

受理号：CQZ2301701

医疗器械产品注册技术审评报告

产品中文名称：颅内动脉瘤磁共振图像辅助检测软件

产品管理类别：第三类

申请人名称：复影（上海）医疗科技有限公司

国家药品监督管理局

医疗器械技术审评中心

目 录

基本信息.....	3
一、 申请人名称.....	3
二、 申请人住所.....	3
三、 生产地址.....	3
技术审评概述.....	4
一、 产品概述.....	4
二、 临床前研究概述.....	4
三、 临床评价概述.....	6
四、 产品受益风险判定.....	11
综合评价意见.....	13

基本信息

一、申请人名称

复影（上海）医疗科技有限公司

二、申请人住所

中国（上海）自由贸易试验区临港新片区环湖西二路 888

号 C 楼

三、生产地址

上海市青浦区华青路 1269 号 101-104 室

技术审评概述

一、产品概述

(一) 产品结构及组成

该产品由安装程序、外部环境软件安装程序组成，功能模块包括用户登录、病例浏览、动脉瘤辅助检测模块（基于深度学习算法）。

(二) 产品适用范围

该产品用于对颅脑磁共振造影图像的显示、处理、测量和分析，可对3mm及以上颅内囊性动脉瘤辅助检测，供经培训合格的医师使用，不能单独用作临床诊疗决策依据。

(三) 型号/规格

AIneurysm，发布版本 1。

(四) 工作原理

该产品基于颅脑磁共振造影图像，利用医学图像处理技术和深度学习算法，自动提取颅内动脉血管区域，在该区域检测疑似颅内囊性动脉瘤，并将疑似病灶的区域进行展示。医生可手动编辑颅内囊性动脉瘤信息，最终形成文本报告。

二、临床前研究概述

(一) 产品性能研究

该产品性能指标包括软件功能、使用限制、输入输出、接口、运行环境、性能效率、最大并发数、用户界面、消息、用户差错防御、访问控制、版权保护、可靠性、维护性。

申请人针对上述性能指标提交了产品性能研究资料，同时提交了产品技术要求与产品检测报告，检测结果与产品技术要求相符。

(二) 产品有效期和包装

该产品使用期限为 10 年，申请人提交了包装验证资料，证实包装性能满足设计要求。

(三) 软件研究

该软件发布版本为 1，完整版本为 1.0.0.20240116，软件安全性级别为严重。

申请人依据《医疗器械软件注册审查指导原则（2022 年修订版）》提交了相应级别的自研软件研究资料、外部软件环境评估报告、互操作性研究报告、GB/T 25000.51-2016 测试报告，证明该软件设计开发过程规范受控，综合剩余风险均可接受。

申请人依据《医疗器械网络安全注册审查指导原则（2022 年修订版）》提交了相应级别的网络安全研究报告，证明该软件现有网络安全风险可控，已制定网络安全事件应急响应预案。

申请人依据《人工智能医疗器械注册审查指导原则》提交了基于深度学习技术的颅内动脉瘤检测算法的研究报告，包括训练集信息、训练数据量-评估指标曲线、训练结果、数据标注质控、算法性能评估、算法性能影响因素分析等研究资料，证实软件算法性能能够满足设计要求。

三、临床评价概述

申请人选择临床试验路径进行临床评价，临床试验的目的为评价申报产品临床使用的安全性与有效性。

临床试验采用回顾性、多中心、全交叉多阅片者多病例、优效性的试验设计，三家临床试验中心筛选符合标准的 MRI 影像病例，MRI 影像病例入组完成后进行动脉瘤检出阅片试验，由全部阅片医生各自对全部入组 MRI 影像病例进行两轮全层面阅片（医生单独阅片与 AI 辅助医生阅片，两轮阅片之间间隔不少于一个月的洗脱期）并记录所有主观判定为动脉瘤的病灶，以专家组的动脉瘤标记结果为金标准，验证医生单独阅片结果（对照组）与 AI 辅助医生阅片结果（试验组）对动脉瘤检出的安全性与有效性。受试医生进行两阶段的阅片，第一阶段阅片完成后，需要至少 4 周的洗脱期，再进行第二阶段的阅片；使用软件辅助阅片时，需对软件性能进行评价。本试验阳性病灶及病例定义如下：1.以病灶为单位：阳性动脉瘤定义：病灶最大

