用可能会导致数据量增加，进而提高对数据管理的实践要求。另外，申请人需平衡数据完整性和保留需求。对于设备生成的数据，还需考虑相关管理责任以及隐私安全问题。申请人需要确保 AI 的使用遵循质量风险管理、数据完整性以及其他相关法规和指南的规定。②申请人须制定并验证用于过程控制和放行检测的

模型标准。同时，可能需要在模型开发和验证中考虑模型间学习迁移的潜力。申请人必须定义验证模型，并维持模型输出与产品质量影响的标准。③申请人需要说明模型生命周期的验证策略，以及向监管机构说明如何检查持续更新的 AI 控制模型。此外，当 AI 模型引入生产条件变化时，申请人还需要提供预案，确保产品可比性，尤其是生物制品 [5]。

3.7 上市后阶段

药品上市许可持有人（Marketing Authorization Holder，MAH）有责任验证、监测和记录模型性能，并将 AI 和 ML 操作纳入药物警戒系统，以降低相关风险。

EMA 和 FDA 在对于 AI 应用于制药领域的讨论文件中指出了风险点、监管考虑以及适用的技术要求。EMA 指出了 AI 在药物发现、非临床研究、临床研究以及上市后阶段的应用风险、风险等级、监管考虑和适用的技术指导原则，FDA 则侧重于提出药品生产领域的监管考虑和适用的技术指导原则 [2,5]。

4　AI 和 ML 模型数据采集、保护与完整性的风险考量及监管

FDA 从数据特性阐述数据的质量、可靠性和代表性，并指出 AI 和 ML 可能存在放大基础输入数据中存在的偏见的风险（主要包括 3 类偏见，即人为偏见、系统偏见和统计 / 计算偏见），但尚未提供具体的监管措施 [4]。

EMA 指出数据采集与增强主要存在数据泄露的风险，会导致模型过度拟合；数据完整性方面主要存在数据未完全表征以及大型语言模型的记忆风险。EMA 在数据采集与增强、数据保护以及数据完整性方面给出了应对风险的具体措施和监管考虑 [2]，表述如下。

4.1 数据采集与增强

数据来源和数据采集，如清理、转换、插补、注释和规范化，应符合良好的临床、实验室和生产管理规范（GxP）要求，并以详细、可追溯的方式进行记录。探索性数据分析应描述数据特征、代表性、公平性和与预期任务的相关性，并在记录时注意以下 3 个方面：①数据的相关性和人群代表性，以及所做的内部 / 外部假设；②分类不平衡和相应的处理措施；③因数据产生的不公平或歧视性结果所导致的潜在风险。

数据增强可采用增强技术，并扩展训练数据集。当训练数据集存在限制，并影响了模型的可推广性或公平性时，应在模型文档中明确说明这些限制，并在模型不适用的情况下使用替代方法。用于高风险环境的模型（特别是用于临床的不透明模型），应使用新获得的数据进行前瞻性测试。

4.2 数据保护

申请人或 MAH 有责任确保所有的个人数据，包括 AI 和 ML 模型中间接持有的个人数据，都按照数据保护立法进行存储和处理。数据处理活动有以下原则要求：合法性；公平性和透明度；目的限制；数据最小化；准确性；存储限制；完整性和保密性；问责制；数据主体的权利；默认的数据保护原则。

在应用 AI 处理任何个人数据前，一般应预先对 AI 系统进行特定的风险评估；应说明并记录对数据主体权利和自由的可能影响，评估并证明符合上述原则。同时，需要评估预期使用个人数据的必要性和相称性。必要性评估应考虑使用匿名、合成数据或部署差异化隐私技术的可能性，如不可行，应说明理由；相称性评估应说明待处理数据的数量和类型是否充分符合数据最小化原则，并确定数据使用时侵入性最小的方法，以最大限度地减少对数据主体的影响。

4.3 数据完整性

若个人数据已用于模型训练，则必须进一步评估，通过成员身份、推理和模型逆向攻击是否可以提取此类信息，且需降低重新识别的风险。通过正则化、剔除和添加随机噪声，可提供部分到完全的匿名化处理，降低数据记忆风险。

