

ICS 点击此处添加 ICS 号

CCS 点击此处添加 CCS 号

T/ZSMM

浙江省数理医学学会团体标准

T/ZSMM XXXX—2025

医学人工智能治理综合评价指南 第 1 部分：总则

Guideline for comprehensive evaluation of medical artificial intelligence
governance—Part 1:General specification

（征求意见稿）

（本草案完成时间：2025 年 11 月 28 日）

在提交反馈意见时，请将您知道的相关专利连同支持性文件一并附上。

XXXX – XX – XX 发布

XXXX – XX – XX 实施

浙江省数理医学学会 发布

目次

前言 II

1 范围 1

2 规范性引用文件 1

3 术语和定义 1

4 基本原则 1

 4.1 安全有益 1

 4.2 动态评价 1

 4.3 短期与长期结合 1

 4.4 整体与部分结合 1

5 评价机构和人员 1

 5.1 评价机构 1

 5.2 评价人员职业种类和等级 2

6 评价流程 2

 6.1 成立评价组 2

 6.2 评价对象定义描述符合性评价 2

 6.3 按评价内容模块分类评价 3

 6.4 评价材料收集 3

 6.5 形成评价意见与结论 3

 6.6 资料归档 3

 6.7 评价结果管理与应用 3

7 评价指标体系 3

 7.1 评价指标体系框架 3

 7.2 安全评价 4

 7.3 风险评价 4

 7.4 效用评价 4

 7.5 效率评价 4

 7.6 效益评价 5

前 言

本文件按照GB/T 1.1—2020《标准化工作导则 第1部分：标准化文件的结构和起草规则》的规定起草。

《医学人工智能治理综合评价指南》分为以下六个部分：

- 第1部分：总则；
- 第2部分：安全评价；
- 第3部分：风险评价；
- 第4部分：效用评价；
- 第5部分：效率评价；
- 第6部分：效益评价；

本部分为《医学人工智能治理综合评价指南》的第1部分。

本文件由浙江省数理医学学会提出并归口。

本文件起草单位：南方医科大学、南方科技大学、深圳市人民医院、浙江数字内容研究院、深圳市卫生健康委员会、深圳市卫生健康发展研究与数据管理中心、上海市第六人民医院、中国医学科学院医学信息研究所、南方医科大学第三附属医院、南方医科大学珠江医院、南方医科大学第八附属医院、南方医科大学南方医院赣州医院、南方医科大学中西医结合医院、东莞市石碣医院、广东医科大学附属医院、深圳市第四人民医院、深圳市妇幼保健院、四川大学华西第二医院、香港大学深圳医院、中山市人民医院、珠海市人民医院、佛山市第一人民医院、前海人寿广州总医院、广州市红十字会医院、北京大学深圳医院。

本文件主要起草人：王冬、姜虹、毛燕娜、耿庆山、汤昊宸、朱春艳、许彬彬、丁万夫、郑静、曹艳林、吕群蓉、崔书亭、张冬云、李晨程、陈宝颖、潘鑫、吴超梅、王亚琴、郭洪波、曹蓓、戴辉、杜庆锋、刘仲文、蔡定彬、王诚、李笑天、张少毅、徐小平、黄晓星、郭煜、段光荣、周宏峰、张立贤、赵永胜。

医学人工智能治理综合评价指南

第1部分：总则

1 范围

本文件确立了开展医学人工智能治理综合评价的基本原则，提供了评价机构和人员、评价流程和评价内容等方面的指导与建议。

本文件适用于组织开展医学人工智能技术对社会活动产生的现实或潜在影响进行综合评价。

2 规范性引用文件

下列文件中的内容通过文中的规范性引用而构成本文件必不可少的条款。其中，注日期的引用文件，仅该日期对应的版本适用于本文件；不注日期的引用文件，其最新版本（包括所有的修改单）适用于本文件。

