

附件

人工智能医疗器械注册审查指导原则

本指导原则旨在指导注册申请人建立人工智能医疗器械生存周期过程和准备人工智能医疗器械注册申报资料，同时规范人工智能医疗器械的技术审评要求，为人工智能医疗器械、质量管理软件的体系核查提供参考。

本指导原则是对人工智能医疗器械的一般要求。注册申请人需根据产品特性和风险程度确定本指导原则具体内容的适用性，若不适用详述理由。注册申请人也可采用其他满足法规要求的替代方法，但需提供详尽的支持资料。

本指导原则是在现行法规、强制性标准体系以及当前科技能力、认知水平下制定的，随着法规、强制性标准体系的不断完善以及科技能力、认知水平的不断发展，本指导原则相关内容也将适时调整。

本指导原则是供注册申请人、审评人员和检查人员使用的指导文件，不涉及行政审批事项，亦不作为法规强制执行，应在遵循相关法规的前提下使用本指导原则。

本指导原则作为数字医疗（Digital Health）指导原则体系的重要组成部分，采用和遵循医疗器械软件、医疗器械网络安全、移动医疗器械、医疗器械人因设计、医疗器械独立软件生产质量现场检查等相关指导原则的概念和要求。

本指导原则是人工智能医疗器械的通用指导原则¹，其他含有或涉及人工智能技术的医疗器械指导原则可在本指导原则基础上结合具体情况进行有针对性的调整、修改和完善。

一、适用范围

本指导原则适用于人工智能医疗器械的注册申报，包括第二类、第三类人工智能独立软件和含有人工智能软件组件的医疗器械（包括体外诊断医疗器械）；适用于自研软件的注册申报，现成软件组件参照执行，不适用于外部软件环境。

本指导原则也可用作人工智能医疗器械的体系核查参考。质量管理软件若采用人工智能技术实现其功能或用途，亦可参考本指导原则的适用要求。

二、主要概念

（一）人工智能医疗器械

本指导原则所述人工智能医疗器械是指基于“医疗器械数据”，采用人工智能技术实现其预期用途（即医疗用途）的医疗器械。

医疗器械数据是指医疗器械产生的用于医疗用途的客观数据，如医学影像设备产生的医学图像数据（如X射线、CT、MRI、超声、内窥镜、光学等图像）、医用电子设备产生的生理参数数据（如心电、脑电、血压、无创血糖、心音等波形数据）、体外诊断设备产生的体外诊断数据（如病理图像、显微图像、有创血糖波形数据等）；在特殊情形下，通用设备（非监管对象）

¹取代《深度学习辅助决策医疗器械软件审评要点》非临床部分的要求。

产生的用于医疗用途的客观数据亦属于医疗器械数据，如数码相机拍摄的用于皮肤疾病诊断的皮肤照片、健康电子产品采集的用于心脏疾病预警的心电数据等。基于医疗器械数据包括医疗器械数据的生成、使用等情况，其中使用情况含单独使用医疗器械数据，或者以医疗器械数据为主联合使用非医疗器械数据（如患者主诉信息、检验检查报告结论、电子病历、医学文献等）。

人工智能是指机器表现出与人类智能相关行为的能力，通常是指通过感知周围环境做出合理行动以达到预期目标的计算机软件或系统。机器学习是指与人类学习行为相关的人工智能，通常是指通过整理现有数据和/或获取新数据以提升性能的计算机软件或系统。机器学习虽是人工智能的子集，但却为人工智能的核心领域，当前二者对于医疗器械而言含义基本相同，故本指导原则从医疗器械安全有效性评价角度出发对二者不做严格区分，统一采用人工智能进行表述。

基于非医疗器械数据的医学人工智能产品，或者采用人工智能技术实现非医疗用途和非医疗器械功能（详见医疗器械软件指导原则）的医疗器械均非人工智能医疗器械。医学人工智能产品是否按医疗器械管理，根据相应分类界定指导原则进行判定，必要时申请医疗器械分类界定。

（二）人工智能医疗器械类型

从医疗器械软件角度，人工智能医疗器械可分为人工智能独立软件和人工智能软件组件，故其类型划分可参考医疗器械

软件指导原则相关维度。

人工智能医疗器械从用途角度可分为辅助决策类和非辅助决策类。其中，辅助决策是指通过提供诊疗活动建议辅助用户（如医务人员、患者）进行医疗决策，如通过病灶特征识别、病灶性质判定、用药指导、治疗计划制定进行辅助分诊、辅助检测、辅助诊断、辅助治疗等，相当于用户的“助手”。反之，仅提供医疗参考信息而不进行医疗决策即为非辅助决策，包括流程优化、诊疗驱动，前者如成像流程简化、诊疗流程简化等，后者如成像质量改善、成像速度提高、自动测量、自动分割、三维重建等，相当于用户的“工具”。此外，辅助决策和非辅助决策从实时性角度均可细分为实时和非实时，前者风险通常高于后者。

人工智能医疗器械从功能角度大体上可分为处理功能、控制功能、安全功能。其中，处理功能又可分为前处理功能和后处理功能，前处理功能是指采集人体解剖、生理信息生成医疗器械数据过程的处理功能，如成像流程简化、成像质量改善、成像速度提高等；后处理功能是指利用医疗器械数据生成诊疗信息或进行医疗干预过程的处理功能，如诊疗流程简化、自动测量、自动分割、三维重建、病灶特征识别、病灶性质判定、用药指导、治疗计划制定等。控制功能是指控制/驱动医疗器械硬件运行的功能，如闭环控制、机械臂运动控制等。安全功能是指保证医疗器械安全性的功能，如风险预警、急停控制等。

人工智能医疗器械从算法角度具有多种类型划分维度。从

学习策略角度可分为有监督学习和无监督学习，前者需要对训练数据进行标注，如线性回归、逻辑回归、决策树、朴素贝叶斯、K近邻、支持向量机等经典回归、分类算法，后者无需对训练数据进行标注，如K均值、主成分分析等经典聚类、降维算法，前者对于数据标注的要求高于后者。从学习方法角度可分为基于模型的算法和基于数据的算法，前者采用统计模型、规则推理等方法，后者主要采用大数据方法，前者对于训练数据量的要求低于后者。从可解释性角度可分为白盒算法和黑盒算法，前者特征提取需要人为干预，可与现有医学知识建立关联，后者自动完成特征提取，难与现有医学知识建立关联，前者可解释性优于后者。

上述类型划分维度相互交叉，例如：前处理和后处理均可采用不同类型的人工智能算法实现辅助决策、非辅助决策用途，有监督学习和无监督学习既可采用基于模型的算法、基于数据的算法，又可采用黑盒算法、白盒算法。同时，同一维度亦不存在严格的划分界线，例如：在用途方面，通过图像识别技术进行流程优化则需考虑诊疗驱动相关要求，自动测量结果若为医疗决策重要指标（如血流储备分数 FFR）则属于辅助决策范畴；在功能方面，控制功能、安全功能可与处理功能相结合，前处理过程可包含后处理功能；在算法方面，某些算法既可用于有监督学习又可用于无监督学习，有监督学习和无监督学习可结合为半监督学习；基于模型的算法亦需数据的支持，基于数据的算法亦可生成模型；白盒算法和黑盒算法可组合使用成

