

《基于大语言模型的药物警戒个例安全性报告 人工智能辅助信息提取规范》团体标准 编制说明

一、任务来源

由上海市药品和医疗器械不良反应监测中心申请，中国药品监督管理研究会研究立项课题《大语言模型在中国药物警戒数据处理的应用研究》于2024年3月批准立项。其中标准题目为《基于大语言模型的药物警戒个例安全性报告人工智能辅助信息提取规范》，团体标准号为（T/SBIAORG 0002—2025）。

该团体标准由上海市药品和医疗器械不良反应监测中心提出，由上海市生物医药行业协会归口，于2024年9月由标委会完成团体标准立项。上海市药品和医疗器械不良反应监测中心、上海罗氏制药有限公司等标准起草单位成立了标准编制工作组，共同完成本次团体标准编制工作。

本文件主要起草人：许晋、胡骏、巢艾伦、杨依晗、栾国琴、黄天娇、袁博、步伐、浦义虎、韩澎、黎洋。

本标准为首次发布。

二、制定背景

2019年新修订的《中华人民共和国药品管理法》中提出国家建立药物警戒制度，而良好的药物警戒实践需要构建完整的生态系统。推动药物警戒生态圈各环节之间（包括药品上市许可持有人、经营企业、医疗机构、监管机构等）的沟通协同，实现数据、信息、知识和价值的有效流动，才能促进各方药物警戒能力提升，保障公众用药安全和健康福祉。其中，做好药品不良反应及其他与用药有关的有害反应的监测、识别、评价和报告是基础也是基石，是进一步开展风险管理的依据。

随着中国生物医药产业蓬勃发展和中国药物警戒制度建设完善，不良反应个例报告数量持续高速增长，给不良反应的识别和报告带来了巨大的挑战。目前药物警戒活动的各方仍主要依赖于人工和手动的方式来识别、处理和审阅不良反应个例报告（ICSRs）。从企业和医疗机构角度而言，这种方式成本高、耗时费力，且容易出现人为错误和主观判断，严重限制了药物警戒活动的效率和规模，也增加了不良反应识别和处理的风险。从监管机构的角度来看，不断涌人的不良反应个例报告（ICSRs）给不良反应监测机构带来了巨大的工作压力，远超已有的监测人力资源。

在全球范围内，监管机构和行业专家正积极探讨如何人工智能（AI）技术来提升药物警戒工作效能。特别是在自然语言模型（Natural Language Processing, NLP）方面，新一代大语言模型（Large Language Models, LLM）展现出卓越的语言理解和生成能力，

使得它在药物警戒领域，特别是不良事件报告自动化处理中，展现出了广阔的应用前景和极大的优势。例如，LLM技术可以用来自动化完成一些PV领域的基础任务（如不良反应报告的接收、录入和处理），降低相关的行政负担和成本，使不良反应监测工作更加高效。美国FDA就在药物警戒上市后安全性监测领域推出了一系列平台和工具，探索自然语言技术在不良事件监测、数据分析、哨点计划等方面的具体应用。

尽管LLM和AI技术在药物警戒领域具有潜力，但它们在应用过程中仍面临诸多合规风险，这也为监管带来新的挑战。使用LLM和AI技术主要风险体现在数据获取与标注的困难、语义理解与上下文处理的复杂性、模型的可解释性、数据隐私与安全等多个方面。全球主要药品监管机构，如美国食品药品监督管理局（FDA）和欧洲药品管理局（EMA），正通过发布指导文件和实施具体项目积极促进人工智能（AI）在药物警戒中的合规应用。例如欧盟最近发布了首个《人工智能法案》涵盖了生物医药领域，标志着该领域监管迈出了重要一步，与此同时，国际组织也在努力构建统一标准并加强跨国合作，共同探索AI技术在医药领域的机会与挑战。我国现有的监管要求大多集中于人工智能或自然语言处理模型的范畴内，尚未针对药物警戒领域出台监管细则，这些国际经验为我国的相关工作提供了重要的参考。

三、制定原则

本标准制定过程坚持了以下基本原则：

（一）规范化原则。遵循相关的标准化规则和要求，包括标准的结构、格式、符号、术语等方面。从而保证标准的统一性、规范性和易读性，以方便使用者理解和使用。

（二）适用性原则。注意标准的适用范围，既不要让标准所涵盖的领域过宽，使编制的标准没有实际技术内容；也不要让标准所涵盖的领域过窄，造成对标准的肢解，无谓地增加标准项目。

（三）先进性原则。制定标准时需确保标准所涉及的技术内容能够满足既定的需求。编写标准时会在充分调查研究的基础上，认真分析国内外同类技术标准的技术水平，在预期可达到的条件下，积极地把先进技术纳入标准，提高产品技术水平。