直径 $\geq 3\text{mm}$ 的动脉瘤；2.以病例为单位：阳性动脉瘤病例定义：此病例可见直径 $\geq 3\text{mm}$ 的动脉瘤；临床试验在三家临床机构开展，计划入组486例，实际入组486例受试者，其中，FAS集、PPS集均为484例。

临床试验的主要有效性评价指标为：以专家组标记结果为参考金标准，通过AFROC-AUC指标评价AI软件对MRI影像病例中的动脉瘤进行辅助检出的有效性，对动脉瘤检出的AI辅助阅片与单独阅片的AFROC-AUC差值大于优效界值。以专家组的动脉瘤标记结果为金标准，计算医生对动脉瘤检出的AI辅助阅片结果与单独阅片结果的AFROC-AUC并进行比较，使用优效性检验对结果进行统计分析，如果统计分析显示AFROC-AUC差值大于预先设定的优效界值，则认为该主要指标达成。

次要评价指标为：对动脉瘤检出的AI辅助阅片与单独阅片的灵敏度差值大于优效界值（病灶为单位）；对动脉瘤检出的AI辅助阅片与单独阅片的特异度差值大于非劣效界值（病例为单位）；以专家组结果为金标准，评价医生单独阅片和AI辅助阅片的总符合率、阳性预测值（PPV）、阴性预测值（NPV）；以专家组结果为金标准，评价软件单独阅片时的灵敏度、特异度（分别以病灶、病例为单位），评价软件单独阅片时的总符

合率、KAPPA 值、阳性预测值 (PPV)、阴性预测值 (NPV) (以病例为研究单位); 判定时间: 记录医生单独阅片和 AI 辅助阅片两组的阅片时间; 软件性能评价, 从功能使用、操作便捷性、稳定性方面来评价。

临床试验结果:

1. 主要有效性评价指标:

FAS: 医师使用与不使用 AI 的动脉瘤检出 AFROC-AUC 差值为 0.0409, 统计量 $F=16.9296$, $P=0.0012$, 95% 置信区间为 (0.0195, 0.0623), 区间下限大于目标值 0, 可以判定医师使用 AI 的动脉瘤检出 AFROC-AUC 优于不使用 AI 的动脉瘤检出 AFROC-AUC。

PPS: 医师使用与不使用 AI 的动脉瘤检出 AFROC-AUC 差值为 0.0409, 统计量 $F=16.9296$, $P=0.0012$, 95% 置信区间为 (0.0195, 0.0623), 区间下限大于目标值 0, 可以判定医师使用 AI 的动脉瘤检出 AFROC-AUC 优于不使用 AI 的动脉瘤检出 AFROC-AUC。

2. 次要评价指标:

(1) 动脉瘤检出的灵敏度 (以病灶为单位) FAS 和 PPS 集结果:

医师使用与不使用 AI 的病灶水平动脉瘤检出灵敏度差值

为 0.0698, 统计量 $F=9.1397, P=0.0102$, 95% 置信区间为(0.0358, 0.1037), 病灶水平下 AI 辅助阅片与单独阅片对动脉瘤检出灵敏度差值的 95% 置信区间下限大于优效界值。

(2) 动脉瘤检出的特异度 (以病例为单位) FAS 和 PPS 集结果:

医师使用与不使用 AI 的病例水平动脉瘤检出特异度差值为 0.0118, 统计量 $F=0.4568, P=0.5131$, 95% 置信区间为(-0.0164, 0.0400), 病例水平下 AI 辅助阅片与单独阅片对动脉瘤检出特异度差值的 95% 置信区间下限大于非劣效界值。

(3) 以专家组结果为金标准, 评价医生单独阅片和 AI 辅助阅片的总符合率、阳性预测值 (PPV)、阴性预测值 (NPV) 的 FAS 和 PPS 集结果:

医师使用与不使用 AI 的动脉瘤检出病例水平总符合率差值为 0.0396, 95% 置信区间为 (0.0287, 0.0505), 统计量 McNemar $\chi^2=3571.834$, $p<0.0001$, 病例水平下医师使用 AI 的动脉瘤检出的总符合率高于不使用 AI 的动脉瘤检出的总符合率。

医师使用与不使用 AI 的动脉瘤检出病例水平 PPV 差值为 0.0206, 95% 置信区间为 (0.0053, 0.0359), 统计量 McNemar $\chi^2=1775.852$, $p<0.0001$, 病例水平下医师使用 AI 的动脉瘤检出的 PPV 高于不使用 AI 的动脉瘤检出的 PPV。

医师使用与不使用 AI 的动脉瘤检出病灶水平 PPV 差值为 0.0150, 95% 置信区间为 (-0.0021,0.0321), 统计量 McNemar $\chi^2=1690.730$, $p<0.0001$, 病灶水平下医师使用 AI 的动脉瘤检出的 PPV 高于不使用 AI 的动脉瘤检出的 PPV。

医师使用与不使用 AI 的动脉瘤检出病例水平 NPV 差值为 0.0521, 95% 置信区间为 (0.0369,0.0672), 统计量 McNemar $\chi^2=1812.943$, $p<0.0001$, 病例水平下医师使用 AI 的动脉瘤检出的 NPV 高于不使用 AI 的动脉瘤检出的 NPV。

(4) 以专家组结果为金标准, 评价软件单独阅片时的灵敏度、特异度 (分别以病灶、病例为单位), 评价软件单独阅片时的总符合率、KAPPA 值、阳性预测值 (PPV)、阴性预测值 (NPV) (以病例为研究单位) 的 FAS 和 PPS 集结果:

以专家组结果为金标准, 在病例水平上, 软件单独阅片的灵敏度为 0.9144, 95% 置信区间为 (0.8696,0.9477); 特异度为 0.8550, 95% 置信区间为 (0.8123,0.8976); 总符合率为 0.8822, 95% 置信区间为 (0.8535,0.9109); Kappa 值为 0.7644, 95% 置信区间为 (0.7071,0.8217); 阳性预测值为 0.8423, 95% 置信区间为 (0.7963,0.8883); 阴性预测值为 0.9218, 95% 置信区间为 (0.8806,0.9523)。在病灶水平上, 软件单独阅片的灵敏度为 0.8140, 95% 置信区间为 (0.7665,0.8614)。

(5) 判定时间的 FAS 和 PPS 集结果:

医师使用 AI 的中位阅片时间为 2.02 (1.57,2.53) 分钟, 医师不使用 AI 的中位阅片时间为 1.47 (1.13,1.82) 分钟, 统计量符号秩=58338.5, $P<0.0001$, 医师使用 AI 的阅片时间高于不使用 AI 的阅片时间, 具有统计学意义。

(6) 软件性能评价

软件的界面友好性、操作便捷性、使用稳定性、图像存储方便安全性评分均 ≥ 4 分。

综上, 申请人提供的临床评价资料符合目前的临床审评要求。

四、产品受益风险判定

该产品主要受益为: 该产品用于对颅脑 TOF-MRA 血管造影图像的显示、处理、测量和分析, 可对 3mm 及以上颅内囊性动脉瘤进行检出, 供经培训合格的医师使用, 不能单独用作临床诊疗决策依据。

该产品主要风险为: (1) 假阳性结果导致不必要后续过度治疗风险。(2) 假阴性结果导致患者颅内动脉瘤病灶出现漏诊。以上风险通过设计、防护等措施进行控制, 相关警示、注意事项均已在说明书中进行提示。

综上，该产品上市带来的受益大于风险，综合剩余风险可接受。

综合评价意见

申请人申请境内第三类医疗器械注册，注册申报资料齐全，符合要求。

依据《医疗器械监督管理条例》（国务院令 739 号）、《医疗器械注册与备案管理办法》（国家市场监督管理总局令第 47 号）等相关医疗器械法规及配套规章，技术审评经系统评价产品安全性和有效性后，基于当前认知水平，认为该产品临床使用受益大于风险，注册申报资料符合现行技术审评要求，建议予以注册。

2024 年 8 月 13 日