为灰盒算法。

同样，人工智能医疗器械从成熟度角度可分为成熟和全新两种类型，其中成熟是指安全有效性已在医疗实践中得到充分证实的情形，全新是指未上市或安全有效性尚未在医疗实践中得到充分证实的情形。人工智能医疗器械的算法、功能、用途若有一项为全新则属于全新类型，反之属于成熟类型。

人工智能医疗器械可同时采用多种、多个人工智能算法，在前处理、后处理过程中实现辅助决策、非辅助决策用途。因此，注册申请人需结合人工智能医疗器械的预期用途、使用场景、核心功能以及所用算法的类型特点、技术特征、组合形式开展相应产品质控工作，以保证产品的安全有效性。

（三）人工智能算法更新

人工智能算法特别是基于数据的算法，具有快速迭代更新的特性。人工智能算法更新属于软件更新范畴，故遵循软件更新的基本原则及要求：人工智能算法更新若影响到人工智能医疗器械的安全性或有效性则属于重大软件更新，应申请变更注册；反之，人工智能算法更新若未影响到人工智能医疗器械的安全性和有效性则属于轻微软件更新，通过质量管理体系进行控制，无需申请变更注册，待下次变更注册时提交相应注册申报资料。

人工智能算法更新可分为算法驱动型更新和数据驱动型更新。其中，算法驱动型更新是指人工智能医疗器械所用算法、算法结构、算法流程、算法编程框架（详见后文）、输入输出

数据类型等发生改变，通常属于重大软件更新。此外，算法重新训练即弃用原有训练数据而采用全新训练数据进行算法训练，亦属于算法驱动型更新。

数据驱动型更新是指仅由训练数据量增加而发生的算法更新。数据驱动型更新是否属于重大软件更新原则上以算法性能评估结果（基于相同的测试集和算法性能评估指标）为准，算法性能评估结果若发生显著性改变则属于重大软件更新，即算法性能评估结果与前次注册（而非前次更新）相比存在统计学差异，反之属于轻微软件更新。

人工智能医疗器械其他类型的算法更新、软件更新以及重大软件更新判定原则详见医疗器械软件指导原则、医疗器械网络安全指导原则。

软件版本命名规则原则上应涵盖算法驱动型更新和数据驱动型更新，明确并区分重大软件更新和轻微软件更新，其中重大软件更新列举常见典型情况。软件版本命名规则的基本要求详见医疗器械软件指导原则、医疗器械网络安全指导原则。

三、基本原则

（一）基于算法特性

人工智能技术从发展驱动要素角度是基于模型/数据和算力的算法，其中模型/数据是人工智能技术的基础，算力是人工智能技术的保证，算法是人工智能技术的核心。

由于算力所用计算资源本身不属于监管对象，计算资源的监管要求取决于其所属的计算平台类型。故从监管角度出发，

人工智能医疗器械安全有效性评价基于其预期用途、使用场景、核心功能，以算法特性为核心重点关注其泛化能力，以模型/数据为基础重点关注其质控情况，同时从风险管理角度兼顾算力不足与失效的影响。

人工智能算法的类型不同，其算法特性、适用场景也不同，评价重点亦有所侧重；同时，不同类型的人工智能算法可组合使用，需结合各算法特性和算法组合形式进行整体评价。因此，注册申请人需结合人工智能医疗器械的预期用途、使用场景、核心功能选择与之相适宜的人工智能算法或算法组合，基于算法特性并结合风险管理开展相应验证与确认工作。

以深度学习为例，其是指通过训练具有多个隐层的神经网络而获得输入输出映射关系的人工智能算法，亦是基于海量数据和高算力的黑盒算法，既可用于有监督学习又可用于无监督学习。因此，对于采用深度学习技术的人工智能医疗器械，基于其预期用途、使用场景、核心功能，重点关注其算法泛化能力、数据质控、可解释性等问题，同时，深度学习若与其他类型的人工智能算法组合使用，还需基于各算法特性重点关注算法组合的整体评价问题。

人工智能新算法研究处于深入发展阶段，如基于小样本数据、基于弱标注数据、基于非结构化数据、黑盒算法透明化等算法。人工智能医疗器械若使用人工智能新算法，亦需基于算法特性并结合风险管理开展相应验证与确认工作，以保证产品的安全有效性。

（二）风险导向

人工智能医疗器械的风险水平亦可用软件安全性级别进行表述，软件安全性级别越高，其生存周期质控要求越严格，注册申报资料越详尽，同时由于全新类型的潜在未知风险多于成熟类型，故需结合成熟度予以综合考虑，具体要求详见医疗器械软件指导原则。

人工智能医疗器械的软件安全性级别可基于产品的预期用途、使用场景、核心功能进行综合判定，其中预期用途主要考虑用途类型、重要程度、紧迫程度等因素，使用场景主要考虑使用场合、疾病特征、适用人群、目标用户等因素，核心功能主要考虑功能类型、核心算法、输入输出、接口等因素。亦可根据风险管理所确定的风险等级进行判定，软件安全性级别与风险等级的分级可以不同，但二者存在对应关系，因此可根据风险等级来判定软件安全性级别，但应在采取风险控制措施之前进行判定。

人工智能医疗器械的主要风险从算法角度包括过拟合和欠拟合，其中过拟合是指算法对于训练数据过度学习而将非普遍规律作为重要特征，欠拟合是算法对于训练数据学习不充分而遗漏重要特征，均会降低算法泛化能力。从用途角度，辅助决策主要包括假阴性和假阳性，其中假阴性即漏诊，可能导致后续诊疗活动延误，特别是要考虑快速进展疾病的诊疗活动延误风险，而假阳性即误诊，可能导致后续不必要的诊疗活动；非辅助决策从算法设计目标能否得以实现角度，亦可参考辅助决

策分为假阴性和假阳性。此外，进口人工智能医疗器械还需考虑中外差异风险，如人种、流行病学特征、临床诊疗规范等差异。

注册申请人应结合人工智能医疗器械的预期用途、使用场景、核心功能开展风险管理活动，采取风险控制措施将风险降至可接受水平，并贯穿于人工智能医疗器械全生命周期过程。

（三）全生命周期质控

注册申请人应结合质量管理体系要求，参考软件、人工智能相关标准和良好工程实践，建立人工智能医疗器械生存周期过程，开展与软件安全性级别相匹配的产品质量保证工作，将风险管理、可追溯分析贯穿于生存周期全程，形成记录以供体系核查。

上市前开展充分、适宜、有效的验证与确认活动，保证算法泛化能力满足用户需求，识别可预见风险并将其降至可接受水平，明确产品使用限制（含技术限制，下同）和必要警示提示信息。上市后持续开展算法泛化能力研究，同时结合用户投诉、不良事件和召回等情况识别前期未预见的风险，并采取有效的风险控制措施将风险降至可接受水平。此外，根据产品更新需求，经评估后实施更新活动，开展与之相适宜的验证与确认活动，保证算法泛化能力持续满足用户需求。