（四）统一性和协调性原则。编制过程中注意符合法律法规的规定以及与相关标准协调，避免与法律法规、相关标准之间出现矛盾，给标准的实施造成困难。

（五）经济性原则。制定标准时要以满足实际需要出发，不要一味地追求高性能、高指标，避免造成经济浪费。

四、制定过程

（一）启动阶段：2024年5月-2024年7月

标准制定项目组内成员的讨论和分工。调研和挖掘大语言模型在药物警戒领域的潜在应用价值及研究动态，梳理国内外相关监管

政策。

（二）草案阶段：2024年8月-2024年9月

《应用大语言模型自动提取个例安全性报告信息》标准草案编制。

（三）征求意见阶段：2024年9月-2024年12月

组织药物警戒药领域专家、信息化和人工智能领域专家等，进行讨论和研究，广泛征求和收集专家的意见。

选取3-5家已经使用或计划在药物警戒领域使用大语言模型的药品上市许可持有人、信息化技术提供商，进行深度交流和讨论，征求和收集行业意见。

五、标准主要结构和关键技术指标说明

1 范围

2 规范性引用文件

3 术语和定义

3.1 药物警戒 pharmacovigilance [来源：药品管理法]

3.2 大语言模型 Large Language Model, LLM [来源：Guiding principles on the use of large language models in regulatory science and for medicines regulatory activities]

3.3 生成式人工智能技术 generative artificial intelligence [来源：生成式人工智能服务管理暂行办法]

3.4 微调 fine-tuning [来源：GB/T 41867—2022]

3.5 多模态数据 multi-modal data [来源：GB/T 42135—2022]

3.6 知识 knowledge [来源: GB/T 23703.2-2010]

3.7 知识库 knowledge base

3.8 词元 Token[来源: ISO 24611:2012]

3.9 提示信息 prompt[来源: Guiding principles on the use of large language models in regulatory science and for medicines regulatory activities]

3.10 提示工程 prompt engineering

4 总体考量

5 大语言模型在个例安全性报告信息提取场景下的系统实施

5.1 模型部署和架构设计

5.2 模型性能评估

5.2.1 评估指标定义

5.2 评估指标计算 (不同的指标, 应用场景)

6 系统应用和质量控制

6.1 制定大语言模型使用策略

6.1.1 使用策略建议

6.2 标准操作流程 (SOP)

6.3 结果复核

6.4 定期回顾和更新

6.5 文档记录

6.6 培训

6.7 合规性

- 6.8 审计
- 6.9 业务持续性
- 7. 数据安全和隐私保护
 - 7.1 缓解措施
 - 7.2 基模型云服务提供商的报告 or 证书要求
 - 7.3 条件限制
- 8、利益相关方沟通

六、重大分歧意见的处理结果及理由

本标准制订编写过程中未产生重大分歧意见。

七、其他需要说明的情况

无

八、与现行法律、法规、文件和标准相关情况的说明

总体而言，我国现有的政策和法律鼓励人工智能技术在生物医药特别是药物警戒领域的应用，但目前尚无详细的法律法规或标准指南专门针对药物警戒领域中大语言模型的应用。

（一）我国相关政策情况说明

我国现有的政策和法律鼓励自然语言处理等人工智能技术在生物医药特别是药物警戒和用药安全性方面的领域的应用。2017 年 7 月，国务院发布了《新一代人工智能发展规划》，要求充分发挥信息化技术和人工智能的优势。2024 年 6 月，《国家药监局综合司关

于印发药品监管人工智能典型应用场景清单的通知》药监综函〔2024〕313号，列出了15个典型应用场景，其中明确了在药物警戒运用人工智能技术的可行性：“利用人工智能技术可以辅助监管人员开展不良反应和不良事件报告的评估工作”。2024年9月《国家卫生健康委员会办公厅关于印发卫生健康行业人工智能应用场景参考指引的通知》国卫办规划函〔2024〕420号，提出了十三个大方向，83个具体场景。其中（十）“人工智能+”药物研发 **70.智能药物临床试验辅助：**“患者入组后……进行智能预警并对需要判读的不良事件自动生成判读任务……进行严重不良事件上报及可疑且非预期严重不良反应上报”。**71.智能药品临床综合评价辅助：**“应用大数据、自然语言处理、机器学习和数据挖掘等人工智能技术，迅速抓取药品在治疗疾病中研究进展和评价情况，……深入挖掘真实世界数据，快速分析评估药品安全性、有效性、……等多维信息，自动完成数据的提取整合和定性定量分析，快速自动生成评价结果报告……”。

（二）我国相关法律法规情况说明

目前尚无详细的法律法规专门针对药物警戒领域大语言模型的应用。其中最相关的是2023年7月颁布的《生成式人工智能服务管理暂行办法》，将生成式人工智能服务监管纳入了《中华人民共和国个人信息保护法》、《中华人民共和国数据安全法》和《中华人民共和国网络安全法》这三部基础法律的框架之下。如果数据涉及跨境传输，还需符合国家互联网信息办公室正式发布的《促进和规