5　AI 和 ML 模型在制药领域的应用及监管

EMA 和 FDA 均从模型应用的全过程提出了相应的风险和监管考虑；EMA 发布的内容较为具体，FDA 则相对笼统。EMA 从模型开发、性能评估、模型使用和模型部署方面介绍了对模型的监管考

虑，具体措施如下[2]。

5.1 模型开发

申请人或 MAH 有责任制定 SOP，促进模型通用性和稳健性的开发，特别是在部署期间不能更新模型的设置，应保存可追踪的文档和开发日志，以便对开发和实践进行二次评估。鼓励探索和实施通用性方法，包括正则化技术和基于日历时间对训练数据进行分层的敏感性分析。应清晰地描述模型的预期用途，以便对特征性项目进行有效评估。

5.2 性能评估

通常，指标集应包含对分类不敏感的参数，并描述完全混淆矩阵。为识别、训练和测试分割相关的随机效应，应提供交叉验证生成的性能指标分布。应支持模型对不同类别比例数据的通用性，并确保在部署期间面对不可控的长期趋势时具有稳健性，预期将对少数类别和与日历时间相关的因素进行敏感性分析。性能指标的预先定义阈值可与使用环境相关，以进一步支持模型性能的可信度。

5.3 可理解性与可解释性

模型使用应提供关于模型架构、超参数调整、训练指标、验证和测试结果的基本原理和详细信息，以及用于减轻非透明度问题的预定义监控和风险管理计划。为了进行审查和监测，应尽可能使用可解释 AI 的方法，这包括提供特征重要性列表、沙普利加性解释（SHapley Additive exPlanation，SHAP）和 / 或模型无关的局部解释（Local Interpretable Model-Agnostic Explanation，LIME）的分析或类似的可解释性指标，用于模型和部署期间的个别推断。计算机视觉模型以及使用注意力机制的其他方式，应尽可能提供注意力图的相关支持，确保特征是从图像中的相关位置提取的。

5.4 模型部署

AI/ML 模型的部署是基于风险的方法。对于高风险的案例，模型的软件和硬件堆栈中所有非平凡变更，包括关键依赖项的版本变更，都需要对模型性能进行重新评估。推理时的数据采集硬件、软件和数据转换通道应符合预定义的规范。应对模型性能进行监测，以便尽早发现退化，并明确可接受的阈值。用于人工分类的数据常规采样，或使用外部质量控制程序提供的测试数据集，应定期评估性能，特别是自主增量学习系统。对所有模型都应制定风险管理计划，定义算法失效的可能风险，例如不正确的预测 / 分类的后果、监测和减轻 / 纠正方法、如何触发模型的暂停 / 解除操作等。

FDA 认为在模型开发、性能监控和验证中主要存在的 2 个风险，即模型影响和决策后果，这 2 个因素则由模型的使用环境决定[6]。FDA 主要通过规范步骤和明确标准文件对模型开发和性能进行评估。在平衡性能和可解释性方面考虑复杂性，应优选传统和简单的模型。监控和记录 AI 和 ML 模型的结果时，应采取纠正措施确保预期结果。进行软件代码和计算的验证时，应验证模型并评估适用性，同时考虑模型预测的不确定性。

6 我国在制药领域 AI 应用的监管建议

6.1 建立指导机构和合作机制

建议药品监管机构建立专门的指导机构或工作组，与学术界、行业的其他利益相关者密切合作，召开研讨会，举办 AI 相关论坛，共同探索 AI 在药物开发中的应用，并制定指导方针和标准。2019 年 7 月 17 日，NMPA 医疗器械技术审评中心联合国家计算机网络与信息安全管理中心等 14 家单位共同发起成立 AI 医疗器械创新合作平台，以积极应对 AI 快速发展给监管和产业带来的风险和挑战，并加快 AI 成果在医疗器械领域的应用[7]。虽然我国目前尚未成立 AI 药品的指导机构或工作组，但 AI 医疗器械创新合作平台的监管经验可提供一定的参考和借鉴。