四、人工智能医疗器械生存周期过程

人工智能医疗器械生存周期（又称生命周期）过程可基于软件生存周期过程予以建立，具体要求详见医疗器械独立软件

生产质量管理规范及其现场检查指导原则。

考虑到有监督深度学习是当前人工智能医疗器械的主流算法，故本指导原则以有监督深度学习为例详述人工智能医疗器械生存周期过程质控要求，主要包括需求分析、数据收集、算法设计、验证与确认、更新控制等阶段。其他类型的人工智能算法可参照执行，不适用内容详述理由并予以记录。

（一）需求分析

需求分析以用户需求与风险为导向，结合产品的预期用途、使用场景、核心功能，综合考虑法规、标准、用户、产品、数据、功能、性能、接口、用户界面、网络安全、警示提示等需求，重点考虑数据收集、算法性能、使用限制等要求。

为保证数据质量和控制数据偏倚，数据收集需考虑数据来源的合规性、充分性和多样性，数据分布的科学性和合理性，数据质控的充分性、有效性和准确性。数据来源在合规性（如个人信息保护法）基础上保证充分性和多样性，以提高算法泛化能力，例如：根据产品的预期用途和使用场景尽可能多采集数据，如来源于多家、多地域、多层级的代表性临床机构，以及多家、多种、多参数的代表性采集设备。数据分布的科学性和合理性需结合目标疾病流行病学特征予以考虑，包括但不限于疾病构成（如分型、分级、分期）、人群分布（如健康、患者，性别、年龄、职业、地域、生活方式）、统计指标（如发病率、患病率、治愈率、死亡率、生存率）等情况，以及目标疾病并发症与类似疾病的影响情况。数据质控的充分性、有效性和准

确性需结合数据采集与数据标注的人员、设备、过程等影响因素予以考虑，具体要求详见后文。

算法性能需结合医疗实际和产品定位，综合考虑假阴性与假阳性、重复性与再现性、鲁棒性/健壮性、实时性等性能指标的适用性及其要求，兼顾不同性能指标的制约关系，如假阴性与假阳性等。同时，结合当前医疗水平情况，考虑金标准或参考标准的确定依据、实现方法和质控要求，以保证算法性能评估的准确性，必要时纳入数据收集过程。

使用限制需考虑产品禁用、慎用等场景，准确表述产品使用场景，提供必要警示提示信息。

（二）数据收集

数据收集基于合规性要求，主要考虑数据采集、数据整理、数据标注、数据集构建等活动的质控要求，以保证数据质量和算法训练效果。

1. 数据采集

数据采集需考虑采集设备、采集过程、数据脱敏等质控要求，并建立数据采集操作规范。数据采集亦可使用历史数据，需结合样本规模、采集难度等影响因素合理选择数据采集方式。若适用，数据采集需经伦理委员会批准。

采集设备质控包括采集设备的兼容性和采集特征等要求。兼容性基于数据生成方式（直接生成、间接生成）考虑采集设备的兼容性要求，如采集设备的名称、型号规格、制造商、性能指标等要求，若无需考虑兼容性要求详述理由并予以记录。

采集特征考虑采集设备的采集方式（如常规成像、增强成像）、采集协议（如 MRI 成像序列）、采集参数（如 CT 加载电压、加载电流、加载时间、层厚）、采集精度（如分辨率、采样率）等要求。

采集过程质控包括人员管理、采集流程、采集质量评估等要求。人员管理考虑采集人员、审核人员的选拔（如职称、工作年限、工作经验、所在机构，若有国外人员则需明确其资质要求）、培训（如培训材料、培训方案）、考核（如方法、频次、指标、通过准则、一致性）等要求。采集流程考虑人员职责、采集步骤、结果审核等要求。采集质量评估考虑评估人员、评估方法、评估指标、通过准则等要求，并记录评估结果。

数据采集若使用历史数据，需列明采集设备及采集特征要求，并开展数据采集质量评估工作。

采集的数据应进行数据脱敏以保护患者隐私，数据脱敏需明确脱敏的类型（静态、动态）、规则、方法以及脱敏内容的确定依据。

脱敏数据汇总形成原始数据库，不同模态的数据在原始数据库中需加以区分（下同）。原始数据库需考虑样本规模的充分性、样本多样性等问题。

2. 数据整理

数据整理基于原始数据库考虑数据清洗、数据预处理的质控要求。数据清洗需明确清洗的规则、方法、结果，数据预处理需明确处理的方法（如滤波、增强、重采样、尺寸裁剪、均

一化等)、结果。数据整理所用软件工具(含脚本,下同)均需明确名称、型号规格、完整版本、制造商、运行环境,并进行软件确认。

数据经整理后形成基础数据库,需明确样本类型、样本量、样本分布等信息。样本类型以适用人群为单位可分为单一数据、数据序列(由多个单一数据组成,如结构序列、功能序列、时间序列)。样本量需考虑样本规模的充分性,明确样本总量及其确定依据。样本分布需考虑样本的科学性和合理性,依据适用人群、数据来源机构、采集设备、样本类型等因素明确疾病构成的数据分布情况。

3.数据标注

数据标注作为有监督学习数据质控的关键环节,需建立数据标注操作规范,明确标注资源管理、标注过程质控、标注质量评估等要求。

标注资源管理包括人员管理和基础设施管理。人员管理考虑标注人员、审核人员和仲裁人员的选拔(如职称、工作年限、工作经验、所在机构,若有国外人员则需明确其资质要求)、培训(如培训材料、培训方案)、考核(如方法、频次、指标、通过准则、一致性)等要求。基础设施管理考虑标注场所(真实场所或模拟场所,模拟场所可根据产品实际情况调整模拟程度,详述调整理由并予以记录)、标注环境条件(如空间、照明、温度、湿度、气压)、标注软件(名称、型号规格、完整版本、制造商、运行环境、软件确认)等要求。

标注过程质控包括人员职责（如人员资质、人员数量、职责分工）、标注规则（如临床指南、专家共识、专家评议、文献分析）、标注流程（如标注对象、标注形式、标注轮次、标注步骤、结果审核）、分歧处理（如仲裁人员、仲裁方式）、可追溯性（如数据、操作）等要求。

标注质量评估包括评估人员、评估方法、评估指标、通过准则等要求，并记录评估结果。

数据经标注后形成标注数据库，样本类型可分为数据块（如图像区域、数据片段）、单一数据（由多个数据块组成）、数据序列（由多个单一数据组成）。标注数据库的样本量、样本分布等要求及风险考量与基础数据库相同。

数据标注可使用自动标注软件，但自动标注结果不得直接使用，应由标注人员审核后方可使用；同时，自动标注软件亦需明确名称、型号规格、完整版本、制造商、运行环境等信息，并进行软件确认。