范数据跨境流动规定》的相关要求。

（三）我国相关现行标准情况说明

目前尚无详细的标准指南专门针对药物警戒领域中大语言模型的应用。目前仅查阅到中国信息通信研究院自 2023 年起牵头编制了大语言模型相关的企业标准。2024 年 6 月，工业和信息化部等四部门印发《国家人工智能产业综合标准化体系建设指南（2024 版）》要求加强人工智能标准化工作系统谋划。2024 年 3 月 25-27 日，由中国电子技术标准化研究院主办的《人工智能 预训练模型 第 1 部分 通用要求》等 8 项国家标准编制启动会在北京召开。2024 年，全国标准信息公共服务平台共发布 4 项推荐性国家标准，其中 11 月 28 日发布的 GB/T45081-2024 人工智能管理体系适用于各种规模、类型和性质的提供或使用人工智能系统产品或服务的组织，并对跨领域或跨行业使用人工智能管理体系的。

（四）国际法规标准的相关情况说明

在全球范围内，药物警戒（PV）领域中人工智能（AI）的应用正在受到广泛关注，各国和地区正积极构建相应的监管框架与应用实践，以确保 AI 技术的安全、有效使用。美国食品药品监督管理局（FDA）、欧洲药品管理局（EMA）等主要监管机构参与撰写、发布了多份指导文件或征求意见稿。此外，国际组织如 ICMRA、CIOMS 工作组以及 PIC/S 也在推动全球统一标准的建立，促进跨国界的交流合作，共同应对 AI 带来的机遇与挑战。

表 1 美、欧、国际组织涉及药物警戒领域的法规政策对比表

国家或地区	法规政策	目前阶段	发布时间
-------	------	------	------

美国	法规指南：《在药品和生物制品研发中的人工智能和机器学习》Using Artificial Intelligence and Machine Learning in the Development of Drug and Biological Products	征求意见	2023
	法规指南：《使用人工智能支持药品和生物制品监管决策的考虑因素——行业及其他利益相关方指南》Considerations for the Use of Artificial Intelligence to Support Regulatory Decision-Making for Drug and Biological ProductsGuidance for Industry and Other Interested Parties	征求意见	2025
欧盟	法律：《人工智能法案》（Artificial Intelligence Act）附件 3 包含医疗保健的公共服务	生效	2021
	法规指南：《2023 年-2028 年人工智能工作计划》Multi-annual artificial intelligence work plan 2023-2028	生效	2023
	思考性文件：《人工智能在药品生命周期中的使用的思考性文件》AI in medicinal product lifecycle reflection paper	生效	2024
	指导原则：《大型语言模型在药品监管科学及监管活动中的使用指导原则》Guiding principles on the use of large language models in regulatory science and for medicines regulatory activities	生效	2024
国际组织 CIOMS	指导原则：第 XIV 工作组药物警戒中的人工智能：在药物警戒领域使用人工智能的原则和指导方针 Principles and guidance for the use of artificial intelligence or intelligence augmentation in the field of pharmacovigilance	初稿撰写	预期 2025 征求意见
国际组织 ICMRA	报告：《前瞻性扫描评估报告——人工智能》Horizon Scanning Assessment Report - Artificial Intelligence	生效	2021

尽管不同地区在法规制定的进展和侧重点上存在差异，但普遍趋势是采取基于风险的方法来评估和管理 AI 系统，并强调透明度、责任机制、数据隐私及人类监督的重要性。美、欧、国际组织涉及药物警戒领域的指导原则具体要求对比见表 2.

表 2 美、欧、国际组织药物警戒领域指导原则具体要求对比

相关要求	美国	欧盟	国际组织
	管理原则		
人类监督	√	√	√ CIOMS (人类监督、治理)
明确的责任机制	√		√ CIOMS
透明度	√	√ 保持透明和准确	√ CIOMS、ICMRA
可解释性	√		√ CIOMS
基于风险的方法	√	√ 区分风险等级、 系统必须评估和减少 风险、 PV 系统纳入药物警 戒系统、质量控制措 施	√ CIOMS、ICMRA

必要时和监管机构沟通	√	√	
公平与公正			√ CIOMS
数据要求			
偏倚	√		
完整性	√		
隐私和安全	√		√ CIOMS
来源	√		√ ICMRA
相关性	√		
可复制性	√	√ 交叉验证	
可重复性	√		
代表性	√		
可靠性	√		√ ICMRA
模型的开发、验证、监控和记录			
模型验证	√ 根据风险模型进行可信度评估	√	√ CIOMS (有效性和稳健性)
监控和记录性能	√ 监控并记录 AI/ML 模型的表现、确认所需的证据水平和记录保存	√ 监控和记录模型性能 / 高风险系统维护日志	
模型开发	√ 考虑模型的实际应用、平衡性能和可解释性	√ PV 领域可以有更灵活的建模和部署方法、通过增量学习持续提升效果	

九、实施标准的措施建议

强化标准的学习，抓好系统功能设计、应用环节。在标准发布后，应积极开展对本标准的宣贯，以多种形式组织学习和研究。系统应用单位可以依据本标准的功能规范对系统的开发商提出相应功能设计、研发的要求。

(完)