附件2

病理图像人工智能分析软件性能评价审评要点

本要点旨在指导注册申请人对病理图像人工智能分析软件注册申报资料中非临床评价部分的准备及撰写，同时也为技术审评部门提供参考。

本要点是对病理图像人工智能分析软件申报资料的一般要求，申请人需依据产品的具体特性确定其中内容是否适用。若不适用，需具体阐述理由及相应的科学依据，并依据产品的具体特性对注册申报资料的内容进行充实和细化。

本要点是供注册申请人和技术审评人员使用的指导性文件，但不包括审评审批所涉及的行政事项，亦不作为法规强制执行，需在遵循相关法规的前提下使用本指导原则。如果有能够满足相关法规要求的其他方法，也可以采用，但是需要提供详细的研究资料和验证资料。

本要点是在现行法规和标准体系以及当前认知水平下制定的，随着法规和标准的不断完善，以及科学技术的不断发展，本要点的相关内容也将进行适时的调整。

一、适用范围

病理图像人工智能分析软件是指基于数字病理图像，采用深度学习等人工智能技术实现对病理图像的分割、检测等功能的医疗器械**。**数字病理图像包括经数据成像设备获得的显微镜下细胞或组织图像、全玻片数字扫描技术（Whole Slide Imaging, WSI）的病理图像等。产品在医疗机构和/或医学实验室使用，可辅助病理医师为疾病的诊断、预后、治疗等提供信息，不能作为临床诊断决策的唯一依据。

根据《人工智能医用软件产品分类界定指导原则》，本文所述软件用于辅助决策，可为病理医生提供病灶特征识别、病变性质等信息，按第三类医疗器械管理。

分类编码：21-04-02。

产品名称：建议依据软件医疗用途进行产品名称命名，体现处理对象和临床用途，如宫颈细胞学数字病理图像计算机辅助分析软件、免疫组化病理图像计算机辅助诊断软件。申请人可结合申报产品的特征，根据《医用软件通用名称命名指导原则》，参考《医疗器械分类目录》中的品名举例确定命名。

若作为软件组件集成于其他医疗器械（如病理切片扫描仪、医学显微图像扫描分析产品）中，可以参考本要点的要求。

本要点基于《人工智能医疗器械注册审查指导原则》的基本框架要求，其中关于人工智能医疗器械的的术语定义、相关要求参照上述指导原则。

二、产品简介

产品的工作原理：产品为独立软件，采用深度学习技术对数字病理图像进行前处理（如图像质量改善）和常规后处理（如图像分割、细胞/组织定位、细胞分类等），实现对数字病理图像中疑似异常细胞的计数和定位识别功能或异常组织区域的定位和标记功能；通过算法对细胞/组织图像块的分类结果进行分析，实现对病理图像的辅助诊断功能。

主要结构及组成：结构组成明确交付内容和功能模块，其中交付内容包括软件安装程序、授权文件、外部软件环境安装程序等软件程序文件，功能模块包括客户端、服务器端（若适用）等，若适用注明选装、模块版本。

服务器端：负责接收并处理请求，对请求进行数据处理，调用人工智能算法（此处应体现核心算法名称，如AlexNet，ResNet，MobileNet，EfficientNet，ResNeXt等）对图像数据进行处理并输出分析结果。把用户请求的数据结果返回客户端。

客户端：将服务器端分析结果呈现到用户界面。用户可在用户界面进行辅助诊断/阅片、任务管理、报告管理、检索查询、用户管理、查看日志等功能操作。

三、基本要求

（一）综述资料

详细描述该产品的图像前处理和后处理功能及功能实现方法，描述配套使用图像采集设备、切片制作流程（包括配套试剂、设备）和图像质量的要求，简述软件采用的主要算法。

简述软件研发背景，包括对可辅助诊断疾病的介绍，目前相关疾病临床诊断参考标准、病理诊断类型或细胞异常类型等依据。申报产品的病理阅片标准和病理分型依据应与临床标准一致。

（二）非临床资料

1.产品技术要求及检验报告

产品技术要求可参考《医疗器械软件注册审查指导原则》（2022年修订版）附录中的模板进行编写，此外性能指标还应包括“临床功能”，如图像处理、图像分析、细胞/组织识别、病理图像诊断提示等内容。

2.软件研究资料

按照《医疗器械软件注册审查指导原则》（2022年修订版）、《医疗器械网络安全注册审查指导原则》（2022年修订版）、《人工智能医疗器械注册审查指导原则》的要求，提交研究资料。明确软件发布版本号。提交的研究资料主要包含软件研究资料、算法研究资料、网络安全研究资料。研究资料内容应涵盖申报产品适用范围中描述的全部内容。

2.1风险管理

参考YY/T 0316、YY/T 1406.1和YY/T 0664、GB/T 42062进行产品的风险管理活动。风险管理活动应当基于软件的预期用途、使用场景、核心功能予以实施，并贯穿软件全生命周期过程。此类产品临床使用中的风险主要包括图像分析过程中的识别、标记错误，非病理医师使用和样本采集错误等，造成的假阴性和假阳性错误。算法风险主要包括过拟合和欠拟合，数据扩增和数据偏移等。

2.2需求规范

提供软件需求规范文档，明确软件的功能、性能、接口、用户界面、运行环境、网络安全、数据采集等需求，另附软件开发所形成的原始文档。

数据采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。数据应来源于不同地域的不少于3家机构。机构应采用软件说明书明确的切片制作流程、组织染色和免疫组织化学技术制备的切片（包含指定的样本来源，即指定的切片制备流程和/或样本保存液、抗体试剂、染色液等）和配套的图像采集设备获得的符合要求的数字病理图片。数据分布的科学性和合理性，应考虑包括但不限于疾病构成（如分型、分级、分期）、人群分布（如高危人群、患者，性别、年龄）等情况。