4.数据集构建

基于标注数据库构建训练集（用于算法训练）、调优集²（若有，用于算法超参数调优）、测试集（用于算法性能评估），明确训练集、调优集、测试集的划分方法、划分依据、数据分配比例。训练集原则上需保证样本分布具有均衡性，测试集、调优集原则上需保证样本分布符合真实情况，训练集、调优集、测试集的样本应两两无交集并通过查重予以验证。

²人工智能领域称之为验证集(Validation set)。为避免与医疗器械领域所用术语验证(Verification)、确认(Validation)相混淆，本指导原则将其改称为调优集。

为解决样本分布不满足预期的问题，可对训练集、调优集小样本量数据进行扩增，原则上不得对测试集进行数据扩增，对抗测试（详见后文）除外。数据扩增需明确扩增的对象、范围、方式（离线、在线）、方法（如翻转、旋转、镜像、平移、缩放、滤波、生成对抗网络等）、倍数，在线扩增亦需予以记录，扩增需考虑数据偏倚的影响及风险，如部分数据扩增倍数过大、数据扩增倍数不均衡等。若采用生成对抗网络（详见后文）进行数据扩增，需明确算法基本信息以及算法选用依据。

数据经扩增后形成扩增数据库，需列表对比扩增数据库与标注数据库在样本量、样本分布（注明扩增倍数）等差异，以证实扩增数据库样本量的充分性以及样本分布的合理性。

（三）算法设计

人工智能算法作为人工智能医疗器械的核心，其设计主要考虑算法选择、算法训练、算法性能评估等要求。对于黑盒算法，算法设计应开展算法性能影响因素分析，同时建议与现有医学知识³建立关联，以提升算法可解释性。

1. 算法选择

算法选择提供所用算法的名称、类型（如有监督学习、无监督学习，基于模型、基于数据，白盒、黑盒）、结构（如层数、参数规模）、输入输出数据类型、流程图、算法编程框架、运行环境等基本信息，并明确算法选用依据，包括选用的理由和基本原则。

³医学知识作为医疗器械的外部参考准则，其自身评价问题不属于医疗器械安全有效性评价范畴。

若组合使用集成学习、迁移学习、强化学习等，亦需提供算法基本信息以及算法选用依据，详见后文。

2. 算法训练

算法训练需基于训练集、调优集进行训练和调优，考虑评估指标、训练方式、训练目标、调优方式、训练数据量-评估指标曲线等要求。

评估指标建议根据用户需求进行选择，辅助决策可选择敏感性、特异性等指标，非辅助决策可选择图像质量、测量准确性等指标。训练方式包括但不限于留出法和交叉验证法，若组合使用联邦学习（详见后文）亦需明确算法选用依据，并提供算法基本信息。训练目标根据医疗情况详述目标确定依据，提供受试者工作特征（**ROC**）曲线或其衍生曲线、混淆矩阵及其衍生参数等证据予以证实。调优方式明确优化策略和实现方法。训练数据量-评估指标曲线用于证实算法训练的充分性和有效性，若无法提供则需详述理由并提供替代证据。

3. 算法性能评估

算法性能评估作为软件验证的重要组成部分，需基于测试集对算法设计结果进行评估，综合考虑假阴性与假阳性、重复性与再现性、鲁棒性/健壮性、实时性等适用评估要求，以证实算法性能满足算法设计目标，并作为软件验证、软件确认的基础。亦可基于第三方数据库（详见后文）开展算法性能评估。

同时，开展算法性能影响因素分析以提升算法可解释性，详述影响算法性能的主要因素及其影响程度，如采集设备、采

集参数、疾病构成、病变特征等因素，基于分析结果明确产品使用限制和必要警示提示信息。

此外，根据产品实际情况开展压力测试（详见后文）、对抗测试等测试，以全面深入评估算法性能。

（四）验证与确认

1. 软件验证

软件验证是指通过提供客观证据认定软件开发、软件更新某一阶段的输出满足输入要求，包括软件验证测试（单元测试、集成测试、系统测试）、设计评审等系列活动。

软件验证基于软件需求予以开展，保证软件的安全有效性，并作为软件确认的基础。

2. 软件确认

软件确认是指通过提供客观证据认定软件满足用户需求和预期目的，包括软件确认测试（用户测试）、临床评价、设计评审等系列活动。根据产品实际情况，软件确认方式可单一使用，亦可组合使用。

软件确认测试基于用户需求，由预期用户在真实或模拟使用场景下予以开展，亦可基于测评数据库（详见后文）予以开展。

临床评价基本原则详见医疗器械软件指导原则，基于核心功能或核心算法，结合预期用途和成熟度予以综合考虑：非辅助决策类功能基于核心功能开展同品种医疗器械比对，全新的功能、算法和用途原则上均需开展临床评价；辅助决策类功能

基于核心算法开展同品种医疗器械比对，所选同品种医疗器械的临床证据原则上需基于临床试验（含回顾性研究，下同），全新的功能、算法和用途原则上均需开展临床试验。具体要求详见医疗器械临床评价等通用指导原则，以及人工智能医疗器械临床评价等专用指导原则。

同时，开展算法性能比较分析，若各类测试场景（含临床评价）算法性能变异度较大，详述原因并基于分析结果明确产品使用限制和必要警示提示信息。

最后，结合算法训练、算法性能评估、临床评价等结果开展算法性能综合评价，针对训练样本量和测试样本量过少、测试结果明显低于算法设计目标、算法性能变异度过大等情况，对产品的适用范围、使用场景、核心功能进行必要限制。

（五）更新控制

人工智能医疗器械若发生算法更新、软件更新，均应当按照质量管理体系的要求，开展与算法更新、软件更新的类型、内容和程度相适宜的验证与确认活动，将风险管理、可追溯分析贯穿于更新全程，形成记录以供体系核查。

对于算法更新，无论算法驱动型更新还是数据驱动型更新，均应开展相应验证与确认活动，如算法性能评估、临床评价等，以保证算法更新的安全有效性。对于软件更新，具体要求详见医疗器械软件指导原则、医疗器械独立软件生产质量现场检查指导原则。

此外，算法更新、软件更新均需考虑引入回滚机制，以保

证医疗业务的连续性，特别是对风险较高的软件。

总之，人工智能医疗器械所含的每个人工智能算法或算法组合，均需独立开展需求分析、数据收集、算法设计、验证与确认、更新控制等活动，同时考虑人工智能算法组合的整体评价要求，以保证产品的安全有效性。

前期已开发的人工智能医疗器械若不满足本指导原则的适用要求，应开展差距分析并予以记录，明确产品使用限制和必要警示提示信息。必要时采取补救措施，以满足质量管理体系要求，并保证产品的安全有效性。

五、技术考量

（一）注册单元与检测单元

人工智能医疗器械的注册单元与检测单元划分原则详见医疗器械软件指导原则，人工智能独立软件、人工智能软件组件分别参照独立软件、软件组件的要求。

若软件核心功能（即软件在预期使用场景完成预期用途所必需的功能）相同，但核心算法（即实现软件核心功能所必需的算法）类型不同，则每类核心算法所对应的核心功能均需检测，即检测对象为核心功能而非核心算法。