DB4403/T 634—2025 医学人工智能治理综合评价指标体系

3 术语和定义

下列术语和定义适用于本文件。

3.1

医学人工智能治理 medical artificial intelligence governance

通过制定和执行原则、标准、指南与政策法规等，促进医学人工智能发展的过程。

3.2

医学人工智能治理评价 evaluation of medical artificial intelligence governance

通过治理评价指标体系考量医学人工智能治理的活动。

注：包括考量医学人工智能的安全与效用现状、风险问题现状、卫生经济问题现状等；不包括医学人工智能技术赋能社会治理活动的评价。

4 基本原则

4.1 安全有益

重点保障人的安全和尊严，同时做好评价数据及相关资料的保密工作，保障相关方的合法权益。

4.2 动态评价

在不同时态（如事前、事中、事后）和不同发展阶段（如试验期、发展期、成熟期）进行动态评价，并基于评价结果动态调整评价活动。

4.3 短期与长期结合

立足现有数据和当前现实情况，制定短期与长期评价目标，选择适宜的评价内容进行评价。

4.4 整体与部分结合

以整体为统领，以各类为支撑，根据各部分的不同特点，寻找适宜的实验场景切入，以点带面，统筹推进评价工作的落实。

5 评价机构和人员

5.1 评价机构

评价机构宜由相关利益方和相关监管与治理部门共同组成，可包括：

- 卫生健康管理、政务服务数据管理、医疗保障管理、市场监督管理、工业和信息化管理、科技创新管理、人力资源与社会保障等政府职能部门；
- 医疗卫生等服务机构；
- 高校、科研院所等科研机构；
- 医师协会、医院协会、人工智能学会等社会团体。

5.2 评价人员职业种类和等级

5.2.1 评价人员职业种类

职业种类宜基于医学人工智能治理评价工作的内在需求、多学科交叉特点及应用场景规律的原则进行分类，见表1。

表1 职业种类

序号	职业种类
01	信息技术服务从业人员（包括但不限于信息技术咨询服务、设计与开发服务、信息系统集成实施服务、运行维护服务、数据处理和存储服务、运营服务、数字内容服务、云计算服务、信息安全服务、其他信息技术服务）
02	专业技术人员（包括但不限于哲学研究人员、数学研究人员、医学研究人员、管理学研究人员、经济学研究人员、法学研究人员、医学设备管理工程技术人员、通信工程技术人员、信息管理工程技术人员、数据分析处理工程技术人员、卫生专业技术人员、其他卫生专业技术人员）
03	国家机关、群众团体、社会组织、企事业单位负责人（包括但不限于国家行政机关负责人、社会团体负责人、企业负责人、事业单位负责人）
04	办事人员（包括但不限于行政业务办理人员、行政事务处理人员、其他办事人员）

5.2.2 职业等级分级方法

在职业种类划分的基础上，可根据医学人工智能治理评价工作的需求，从业人员的职业发展客观规律以及将评价人员职业等级划分为五个等级，作为其能力评价的依据。组织根据自身情况及评价项目特征，结合下表要求对各级人员的能力进行细化与量化。表2给出了评价人员等级分级方法。

表2 职业等级分级方法

职业等级	等级要求
5级	能领导完成高度复杂或开创性的医学人工智能治理评价项目，精通关键领域的评价技术与方法。能够在评价理论、体系或标准上实现创新，并主导其制定与推广，对行业发展产生深远影响。
4级	能独立领导并完成复杂的多模块评价任务，掌握核心评价技能。能够有效带领团队解决评价过程中的疑难问题，提供专业指导，并参与评价制度与流程的优化工作。
3级	能独立承担特定专业领域（如安全、伦理、临床等）的评价工作，并进行深度分析。具备指导初级人员的能力，是保障评价质量与进度的核心骨干。
2级	能在他人指导下，独立完成评价任务中的数据收集、初步分析与报告撰写等具体工作。能准确理解并有效执行既定的评价流程与规范。
1级	能在明确指导下完成评价相关的资料整理与辅助性工作，初步掌握评价基础知识，并通过实践逐步积累经验。

6 评价流程

6.1 成立评价组

宜根据第5章的内容成立评价组，并配置1名组内秘书和1名调查协调员。

6.2 描述评价对象符合性评估

宜给出评价对象是否为人工智能技术医疗设备或人工智能器械的符合性描述。评估不符合时，可终止评价。

6.3 按评价内容模块分类评价

宜根据评价对象的应用场景进行分类评价。评价指标体系设置参见第7章。

6.4 评价材料收集

评价材料宜根据评价内容所需选择，可包括研发合同、伦理报告、研发方案、软件后台数据、医疗卫生机构管理平台数据、区域卫生统计年鉴、社会实验的实验数据和调查数据等。评价组秘书在开展评价前宜对评价材料进行形式审查。宜关注收集材料的数据质量，包括：

