

《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》推荐性国家标准编制说明

一、工作简况

（一）任务来源

按照《国家标准化管理委员会关于下达2024年第七批推荐性国家标准计划及相关标准外文版计划的通知》（国标委发〔2024〕44号）要求，本推荐性国家标准项目由国家药品监督管理局提出并归口，项目计划号20242922-T-464，国家药品监督管理局负责组织该标准的起草工作，由中国食品药品检定研究院（国家药品监督管理局人工智能医疗器械标准化技术归口单位）具体牵头完成。

（二）标准起草单位及主要起草人

标准主要起草单位：中国食品药品检定研究院、海军军医大学第二附属医院（上海长征医院）、辽宁省医疗器械检验检测院、国家药品监督管理局医疗器械技术审评中心、中国人民解放军总医院、北京大学、中国科学技术大学苏州高等研究院、重庆大学、中山大学、上海市医疗器械检验研究院、北京市医疗器械检验研究院（北京市医用生物防护装备检验研究中心）、山东省医疗器械和药品包装检验研究院、湖北省医疗器械质量监督检验研究院、上海西门子医疗器械有限公司、飞利浦（中国）投资有限公司、上海联影智能医疗科技有限公司、北京安德医智科技有限公司、广州柏视医疗科技有限公司、杭州健培科技有限公司、国药集团医疗器械研究院有限公司、河南省药品医疗器械检验院（河南省疫苗批签中心）等。

（三）起草过程

1. 成立工作组

2024年11月以来，牵头单位公开征集起草单位并成立了标准起草组，制定了实施计划。起草组按计划进行了资料收集、工作分工。

2. 组织起草

本文件的起草过程是对IEC 63524 ED1 标准的等同转化。该国际标准在2023年10月立项，立项草案基于我国行业标准YY/T 1858-2022。IEC 63524 ED1目前仍处于起草阶段，在2025年2月完成了委员会征求意见。

2024年11月以来，本文件起草组对IEC 63524 ED1的委员会草案稿进行了等同翻译和编辑性修改。同时，起草组围绕国内相关产品的注册情况、算法测试情况进行了调研，对标准草案的技术内容进行了对照论证，确认其适用于国内产品。

二、 编制原则、推荐性国家标准主要技术要求的依据（包括验证报告、统计数据等）及理由

（一）编制原则

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》和GB/T 1.2—2020 《标准化工作导则 第2部分：以ISO/IEC标准化文件为基础的标准化文件起草规则》给出的规则起草。

（二）标准主要技术要求的依据

本标准在制定过程中，梳理了国内外已上市相关产品的监管法规要求和公布的技术验证方式，测试流程参考了IEC 63450、IEC 63521等人工智能医疗器械国际标准，测试方法借鉴了ISO/IEC TR 29119-11等通用AI领域的标准化文件，调研了学术界对相关指标的定义和计算方法。

本文件规定了对采用人工智能技术的肺部影像辅助分析软件的算法性能测试方法。

本文件适用于采用人工智能技术、具有辅助诊断、辅助检测、辅助筛查、辅助分诊、优先级评定、随访跟踪等后处理功能的肺部影像辅助分析软件。算法的输入模态为肺部放射影像，包括但不限于X线、CT、MRI。

本文件不适用于影像前处理及过程优化产品。

本标准包括5个章节，分别为：

- 1 范围；
- 2 规范性引用文件；
- 3 术语和定义；
- 4 缩略语；
- 5 概述；
- 6 算法性能测试方法。

此外，本标准包括一个规范性附录，对算法测试的准备阶段进行规范。其内容在原YY/T 1858-2022标准第4章的基础上，按照IEC 63524 ED1的起草情况进行了修改，主要包括：

- 原黑盒测试、白盒测试等内容整合为合成数据在算法测试中的使用要求；
- 测试平台等概念整合为测试工具。

本标准包含两个资料性附录，为标准的实施提供参考信息。其中，附录B给出

了胸部CT肺结节测试数据集描述样例，用于帮助测试人员规范对测试集的介绍。附录C给出了测试指标选取与统计分析的一般思路，作为对正文的补充解释。

正文内容考虑了算法性能测试的指标定义、步骤和方法；附录的内容作为补充、举例和解释。在本标准制定的过程中，未涉及有争议项目。

（三）验证情况

本标准作为试验方法标准，根据人工智能算法在肺部影像辅助分析中的应用场景进行技术验证。本次技术验证采用肺结节辅助检测软件作为试验样品，依托中检院胸部CT肺结节标准测试集开展具体测试，测试指标覆盖了检出、分类、分割、测量等场景。

规范性附录A主要是对测试流程的验证，包括过程验证、工具验证、检查记录等实现形式；第6章是对具体测试方法的试验验证，依托相关AI产品和算法模型，在典型应用场景下输入测试集，对AI输出的结果进行分析，在病例层面、病灶层面形成定量的测试结果，客观反映了该产品在检出、分类、分割、测量等各种应用情景下的算法性能指标，同时对鲁棒性、错误统计等内容进行了验证，丰富了对产品算法性能的认识。上述内容可视为对临床试验的补充，为监管决策提供更多的证据。

本标准有助于统一肺部影像辅助分析软件的算法性能指标定义，提升不同产品之间的可比性，规范各企业、检测机构开展算法性能测试的流程、方法和工具，支撑产品各个阶段的质量评价需求。本标准有助于支持多中心测试资源的联动、整合、优化，避免测试集/测试工具的重复建设，节约测试与维护成本，降低企业负担。

三、与有关法律、行政法规和其他强制性标准的关系，配套推荐性标准的制定情况

本标准与我国有关的现行法律、法规和强制性国家标准无冲突。本标准不属于系列标准，不涉及配套的推荐性标准。

四、与国际标准化组织、其他国家或者地区有关法律法规和标准的比对分析

本标准在起草过程中借鉴了IEC 63450、IEC 63521、IEC 62304、ISO 24971-2等人工智能医疗器械相关国际标准的通用思路，提出了肺部影像辅助分析软件算法性能测试的具体方法。本标准与其他国家或地区有关法律法规不冲突。

五、重大分歧意见的处理过程、处理意见及其依据

本标准起草阶段未产生分歧。

六、标准性质（强制性、推荐性）的建议

由于人工智能医疗器械的核心算法、预期用途和应用场景处于快速发展变化中，本标准的技术内容仅代表当前的认识，根据未来业态可能进行更新。为了避免过度限制技术创新，建议本标准为推荐性国家标准。

七、贯彻标准的要求和建议措施

本标准发布后，中国食品药品检定研究院将及时开展标准宣贯，帮助相关各方尽快理解和使用本标准提出的测试方法。

本标准预期不会引发医疗器械变更注册。考虑到标准宣贯、人员培训、检验能力扩项等各个环节需要的周期，建议标准发布一年后实施。

九、废止现行有关标准的建议

本文件自实施之日起，代替行业标准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》。

十、涉及专利的有关说明

目前未收到相关专利的反馈意见。

十一、推荐性国家标准所涉及的产品、过程或者服务目录

涉及的产品主要为采用人工智能技术的肺部影像辅助分析软件，例如肺结节辅助检测软件、肺炎辅助分诊评估软件等。《医疗器械分类目录》为21-04决策支持软件。

十二、其他应当予以说明的事项

本文件的内容可能根据IEC 63524 ED1的后续起草情况进行动态更新，以确保与国际标准保持等同。

国家药品监督管理局

2025年3月4日