（二）网络安全与数据安全

人工智能医疗器械需结合预期用途、使用场景、核心功能，基于保密性、完整性、可得性等网络安全特性，确定网络安全能力建设要求，以应对网络攻击和数据窃取等网络威胁，如算法编程框架漏洞攻击、数据污染等。具体要求详见医疗器械网

络安全指导原则。

除此之外，还需在人工智能医疗器械全生命周期过程中考虑数据安全问题，包括上市前设计开发阶段和上市后使用阶段。数据转移需明确转移方法、数据污染防治措施、数据销毁等要求。数据整理、数据集构建、算法训练、算法性能评估、软件验证等内部活动需在封闭或受控的网络环境下开展以防止数据污染。数据标注、软件确认等涉及外方的活动需明确数据污染防治措施，特别是在开放网络环境下。各数据库（集）均需进行数据备份，明确备份的方法、频次、数据恢复方法。数据采集、上市后使用需考虑医疗机构关于网络安全与数据安全的接口要求。

（三）移动计算与云计算

人工智能医疗器械若使用移动计算、云计算等技术，则遵循相关指导原则要求。具体而言，移动计算相关要求详见移动医疗器械指导原则，云计算相关要求详见医疗器械软件指导原则，移动计算与云计算的网络安全相关要求详见医疗器械网络安全指导原则。

（四）人因与可用性

建议加强人工智能医疗器械的人因设计以提升可用性，将用户错误使用的风险降至可接受水平，特别是软件用户界面。具体要求详见医疗器械人因设计相关指导原则。

（五）压力测试

本指导原则所述压力测试⁴是指采用罕见或特殊的真实数据样本开展的算法性能测试，侧重于评估算法泛化能力的极限。

注册申请人需根据产品实际情况开展压力测试，以全面深入评估算法性能，必要时可引入对抗样本开展对抗压力测试。若未开展相应测试或测试结果不佳，均需对产品的适用范围、使用场景、核心功能进行必要限制，并在说明书中明确产品使用限制和必要警示提示信息。

（六）对抗测试

本指导原则所述对抗测试是指基于真实数据样本利用数据扰动、生成对抗网络等技术产生对抗样本，并采用对抗样本开展的算法性能测试，侧重于评估算法的鲁棒性/健壮性。

若条件具备，建议注册申请人开展对抗测试，以全面深入评估算法性能。若未开展相应测试或测试结果不佳，均需明确产品使用限制和必要警示提示信息。

（七）第三方数据库

第三方数据库可用于算法性能评估，但其类型、用途等情况各不相同，未必能够完全满足软件确认测试的要求。因此，使用第三方数据库进行软件确认测试，需评估其满足软件确认测试的充分性、适宜性和有效性。

可用于软件确认测试的第三方数据库即为测评数据库。测评数据库除满足数据库通用要求（如数据管理、网络安全与数据安全、可扩展性）外，还应满足以下专用要求。

⁴与软件工程领域的压力测试含义不同。

1.权威性：测评数据库的数据采集、数据标注及其质控工作由相应临床专业领域的权威机构（如国家临床医学研究中心等）负责，以保证数据准确性；标注人员、审核人员、仲裁人员需具备与其岗位职责相匹配的临床实践经验，以保证数据质量。

2.科学性：测评数据库的数据样本均为临床真实数据，不得进行数据扩增；样本分布符合目标疾病流行病学特征情况，样本总量基于目标疾病流行病学统计指标、测试抽样误差控制要求，通过统计学计算予以确定；单次测试原则上根据测评数据库样本分布情况等比例随机抽取测试样本，且其样本量满足测试抽样误差控制要求。

3.规范性：测评数据库的数据采集、数据脱敏、数据清洗、数据预处理、数据标注、数据更换、数据管理、数据安全保证、数据备份等数据治理活动以及测评活动均需建立质控程序文件，并满足可追溯性要求。

4.多样性：测评数据库的样本需来源于多家、多地域、多层级的代表性临床机构以及多家、多种、多参数的代表性采集设备，以保证数据多样性能够满足算法泛化能力评估要求；若条件具备，测评数据库可包含适当比例的罕见或特殊的临床真实数据样本，以便开展压力测试深入评估算法泛化能力的极限。

5.封闭性：测评数据库需封闭管理，样本总量需远大于单次测试样本量，测评活动亦需封闭管理，以保证算法性能评价的客观性、公正性、公平性。

6.动态性：测评数据库需定期补充或更换一定比例的数据样

本，以保证其具备持续的科学性、多样性和封闭性以及数据的时效性；数据补充或更换的频率、比例需有确定依据，并满足规范性要求。更换出库的数据样本由测评数据库责任方自行确定其处理方案。

基于测评数据库，利用数据扰动、生成对抗网络等技术，可衍生出对抗测评数据库。若条件具备，建议基于测评数据库建设对抗测评数据库，以全面深入评价算法性能。

注册申请人可依据上述专用要求对第三方数据库进行筛选，若第三方数据库能用作测评数据库（含对抗测评数据库，下同）则可用于软件确认测试，并在产品注册申报时按医疗器械主文档登记事项要求提交测评数据库评估材料。

公开数据库因不具备封闭性而不能用作测评数据库，但可用于算法性能评估，若用于算法训练需重新进行质量评估。使用公开数据库需根据其使用情况开展必要评估工作，并在注册申报资料中予以说明。

其他类型第三方数据库可参照测评数据库和公开数据库的要求予以使用，同时考虑其适用场景、数据质量评估、时效性等要求。

（八）白盒算法

白盒算法的特征提取需要人为干预，可与现有医学知识建立关联，可解释性高，通常为基于模型的算法。

对于此类算法，有监督学习的模型/数据质控要求参照有监督深度学习执行，不适用内容详述理由并予以记录。同理，无

监督学习的模型/数据质控要求亦参照有监督深度学习执行，其中数据标注不适用，其余不适用内容详述理由并予以记录。

此类算法无论有监督学习还是无监督学习，均需明确特征信息，如特征分类（如人口统计学、生物学、形态学）、特征属性（如形态、纹理、性质、尺寸、边界）、特征展现方式（如形状、尺寸、边界、颜色、数量）等信息。

（九）集成学习

集成学习是指产生多个基学习器并以某种结合策略将基学习器进行集成的人工智能算法。集成学习从不同角度具有不同类型划分维度，如同质和异质的基学习器集成算法、串行和并行的集成算法等，并可与深度学习等人工智能算法组合使用。

人工智能医疗器械若使用集成学习算法，注册申报材料需明确算法的名称、类型、输入输出、流程图、运行环境等基本信息以及算法选用依据，并根据集成学习类型及其算法特性提供算法验证与确认资料。