- 安全性：
 - 根据授权使用数据，防止数据窃取和篡改；
 - 加强数据采集与评价的环境、设施和技术的安全管控措施，防止数据泄露；
 - 对数据进行定位溯源，并对过程、手段进行安全审计及检测。
- 完备性：
 - 根据评价方法进行数据完备性评价，确保数据资料详实、完备；
 - 数据不充分的评价指标，宜做出风险性提示。
- 一致性：采集的资料与实施评价的细分指标宜具有高度一致性，具有逻辑上高度置信的因果联系。
- 准确性：采集的资料可充分、真实的反应数据所描述的实体特征。
- 时效性：采集的资料宜反映待评价对象当前的状态，如评价周期较长，宜进行采集数据资料的动态版本控制。

6.5 形成评价意见与结论

根据评价内容形成评价对象关于医学人工智能治理综合评价的评价意见和结论。

6.6 资料归档

宜整理医学人工智能治理综合评价活动的全过程资料，归档留存。

6.7 评价结果管理与应用

整体结果未得到评价组专家2/3及以上票数通过时，医学人工智能产品宜根据评价组整改意见进行整改。1年内完成整改后，可重新评价，仍未通过者，可不予推荐。

7 评价指标体系

7.1 评价指标体系框架

宜参照DB4403/T 634—2025设置评价指标体系，包括一级指标5个、二级指标18个，见图1。

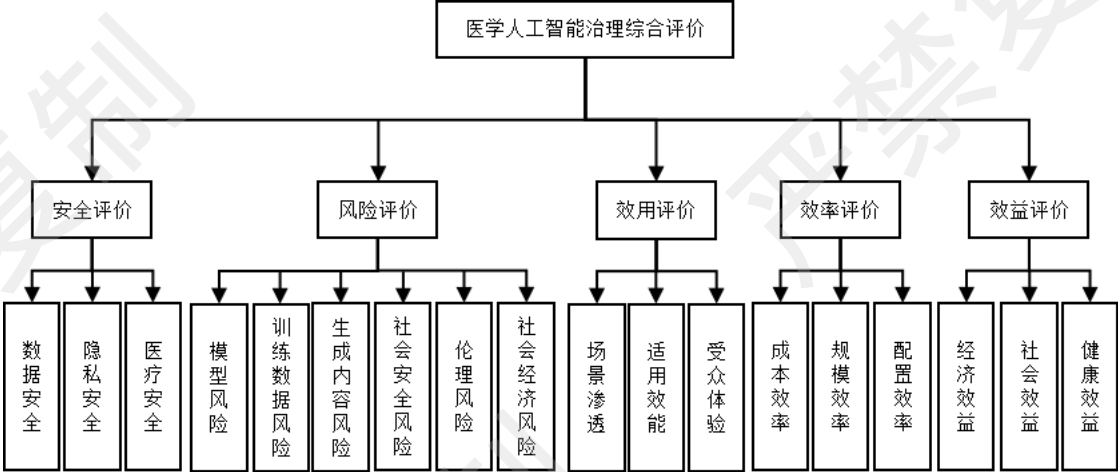


图1 医学人工智能治理综合评价指标体系架构图

- 注1：技术实验期治理评价是指某种医学人工智能软件涉及相关算法仍处于技术测试与开发阶段，其重点在安全评估与风险评估，一般发生在医学人工智能软件未获得医疗器械注册且未被应用于现实医学场景前的事前评价。
- 注2：技术发展期治理评价是指某类医学人工智能软件涉及相关算法正处于同类算法优化、微调、快速迭代的技术应用阶段，其重点在安全评估、风险评估、效用评估，一般发生在医学人工智能软件未获得医疗器械注册且正被应用于现实医学场景中试用的事中评价。
- 注3：技术成熟期治理评价是指某类医学人工智能软件涉及相关算法已适用于医疗卫生多场景技术应用的阶段，其重点在效率评估和效益评估，一般发生在医学人工智能软件获得医疗器械注册已被广泛应用于现实医学场景中的事后评价。
- 注4：动态评估是既包括在不同时态（前、中、后）不同发展阶段（试验期、发展期、成熟期）进行评估，也包括基于不同时态不同阶段的评估结果反馈进行动态整改的评估过程。

7.2 安全评价

安全评价是对评价对象在研发、使用、推广的过程中的数据安全、隐私安全以及医疗安全进行的综合评价指标。

- 数据安全是对评价对象开发与应用过程中涉及健康医疗数据的安全管理的评价指标。
- 隐私安全是对评价对象保护个人敏感信息的处置方式进行评价的评价指标。
- 医疗安全是对评价对象在医学场景应用过程中对医疗卫生服务质量带来的影响进行评价的评价指标。