药品监管机构应充分借鉴医疗器械行业中 AI 应用方面的标准和监管指南，制定药物开发和使用的 AI 监管指南，涵盖责任主体和问责机制，数据质量可靠性和代表性，算法或模型的透明度、可理解性和有效性，伦理原则等方面内容。应积极推动生物医学领域ML 良好实践的国际发展和标准化进程[3]。

6.2 风险评估

监管机构需要制定基于风险的方法来评估和监管 AI。在评估 AI 和 ML 模型是否引入特定的风险

和危害时，应根据使用环境评估模型的风险和可信度[6]，考虑模型影响和决策后果等因素。在使用AI和ML模型时，应该根据模型的风险和使用程度来决定人的参与程度。人的参与程度应与模型风险、不确定性成正比[8]。通过制定风险管理计划，识别和减轻风险。对于高风险的AI软件，监管机构应具有访问所使用算法和底层数据集的权限。在整个AI/ML模型开发过程中，AI和ML专业领域的专家参与并提供的有价值见解和专业知识，可有助于降低相关风险[9]。可采用风险评估工具，如美国注射剂协会（Parenteral Drug Association，PDA）推荐采用的《良好自动化生产实践规范》第二版作为AI生命周期的风险评价工具[10]。艾昆纬公司（IQVIA）推荐了美国国家标准与技术研究所开发的人工智能风险管理框架[11]。对于影响药品的质量、安全性、疗效，并与其获益相关的数字工具，应基于风险定义主要和次要更新。自主学习的AI应用具有较高风险，应研究和记录此类系统的行为。

6.3 关于数据的监管考虑

在数据方面，申请人和利益相关方需要确保数据质量的可靠性和代表性，还需关注数据的来源、隐私和安全、完整性、相关性、重复性和再现性。应在数据采集、增强、保护和完整性方面采取措施，确保数据质量，如偏差的识别和管理等。源数据的多样性和可变性可降低偏差风险。可采用数据去标识、数据加密、限制数据访问及使用联合学习等措施，确保数据完整性、减少偏差、处理数据缺失和提高质量[8]。为确保重现性和复制性，应对数据和模型进行版本化控制，记录训练数据的顺序和随机因素，并在制药质量体系中实施严格的变更控制[8]。另外，在特定条件下，需考虑数据管理的责任、隐私和安全性，平衡数据完整性和数据去留的问题。

6.4 关于模型的监管考虑

总体上，申请人或利益相关方应对模型的开发、性能评估、透明度、部署和验证进行监控和记录，需要确保AI和ML模型的透明度、可解释性和可理解性[2]。应提供模型的相关信息，如识别透明度的阻碍和促进因素，并应明确记录、透明化和跟踪关键步骤，在设计AI技术之前发布或记录透明度要求。国际制药工程协会（The International Society for Pharmaceutical Engineering，ISPE）建议加强和实现模型可追溯性和可审计性的规程，包括模型版本化、模型与训练数据的关联、开发过程和见解的结构化说明，并定期根据药品质量体系审查模型。AI和ML模型的开发应遵循ICH Q8/Q9/Q10的要点，如确定模型目的、建模和实验方法，选择变量/输入，理解模型局限性和假设，记录数据采集，开发和验证模型，评估不确定性并减轻影响，归档模型开发、验证和生命周期管理计划，并在药品质量体系中管理模型的维护和变更[8]。验证AI/ML模型的方法时可考虑对数据集进行严格测试和评价，采用交叉验证技术评估模型的可推广性和稳健性，并将预测结果与既定标准或专家结论进行比较，以实施外部验证[9]。