（十）迁移学习

迁移学习是指将在某领域或任务学习到的模型应用于不同但相关的领域或任务的人工智能算法，如将在自然图像领域学习形成的模型应用于医学图像领域。迁移学习主要包括基于特征和基于模型的算法，并可与深度学习、联邦学习等人工智能算法组合使用。

人工智能医疗器械若使用迁移学习算法，注册申报材料需明确算法的名称、类型、输入输出、流程图、运行环境等基本

信息以及算法选用依据，并根据迁移学习的类型及其算法特性提供预训练模型的数据集构建、算法测试等资料。

（十一）强化学习

强化学习是指基于行动与环境的交互，实现行动从环境所获累积奖励最大化的人工智能算法，采用“试错”的学习策略。强化学习从不同角度具有不同类型划分维度，如有模型与无模型的算法、基于策略和基于价值的算法等，并可与深度学习等人工智能算法组合使用。

人工智能医疗器械若使用强化学习算法，注册申报材料需明确算法的名称、类型、输入输出、流程图、运行环境等基本信息以及算法选用依据，并根据强化学习类型及其算法特性提供算法验证与确认资料。

（十二）联邦学习

联邦学习是基于多个数据源协同建立学习模型的人工智能算法，即每个数据源独立进行本地数据学习，通过交换数据模型共同建立学习模型，实为分布式算法训练技术，亦属于隐私计算技术。联邦学习主要包括横向联邦（基于样本）和纵向联邦（基于特征）等算法，并可与迁移学习等人工智能算法组合使用。

人工智能医疗器械若使用联邦学习算法等隐私计算技术进行算法训练，注册申报材料需明确算法的名称、类型、输入输出、流程图、运行环境等基本信息以及算法选用依据，并根据联邦学习等隐私计算技术的算法类型及其算法特性提供算法验

证与确认资料。

（十三）生成对抗网络

生成对抗网络是通过生成器和判别器的互相博弈学习而产生与真实数据类似样本的人工智能算法，主要用于数据扩增、对抗测试的样本生成（详见前文），并可与深度学习等人工智能算法组合使用。

人工智能医疗器械若使用生成对抗网络进行数据扩增、对抗样本生成，注册申报资料需明确算法的名称、类型、输入输出、流程图、运行环境等基本信息以及算法选用依据，并根据生成对抗网络类型及其算法特性提供算法测试资料。此时，建议开展对抗测试。

（十四）持续学习/自适应学习

持续学习/自适应学习具备自学习能力，部署后可通过持续学习用户数据而进行产品快速更新。此时，用户亦成为产品开发者，与注册申请人共同承担产品质量责任和法律责任；同时，此种更新对于产品安全有效性的影响具有高度不确定性，特别是基于数据的无监督学习。

因此，在当前法律法规体系和技术水平条件下，持续学习/自适应学习应关闭自学习功能，或者虽开放自学习功能但不得投入使用，即用户始终使用产品原有功能，自学习功能仅用于算法训练或医学科研。

注册申请人应按照质量管理体系要求，对自学习功能所产生的产品更新的安全有效性进行验证与确认，必要时申请变更

注册，待批准后方能将自学习功能所产生的产品更新投入使用。

（十五）人工智能算法编程框架

人工智能算法编程框架（以下简称算法框架）从开发者角度可分为自研算法框架和现成算法框架，其中自研算法框架即注册申请人自行研发的人工智能算法框架，属于自研软件；现成算法框架是指注册申请人所使用的由第三方开发的人工智能算法框架，视为现成软件，现成算法框架开发者视为医疗器械供应商。

算法框架更新包括产品更新（即更换算法框架）、版本更新、补丁更新，其中产品更新、非效率型版本更新对于人工智能医疗器械而言属于重大软件更新，效率型版本更新（即运算效率单纯提高）、补丁更新对于人工智能医疗器械而言通常属于轻微软件更新，除非影响到人工智能医疗器械的安全性或有效性。算法框架同时发生多种类型的更新，同样遵循风险从高原则。

算法框架可参照自研软件、现成软件相关要求提交相应注册申报资料，均需明确算法框架的名称、类型（自研算法框架、现成算法框架）、型号规格、完整版本、制造商等信息。算法框架若基于云计算平台，相关要求详见前文关于云计算的要求。

（十六）人工智能芯片

人工智能芯片作为计算资源的组成部分，本身不属于监管对象，根据其所属的计算平台类型考虑监管要求。对于人工智能独立软件，人工智能芯片作为通用计算平台的组成部分，不属于产品结构组成，按照运行环境予以考虑。对于人工智能软

件组件，人工智能芯片作为医用计算平台的组成部分，属于产品结构组成，与产品进行整体评价。

无论何种情况，使用人工智能芯片均需在注册申报材料中明确其名称、型号规格、制造商、性能指标等信息。

六、算法研究资料

（一）算法研究报告

算法研究报告适用于人工智能算法或算法组合的初次发布和再次发布，包括算法基本信息、算法风险管理、算法需求规范、数据质控、算法训练、算法验证与确认、算法可追溯性分析、结论等内容，不适用内容详述理由。

1.算法基本信息

明确算法的名称、类型、结构、输入输出、流程图、算法框架、运行环境等基本信息以及算法选用依据。

其中，算法类型从学习策略、学习方法、可解释性等角度明确算法特性。算法结构明确算法的层数、参数规模等超参数信息。算法框架明确所用人工智能算法框架的基本信息，包括名称、类型（自研算法框架、现成算法框架）、型号规格、完整版本、制造商等信息；若基于云计算平台，明确云计算的名称、服务模式、部署模式、配置以及云服务商的名称、住所、服务资质。运行环境明确算法正常运行所需的典型运行环境，包括硬件配置、外部软件环境、网络条件；若使用人工智能芯片需明确其名称、型号规格、制造商、性能指标等信息。算法选用依据详述人工智能算法或算法组合选用的理由和基本原则。

2.算法风险管理

明确算法的软件安全性级别（轻微、中等、严重）并详述判定理由。提供算法风险管理资料，明确过拟合与欠拟合、假阴性与假阳性、数据污染与数据偏倚（如数据扩增）等风险的控制措施。若无单独文档可提供软件风险管理资料，并注明算法风险管理所在位置。

3.算法需求规范

提供算法需求规范文档，若无单独文档可提供软件需求规范，并注明算法需求所在位置。

4.数据质控

提供数据来源合规性声明，列明数据来源机构的名称、所在地域、数据收集量、伦理批件（或科研合作协议）编号等信息。

提供数据采集操作规范文档，根据数据采集方式明确采集设备、采集过程、数据脱敏等质控要求。

概述数据整理情况，明确数据清洗、数据预处理的质控要求。

提供数据标注操作规范文档，明确标注资源管理、标注过程质控、标注质量评估、数据安全保证等要求。

若适用，提供数据扩增情况说明，明确扩增的对象、方式、方法、倍数等信息。

依据适用人群、数据来源机构、采集设备、样本类型等因素，提供原始数据库、基础数据库、标注数据库、扩增数据库

关于疾病构成的数据分布情况。

若数据来自公开数据库，提供公开数据库的基本信息（如名称、创建者、数据总量等）和使用情况（如数据使用量、数据质量评估、数据分布等）。

5. 算法训练

依据适用人群、数据来源机构、采集设备、样本类型等因素，提供训练集、调优集（若有）关于疾病构成的数据分布情况。

明确算法训练所用的评估指标、训练方式、训练目标、调优方式（若有），提供 ROC 曲线或混淆矩阵等证据证明训练目标满足医疗要求，提供训练数据量-评估指标曲线等证据以证实算法训练的充分性和有效性。