7.3 风险评价

风险评价是在风险管理策划的基础上，对风险进行定性或量化分析的过程，旨在确定风险发生的可能性和影响程度。是对评价对象系统本身所具有的威胁安全、施加影响或逃避监督的可能性、以及对社会、伦理、法制等其他方面产生的伤害事件预期、系统对人的控制程度和对人工智能系统结果的控制及纠错程度等进行评价的评价指标。

- 模型风险是对评价对象的算法模型稳健性、泛化性、鲁棒性、可解释性、可信任性等进行测试，以评价算法的不确定风险的评价指标。
- 训练数据风险是对评价对象所使用训练数据的科学性、有效性、合规性以及制定不良训练数据限制规则和防范毒性训练数据的污染等方面进行评价的评价指标。
- 生成内容风险是对评价对象的输出结果是否有包含违反社会主义核心价值观、歧视性、商业违法违规、侵犯他人合法权益的内容和是否满足特定服务类型方面进行评价的评价指标。
- 社会安全风险是对评价对象在技术使用中潜在的社会安全风险进行评价的评价指标。
- 伦理风险是对评价对象在社会伦理方面影响社会公平与稳定，以及对个体、人类社会、环境等方面造成潜在威胁和危害的风险进行评价的评价指标。
- 社会经济风险是对评价对象在社会经济方面影响社会与对经济发展的潜在风险进行评价的评价指标。

7.4 效用评价

效用评价是对评价对象在研发、使用、推广的过程中的应用场景渗透广度与深度、场景契合程度、受众体验、硬件环境影响以及适用效能进行评价的评价指标。

- 场景渗透是对评价对象在可适用医疗卫生机构中使用的广度深度进行评价的评价指标。
- 适用效能是对评价对象在医学场景应用过程中，与医疗业务情境和业务流程的场景契合程度进行评价的评价指标。
- 受众体验是反映用户对评价对象依赖程度呈现的满意度、不可或缺性以及优先关注程度进行评价的评价指标。

7.5 效率评价

效率评价是采用数据包络分析方法、随机前沿分析法等效率评价算法模型对评价对象应用于医疗领域前后对医疗服务供给的技术效率、规模效率、成本效率、配置效率、规模收益状态以及医疗资源配置状态产生的影响进行评价的评价指标。

- 成本效率是采用数据包络分析法、随机前沿分析法等效率评价算法模型对评价对象应用于医疗领域前后带来医疗资源投入成本变化对医疗服务供给产生的影响进行评价的评价指标。

- 规模效率是采用数据包络分析法、随机前沿分析法等效率评价算法模型对评价对象应用于医疗领域前后医疗资源投入规模变化对医疗服务供给的规模收益状况的影响进行评价的评价指标。
- 配置效率是采用数据包络分析法、随机前沿分析法等效率评价算法模型对评价对象应用于医疗领域前后医疗资源投入成本与规模变化对医疗服务供给产出最优状态的影响进行评价的评价指标。

7.6 效益评价

效益评价是对评价对象产生的经济效益、社会效益以及健康效益进行评价的评价指标。

- 经济效益是对评价对象在医疗卫生服务供给方面带来的经济边际效益影响进行评价的评价指标。
- 社会效益是对评价对象在医疗卫生领域推广使用后，对社会人群健康和社会健康发展带来的影响进行评价的评价指标。重点关注对传染性疾病和慢性非传染性疾病监测、预警、规范管理带来的影响，以及推广使用增加的成本所带来的健康劳动力收益增量情况和非健康劳动力收益损失情况。
- 健康效益是对评价对象在医疗卫生领域推广使用后，对某区域人群、不同年龄结构人群、某疾病患病人群健康状况的影响进行评价的评价指标。重点测量对某疾病患者人群的伤残调整寿命年和某区域全人群的伤残调整寿命年的影响。