算法性能需结合产品预期用途，综合考虑分析速度、敏感性、特异性、重复性与再现性、泛化性等性能指标的适用性及其要求。同时还需考虑因梯度消失、梯度爆炸、过拟合和欠拟合等影响算法性能的因素。

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。

2.3软件性能研究

提供产品性能研究资料以及产品技术要求的研究和编制说明，给出软件相关的功能性、安全性指标的确定依据。注册申请人需在软件研究资料中提交GB/T 25000.51自测报告，亦可提交自检报告或检验报告代替自测报告。产品技术要求中各指标依据产品自身特点确定。

2.4算法研究资料

根据《人工智能医疗器械注册审查指导原则》提交算法研究资料。提交每个人工智能算法或算法组合的算法研究报告。

2.4.1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（基于模型的算法和基于数据的算法）、可解释性（白盒算法和黑盒算法）等角度明确算法特性。算法结构部分应明确算法的层数、参数规模等超参数信息。流程图应包含输入图像示例、学习方法示意图和输出图像示例。算法框架应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，应明确云计算的名称、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。运行环境应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应明确其名称、型号规格、制造商、性能指标等信息。算法选用依据详述算法或算法组合选用的理由和基本原则。

2.4.2算法风险管理

结合预期用途、使用场景、核心功能、算法成熟度、网络安全等因素，此类产品的软件安全性级别为严重级别。提供算法风险管理资料，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。

2.4.3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。

2.4.4数据收集

2.4.4.1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。

2.4.4.2数据采集

建议参考2.2需求规范的要求，提供数据采集操作规范文档，包括数据采集方案和数据采集标准操作规程。

数据采集主要由临床机构实施，应明确病理切片或细胞涂片的制作流程、图像采集方式（如眀场扫描）和采集参数要求（包含扫描倍率、扫描层数、焦点数量要求、聚焦方式、采样阈值等）、采集设备要求（应包含扫描仪的型号、分辨率、图像尺寸等）、病理切片或细胞涂片要求、样本的来源与分布、样本质量要求（包含切片或涂片的保存时限、玻片完整性等）、图像质量要求（应包含图象区域、清晰度等要求）、采集过程、数据脱敏、数据转移等要求。采集过程应对样本数据进行编号并加密，方案中应包含编号规则。

2.4.4.3数据整理

明确数据清洗/预处理程序，对数据处理中应用的软件进行简述，并以附件的形式提交数据处理中各软件的软件研究资料。

2.4.4.4数据标注

明确标注人员和仲裁人员的资质要求和培训内容，标注人员和仲裁人员应为病理医师，数据应经过不少于2人标注。简述标记系统信息，配合病理图像和标注系统界面截图，介绍标注过程、标注对象和标注标准（注意与病理诊断标准的一致性）等。明确标注过程质控、标注质量评估、数据安全保证等要求。标注质量评估可抽选一定比例数据进行非标注人员的评估。

提供原始数据库、基础数据库、标注数据库、扩增数据库（如有），关于疾病构成的数据分布情况，包括适用人群、数据来源机构、采集设备、样本类型等因素。

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）。

2.4.4.5数据集构建

明确各数据集划分的方法及依据。训练集应当保证样本分布具有均衡性，训练样本应涵盖不同病理类型。调优集应保证样本分布符合临床实际情况，如阳性比例，不同病理类型比例、不同临床分型比例等均应符合临床实际情况。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。

进行数据扩增验证时，应当明确扩增的方式、方法、倍数，并考虑数据扩增对软件的影响及风险。列表比较扩增数据库与标注数据库的差异，论证扩增数据库样本量的充分性以及分布的合理性。

2.4.5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。

算法训练基于训练集、调优集进行训练和调优，应明确算法训练所用的评估指标、训练方式、训练目标、调优方法，提供ROC曲线或混淆矩阵等证据（如：迭代次数-训练CountIOU曲线和迭代册数-召回率曲线）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（如迭代次数-Loss曲线）等证据以证实算法训练的充分性和有效性。

2.4.6算法性能评估

基于测试集对算法设计进行评估，确认软件算法性能的效率、敏感性、特异性，性能应满足算法设计要求。

算法性能评估应包括软件对样本满意度评价的能力测试、泛化能力的测试、压力测试（指采用罕见或特殊的真实数据样本开展的算法性能测试）、对抗测试、重复性与再现性测试、敏感性与特异性测试、分析效率测试、算法性能影响因素分析、性能评估结果比较分析、偏差报告等研究。

通过样本量估算确定测试集中阴、阳性样本和压力样本的样本量，提供测试集关于疾病构成（包括年龄和病理类型、异常细胞类型等）、数据来源的数据分布情况。明确对抗样本的选择原则；明确性能评估的可接受标准和标准制定依据。提交测试报告和细胞/组织形态图示例。

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）。提交第三方数据库数据样本来源和本计算机辅助分析软件产品的产品技术要求、说明书中指定的制片流程和图像采集设备的一致性的评价（包含指定的样本来源，即指定的切片制备流程和/或样本保存液、抗体试剂、染色液等）和配套的图像采集设备获得的符合要求的数字病理图片。

根据研究结论，确定软件使用限制并在说明书中进行描述。

2.4.7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。

（三）产品说明书和标签样稿

明确该产品适用范围，明确产品适用的显微镜和图像采集设备的名称和型号，明确图像参数及质量。明确病理切片或细胞涂片的制备流程及试剂厂家信息。明确软件报告内容。对产品带来的假阳/假阴性风险进行提示。

根据算法影响因素分析报告，在说明书明确产品使用限制和必要的警示提示信息。列明算法训练总结和算法性能评估总结以及临床评价总结。