6. 算法验证与确认

依据适用人群、数据来源机构、采集设备、样本类型等因素，提供测试集关于疾病构成的数据分布情况。

提供假阴性与假阳性、重复性与再现性、鲁棒性/健壮性、实时性等适用指标的算法性能评估结果，以证明算法性能满足算法设计目标。

若使用第三方数据库开展算法性能评估，提供第三方数据库的基本信息（如名称、创建者、数据总量等）和使用情况（如测试数据样本量、评估指标、评估结果等）。

若适用，提供算法性能影响因素分析报告，明确影响算法性能的主要因素及其影响程度，以及产品使用限制和必要警示

提示信息。

若适用，提供压力测试、对抗测试等测试报告。若未开展相应测试或测试结果不佳，均需明确产品使用限制和必要警示提示信息。

若基于测评数据库进行算法确认，提供测评数据库的基本信息（如名称、创建者、数据总量等）、评估情况（如评估方法、评估指标、评估结果等）、使用情况（如评估指标、评估结果等）。若基于临床评价方式进行算法确认，指向临床评价资料即可。

提供上述各类测试场景（含临床评价）下的算法性能评估结果比较分析报告，明确产品使用限制和必要警示提示信息。

7. 算法可追溯性分析

提供算法可追溯性分析报告，即追溯算法需求、算法设计、源代码（明确软件单元名称即可）、算法测试、算法风险管理的关系表。

若无单独文档可提供软件可追溯性分析报告，并注明算法可追溯性分析所在位置。

8. 结论

简述算法性能综合评价结果，明确对产品的适用范围、使用场景、核心功能所做的必要限制，并判定人工智能算法或算法组合的安全有效性是否满足要求。

（二）算法更新研究报告

算法更新研究报告仅适用于人工智能算法或算法组合的再次发布，在算法研究报告相应内容中明确算法更新情况。

其中，算法基本信息描述申报算法情况，若适用详述与前次注册相比的变化情况；算法风险管理、算法需求规范、数据质控、算法训练、算法验证与确认、算法可追溯性分析描述算法更新的具体情况；结论简述算法性能综合评价结果，并判定人工智能算法或算法组合更新的安全有效性是否满足要求。

考虑到算法更新具有累积效应，算法更新研究报告需涵盖人工智能医疗器械自前次注册（延续注册除外）以来算法更新的全部内容。

七、注册申报资料补充说明

注册申报资料在符合医疗器械注册申报资料要求等文件要求基础上，满足医疗器械软件、医疗器械网络安全、移动医疗器械等相关指导原则要求，同时重点关注以下要求。

（一）产品注册

1. 申请表信息

（1）人工智能独立软件

产品名称应符合通用名称命名规范要求，通常体现输入数据（如 CT 图像、眼底照片）、目标疾病（含病变、疾病的属性）、预期用途（如辅助分诊、辅助评估、辅助检测、辅助诊断）等特征词。

结构组成所述功能模块需保证用语的规范性，若采用人工智能算法需体现核心算法名称，如深度学习等。

适用范围基于预期用途、使用场景、核心功能予以规范，如处理对象、目标疾病、医疗用途、适用人群、目标用户、使

用场所、采集设备要求、使用限制等。

（2）人工智能软件组件

人工智能软件组件通常无需在注册证载明信息中体现。其软件功能名称可参照人工智能独立软件要求。若有辅助决策类软件功能，结构组成（若适用）和适用范围需予以体现。

2. 算法研究资料

对于软件安全性级别为中等、严重级别的产品，全新类型在软件研究资料中以算法为单位，提交每个人工智能算法或算法组合的算法研究报告，具体要求详见第六章；成熟类型在软件研究资料中明确算法基本信息即可，无需提供算法研究资料。

对于软件安全性级别为轻微级别的产品，在软件研究资料中明确算法基本信息即可，无需提供算法研究资料。

3. 用户培训方案

对于软件安全性级别为严重级别、预期由患者使用或在基层医疗机构使用的产品，原则上需单独提供一份用户培训方案，包括用户培训的计划、材料、方式、师资等。

4. 产品技术要求

产品技术要求若含有基于测评数据库测试的性能指标，需在“附录”中明确测评数据库的基本信息（如名称、型号规格、完整版本、责任方、主文档登记编号等）。

基于其他类型第三方数据库测试的性能指标，原则上无需在产品技术要求中体现。

5. 说明书

根据算法性能综合评价结果，对产品的适用范围、使用场景、核心功能进行必要限制，并在说明书中明确产品使用限制和必要警示提示信息。若适用，明确数据采集设备和数据采集过程相关要求。

对于辅助决策类产品，说明书需明确人工智能算法的算法性能评估总结（测试集基本信息、评估指标与结果）、临床评价总结（临床数据基本信息、评价指标与结果）、决策指标定义（或提供决策指标定义所依据的临床指南、专家共识等参考文献）等信息。此时若采用基于数据的人工智能算法，说明书还需补充算法训练总结信息（训练集基本信息、训练指标与结果）。

若产品采用人工智能黑盒算法，则需根据算法影响因素分析报告，在说明书中明确产品使用限制和必要警示提示信息。

（二）变更注册

1. 算法研究资料

对于软件安全性级别为中等、严重级别的产品，全新类型在软件研究资料中以算法为单位，根据人工智能算法的更新情况，提交每个人工智能算法或算法组合的算法更新研究报告（或算法研究报告），具体要求详见第六章；成熟类型在软件研究资料中明确算法基本信息即可，无需提供算法研究资料。