模型的复杂性、性能和可解释性之间需保持平衡，当传统的简约模型与复杂模型的性能相似时，更倾向于选择传统模型[5]。另外，AI技术应该对开发人员、医疗专业人员、患者、用户和监管机构具有可理解性，并根据接受者的能力进行解释。应采用基于风险的方法处理可解释性问题，考虑复杂性和用户期望。选择模型前，应了解用户对不同可解释性的接受度。ISPE建议采用SHAP分解特征预测效果以增强模型的可解释性。同时，应使用统计检验评估数据偏差，记录模型选择过程，并将可解释性扩展至推理代码。开放源代码可用于模型部署，系统应完整记录代码/模型版本并进行全面验证。记录并传达模型预测中的不确定性评价很重要，具体方法包括使用点估计值或置信区间、设计受控措施和适用操作模式。应定期监测模型性能，并辅以定性手段，如用户接受/拒绝和反馈[8]。

6.5 建立药品全生命周期监管措施

在药品全生命周期可采取不同的措施。①药物发现阶段，主要审查模型和数据集的质量和完整性。②非临床研究阶段，需要将SOP扩展到所有AI和ML应用，并预先分析可能与药品收益-风险平衡评估相关的临床前数据。③临床研究阶段，需对模

型的架构、建模、验证和测试进行全面评估，并遵循适用的临床试验数据统计原则。④药品生产阶段，则主要关注云应用程序的数据完整性、模型的开发和验证，以及管理和监控持续学习的 AI 系统。⑤药品上市后阶段，应采取不同的策略降低风险。这些措施旨在确保 AI 和 ML 模型在药品全生命周期过程中的有效性和安全性，以及方便与药品监管机构的沟通和合作[2,5,12]。

参考文献：

[1] FDA. FDA releases two discussion papers to spur conversation about artificial intelligence and machine learning in drug development and manufacturing [EB/OL]. (2023-05-10) [2023-10-10]. https://www.fda.gov/news-events/fda-voices/fda-releases-two-discussion-papers-spur-conversation-about-artificial-intelligence-and-machine.

[2] EMA. Reflection paper on the use of artificial intelligence (AI) in the medicinal product lifecycle draft [EB/OL]. (2023-07-19) [2023-10-10]. https://www.ema.europa.eu/en/news/reflection-paper-use-artificial-intelligence-lifecycle-medicines.

[3] International Coalition of Medicines Regulatory Authorities. Informal innovation network horizon scanning assessment report—artificial intelligence [EB/OL]. (2021-08-06) [2023-10-10]. https://www.icmra.info/drupal/sites/default/files/2021-08/horizon_scanning_report_artificial_intelligence.pdf.

[4] FDA. Using artificial intelligence & machine learning in the development of drug & biological products [EB/OL]. [2023-10-10]. https://www.fda.gov/media/167973/download.

[5] FDA. Artificial intelligence in drug manufacturing [EB/OL]. [2023-10-10]. https://www.fda.gov/media/165743/download.

[6] SCHWARTZ R, VASSILEV A, GREENE K, *et al*. Towards a standard for identifying and managing bias in artificial intelligence [EB/OL]. [2022-03-15]. https://nvlpubs.nist.gov/nistpubs/SpecialPublications/NIST.SP.1270.pdf.

[7] 国家药品监督管理局. 人工智能医疗器械创新推进会在京召开 [EB/OL]. (2019-07-17) [2023-10-10]. https://www.nmpa.gov.cn/yaowen/ypjgyw/ylqxyw/20190717210701958.html.

[8] FDA. Comment from international society for pharmaceutical engineering (ISPE) [EB/OL]. [2023-08-01]. https://downloads.regulations.gov/FDA-2023-N-0743-0018/attachment_1.pdf.

[9] FDA. Comment from insilico medicine [EB/OL]. [2023-09-10]. https://downloads.regulations.gov/FDA-2023-N-0743-0036/attachment_1.pdf.

[10] FDA. Comment from parenteral drug association [EB/OL]. [2023-08-10]. https://downloads.regulations.gov/FDA-2023-N-0743-0041/attachment_1.pdf.

[11] FDA. Comment from IQVIA [EB/OL]. [2023-08-10]. https://downloads.regulations.gov/FDA-2023-N-0743-0063/attachment_1.pdf.

[12] 姚立新. 浅析人工智能 / 机器学习用于药品开发——基于 FDA 讨论文件的探讨 [J]. 中国食品药品监管，2023, (7): 44-57.