参 考 文 献

- [1] GB/T 20001.8—2023 标准起草规则 第8部分：评价标准
- [2] GB/T 35273—2020 信息安全技术 个人信息安全规范
- [3] GB/T 37373—2019 智能交通 数据安全服务
- [4] GB/T 37696—2019 信息技术服务 从业人员能力评价要求
- [5] GB/T 39720—2020 信息安全技术 移动智能终端安全技术要求及测试评价方法
- [6] GB/T 39725—2020 信息安全技术 健康医疗数据安全指南
- [7] GB/T 41867—2022 信息技术 人工智能 术语
- [8] GB/T 44811—2024 物联网 数据质量评价方法
- [9] GB/T 45288.2—2025 人工智能 大模型 第2部分：评测指标与方法
- [10] GB/T 45288.3—2025 人工智能 大模型 第3部分：服务能力成熟度评估
- [11] GB/T 45654—2025 网络安全技术 生成式人工智能服务安全基本要求
- [12] GB/T 45674—2025 网络安全技术 生成式人工智能数据标注安全规范
- [13] GB/T 45652—2025 网络安全技术 生成式人工智能预训练和优化训练数据安全规范
- [14] GB/T 45438—2025 网络安全技术 人工智能生成合成内容标识方法
- [15] GB/T 46071—2025 数据安全技术 数据安全和个人信息保护社会责任指南
- [16] GB/T 45577—2025 数据安全技术 数据安全风险评估方法
- [17] GB/T 45574—2025 数据安全技术 敏感个人信息处理安全要求
- [18] GB/T 46347—2025 人工智能 风险管理能力评估
- [19] JR/T 0197—2020 金融数据安全 数据安全分级指南
- [20] JR/T 0221—2021 人工智能算法金融应用评价规范
- [21] JR/T 0223—2021 金融数据安全 数据生命周期安全规范
- [22] JR/T 0287—2023 人工智能算法金融应用信息披露指南
- [23] MZ/T 165—2020 居民家庭经济状况核对 数据安全要求
- [24] NY/T 4261—2022 农业大数据安全管理指南
- [25] YD/T 3801—2020 电信网和互联网数据安全风险评估实施方法
- [26] YD/T 3865—2021 工业互联网数据安全保护要求
- [27] YD/T 3956—2021 电信网和互联网数据安全评估规范
- [28] YD/T 4043—2022 基于人工智能的多中心医疗数据协同分析平台参考架构
- [29] YD/T 4070—2022 基于人工智能的接入网运维和业务智能化场景与需求
- [30] YD/T 4392.1—2023 人工智能开发平台通用能力要求 第1部分：功能要求
- [31] YD/T 4921—2024 人工智能医疗器械 冠状动脉CT影像处理软件 算法性能测试方法
- [32] YD/T 4960—2024 移动智能终端可信人工智能安全指南
- [33] YY/T 1833（所有部分） 人工智能医疗器械 质量要求和评价
- [34] YY/T 1858—2022 人工智能医疗器械 肺部影像辅助分析软件 算法性能测试方法
- [35] DB11/T 2251—2024 信息安全 人工智能数据安全通用要求
- [36] DB37/T 4845—2025 人工智能技术应用伦理风险的治理要求
- [37] DB52/T 1726—2023 糖尿病视网膜病变人工智能筛查应用规范
- [38] DB4403/T 643—2025 医学人工智能治理综合评价指标体系
- [39] European Union. General Data Protection Regulation[Z]. Geneva: EU, 2018
- [40] 全国人民代表大会常务委员会. 中华人民共和国个人信息保护法：主席令（2021）91号. 2021年
- [41] 全国人民代表大会常务委员会. 中华人民共和国数据安全法：主席令（2021）84号. 2021年
- [42] 国家卫生健康委员会规划与信息司，国家卫生健康委员会统计信息中心. 全国医院信息化建设标准与规范（试行）：国卫办规划发（2018）4号，2018年
- [43] 国家卫生健康委，国家中医药管理局. 全国基层医疗卫生机构信息化建设标准与规范（试行）：国卫规划函（2019）87号. 2019年

[44] 国家互联网信息办公室, 中华人民共和国国家发展和改革委员会, 中华人民共和国教育部, 中华人民共和国科学技术部, 中华人民共和国工业和信息化部, 中华人民共和国公安部. 生成式人工智能服务管理暂行办法: 国家广播电视总局令第15号. 2023年

[45] 国家卫生健康委办公厅. 关于印发医疗机构临床决策支持系统应用管理规范(试行): 国卫办医政函(2023)268号. 2023年

[46] 深圳市第七届人民代表大会常务委员会. 深圳经济特区数据条例: 深圳市第七届人民代表大会常务委员会公告(第十号). 2022年

[47] 中国劳动和社会保障部. 中华人民共和国职业分类大典. 2025年