对于软件安全性级别为轻微级别的产品，在软件研究资料中明确算法基本信息即可，无需提供算法研究资料。

2. 用户培训方案

若适用，提交用户培训方案变化情况说明。

3.产品技术要求

若适用，产品技术要求变更对比表需体现测评数据库的变化情况⁵。

4.说明书

若适用，提交说明书变化情况说明。

（三）延续注册

延续注册通常无需提交算法相关研究资料。若适用，根据注册证“备注”所载明的要求提交相应算法研究资料。

八、参考文献

[1] 全国人大. 中华人民共和国个人信息保护法[Z], 2021.8

[2] 原国家食品药品监督管理总局.医疗器械说明书和标签管理规定（总局令第6号）[Z], 2014.7

[3] 国家市场监督管理总局.医疗器械注册与备案管理办法（总局令第47号）[Z], 2021.8

[4] 原国家食品药品监督管理总局. 医疗器械生产质量管理规范（2014年第64号公告）[Z], 2014.12

[5] 国家药品监督管理局.医疗器械主文档登记事项（2021年第36号公告）[Z], 2021.3

[6] 国家市场监督管理总局.医疗器械注册申报资料要求和批准证明文件格式（2021年第121号公告）[Z], 2021.9

[7] 原国家食品药品监督管理总局.移动医疗器械注册技术审查指导原则（2017年第222号通告）[Z], 2017.12

⁵测评数据库作为产品评价工具，其自身更新原则上无需申请变更注册，注册申请人可根据情况自行决定。

[8] 国家药品监督管理局. 医疗器械生产质量管理规范附录独立软件 (2019 年第 43 号通告) [Z], 2019.7

[9] 国家药品监督管理局. 医疗器械安全和性能的基本原则 (2020 年第 18 号通告) [Z], 2020.3

[10] 国家药品监督管理局. 人工智能医用软件产品分类界定指导原则 (2021 年第 47 号通告) [Z], 2021.7

[11] 国家药品监督管理局. 医用软件通用名称命名指导原则 (2021 年第 48 号通告) [Z], 2021.7

[12] 国家药品监督管理局. 医疗器械临床评价技术指导原则 (2021 年第 73 号通告) [Z], 2021.9

[13] 国家药品监督管理局. 医疗器械生产质量管理规范独立软件现场检查指导原则 (药监综械管〔2020〕57 号) [Z], 2020.5

[14] 国家药品监督管理局医疗器械技术审评中心. 医疗器械软件技术审查指导原则 (第二版) (征求意见稿) [Z], 2020.6

[15] 国家药品监督管理局医疗器械技术审评中心. 医疗器械网络安全技术审查指导原则 (第二版) (征求意见稿) [Z], 2020.9

[16] 国家药品监督管理局医疗器械技术审评中心. 医疗器械人因设计技术审查指导原则 (征求意见稿) [Z], 2020.5

[17] 国家药品监督管理局医疗器械技术审评中心. 深度学习辅助决策医疗器械软件审评要点 (2019 年第 7 号通告) [Z], 2019.7

[18] 国家药品监督管理局医疗器械技术审评中心. 肺炎

CT 影像辅助分诊与评估软件审评要点（试行）（2020 年第 8 号通告）[Z]，2020.3

[19] 国家卫生健康委员会. 人工智能辅助诊断技术管理规范（国卫办医发〔2017〕7 号）[Z]，2017.2

[20] 国家卫生健康委员会. 人工智能辅助诊断技术临床应用质量控制指标（国卫办医发〔2017〕7 号）[Z]，2017.2

[21] 国家卫生健康委员会. 人工智能辅助治疗技术管理规范（国卫办医发〔2017〕7 号）[Z]，2017.2

[22] 国家卫生健康委员会. 人工智能辅助治疗技术临床应用质量控制指标（国卫办医发〔2017〕7 号）[Z]，2017.2

[23] GB/T 25000.12-2017 系统与软件工程 系统与软件质量要求和评价（SQuaRE）第 12 部分：数据质量模型[S]

[24] YY/T 0287-2017 医疗器械 质量管理体系 用于法规的要求[S]

[25] YY/T 0316-2016 医疗器械 风险管理对医疗器械的应用[S]

[26] YY/T 0664-2020 医疗器械软件 软件生存周期过程[S]

[27] YY/T 1406.1-2016 医疗器械软件 第 1 部分：YY/T 0316 应用于医疗器械软件的指南[S]

[28] YY/T 人工智能医疗器械质量要求和评价 第 1 部分：术语（报批稿）[S]，2020.12

[29] YY/T 人工智能医疗器械质量要求和评价 第 2 部分：数据集通用要求（报批稿）[S]，2020.12

[30] YY/T 人工智能医疗器械质量要求和评价 第 3 部分：数据标注通用要求（报批稿）[S]，2021.12

[31] YY/T 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（报批稿）[S]，2021.12

[32] AIMDICP-WG6-2020-001 基于眼底彩照的糖尿病视网膜病变辅助决策产品性能指标和测试方法[S]，2020.7

[33] AIMDICP-WG6-2020-002 基于胸部 CT 的肺结节影像辅助决策产品性能指标和测试方法[S]，2020.7

[34] IMDRF/AIMD WG/N67, Machine Learning-enabled Medical Devices - A subset of Artificial Intelligence-enabled Medical Devices: Key Terms and Definitions[Z], 2021.9

[35] FDA. Computer-Assisted Detection Devices Applied to Radiology Images and Radiology Device Data[Z], 2012.7

[36] FDA. Considerations for Computer-Assisted Detection Devices Applied to Radiology Images and Radiology Device Data[Z], 2012.7

[37] FDA, Developing a Software Precertification Program: A Working Model[Z], 2019.1

[38] FDA. Proposed Regulatory Framework for Modifications to Artificial Intelligence/Machine Learning(AI/ML) - Based Software as a Medical Device(SaMD) (Draft Guidance)[Z], 2019.4

[39] FDA, Clinical Decision Support Software (Draft Guidance)[Z], 2019.9

[40] FDA. Artificial Intelligence and Machine Learning(AI/ML) Software as a Medical Device(SaMD) Action Plan[Z], 2021.1

[41] FDA/Health Canada/MHRA. Good Machine Learning Practice for Medical Device Development: Guiding Principles[Z], 2021.10

[42] FDA. Technical Considerations for Medical Devices with Physiologic Closed-Loop Control Technology(Draft Guidance)[Z], 2021.12

[43] AAMI/BSI. Artificial Intelligence and Machine Learning - the Emergence of Artificial Intelligence and Machine Learning Algorithmsin Healthcare: Recommendations to Support Governance and Regulation[Z], 2019

[44] AAMI/BSI. Machine Learning AI in Medical Devices: Adapting Regulatory Frameworks and Standards to Ensure Safety andPerformance[Z], 2020

[45] BSI. Recent Advancements in AI - Implications for Medical Device Technology and Certification[Z], 2020

[46] ISO/IEC TR 5469(WD) Artificial intelligence-Functional safety and AI systems[S]

[47] ISO/IEC TR 24028:2020Information technology - Artificial intelligence - Overview of trustworthiness in artificial intelligence[S]

[48] ISO/IEC TR 24029-1:2021 Artificial Intelligence(AI) -

Assessment of the robustness of neural networks - Part 1: Overview[S]

[49] ISO/IEC TR 24030:2021 Information technology - Artificial intelligence(AI) - Use cases[S]

[50] ISO/IEC TR 29119-11:2020 Software and systems engineering - Software testing - Part 11: Guidelines on the testing of AI-based systems[S]

[51] 人工智能医疗器械创新合作平台 [Z].
<https://www.aimd.org.cn>

[52]IMDRF AIMD WG[Z]. <https://www.imdrf.org/workitems/wi-aimd.asp>

[53] ITU/WHO FG-AI4H [Z]. <https://www.itu.int/go/fgai4h>

[54] IEEE AIMD WG[Z]. <https://sagroups.ieee.org/aimdwg>