

T/CVDA

团 体 标 准

T/CVDA XX—XXXX

宠物影像智能辅助诊断软件通用要求

General Requirements for Intelligent Assisted Diagnosis Software in
Veterinary Imaging

（征求意见稿）

XXXX—XX—XX 发布

XXXX—XX—XX 实施

中国兽药协会 发布

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。

请注意本文件的某些内容可能涉及专利。本文件的发布机构不承担识别专利的责任。

本文件为技术性团体标准，不作为行政审批、强制监管或执法依据，其内容主要用于指导软件研发、测试评价、临床应用与行业自律。

本文件由中国兽药协会提出并归口管理。

本文件主要起草单位：深圳市晓闻智慧医疗科技有限公司。

本文件主要起草人：刘欣怡，刘思莹。

宠物影像智能辅助诊断软件通用要求

1 范围

本文件适用于以犬、猫等伴侣动物为主要对象的影像智能辅助诊断软件。

本文件适用于对 X 线、超声、CT、MRI 及其他经声明支持的医学影像进行分析处理，并向执业兽医提供辅助检测、辅助分割、征象描述、量化测量、可疑异常提示、鉴别诊断线索、结构化输出建议等辅助决策信息的软件系统。

软件开发方应明确系统经验证的适用边界，包括但不限于：适用物种、品种或体型组、年龄组、影像模态、检查部位、采集体位、输入图像质量条件、设备类型、是否整合非影像临床数据及输出任务类型。

对超出已验证范围的输入，系统不应输出确定性结论，宜输出“不适用”“证据不足”或“建议人工复核/进一步检查”等提示。

本文件不适用于直接替代执业兽医作出最终诊断、处方、治疗决策或法律责任承担的系统。

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。

GB/T 41867—2022 信息技术 人工智能 术语

GB/T 35273—2020 信息安全技术 个人信息安全规范

GB/T 42062—2022 医疗器械 风险管理对医疗器械的应用

YY/T 0664—2020 医疗器械软件 软件生存周期过程

GB/T 25000.10—2016 系统与软件工程 系统与软件质量要求和评价（SQuaRE）第 10 部分：系统与软件质量模型

GB/T 25000.40—2018 系统与软件工程 系统与软件质量要求和评价（SQuaRE）第 40 部分：评价过程

GB/T 22239—2019 信息安全技术 网络安全等级保护基本要求

DICOM Standard (Digital Imaging and Communications in Medicine), relevant parts of the current edition

3 术语和定义

下列术语和定义适用于本文件。

3.1 宠物影像智能辅助诊断系统

利用人工智能等技术，对宠物医学影像进行分析处理，并为执业兽医提供辅助检测、辅助识别、可疑异常提示、征象描述、量化测量、鉴别诊断线索或结构化输出建议等决策支持信息的软件系统。

该系统可针对单幅影像、影像序列或整套检查进行分析，必要时可结合相关非影像临床数据进行综合辅助判断。

3.2 动物诊疗机构

依法取得动物诊疗许可证，从事动物疾病诊断、治疗和动物绝育手术等经营活动的机构，包括动物医院、兽医诊所等。

3.3 执业兽医

经国家正规兽医资格考试合格，依法注册或备案，在动物诊疗机构中负责疾病诊断、开具处方和承担最终诊疗责任的专业人员。

3.4 感兴趣区

医学影像中与特定诊断问题相关或被用户（执业兽医）所关注的区域或图像序列范围。

3.5 分割

从医学影像中提取感兴趣区的过程。注：通常包括自动分割、半自动分割和手动分割。

3.6 算法可追溯性

能够追溯和复现智能辅助诊断算法从训练数据构成、标注、模型训练到性能验证全过程的属性。这对于算法的可靠性评估和持续改进至关重要。

3.7 金标准

在特定时间和医学知识背景下，被公认的、可重复的、准确性最高的疾病诊断方法或标准。在宠物影像领域，可包括组织病理学检查、长期随访的最终临床诊断、或由多名高级别影像专家形成的独立共识。在宠物影像领域，金标准的选择应充分考虑动物种属、品种差异及伦理可行性。

3.8 多中心数据集

来源于两个及以上相互独立的动物诊疗机构或研究中心的影像数据集合。此类数据集对于提高算法泛化能力、减少数据选择偏倚及提升最终评价结果的可靠度具有关键作用。

3.9 准确率

在分类任务中，被测试系统做出正确判断的样本数占测试样本总数的比例。

3.10 敏感度

实际患病的样本中，被测试系统正确识别为阳性的比例。

3.11 特异度

实际未患病的样本中，被测试系统正确识别为阴性的比例。

3.12 全检查分析

针对同一次检查中的单幅影像、影像序列或整套影像资料进行整体分析，并形成病例级或发现级辅助输出的过程。

3.13 目标异常

在系统预期用途、性能声明或验证方案中明确纳入识别、分析或提示范围的病变、异常或征象。

3.14 非目标可疑异常

系统在整套影像分析过程中额外提示的、未明确纳入正式性能声明范围的其他异常或可疑发现。

3.15 预训练模型

在大规模通用或专项数据上预先训练，并可经微调、适配、蒸馏或其他方式应用于宠物影像智能辅助诊断系统的模型。

3.16 适用范围边界

系统经设计、验证和声明后明确限定的适用条件，包括但不限于物种、影像模态、检查部位、疾病类型、临床场景、输入数据类型和输出内容范围。

3.17 模型迭代

基于新数据或新算法对已有影像认知模型进行重新训练、优化和版本更新的过程，旨在持续提升模型的性能与适应性。

4 基本原则

4.1 合法合规

系统的设计、开发、部署、应用和运维，应符合国家关于动物诊疗、执业兽医管理、病历书写、处方管理、动物防疫、网络安全、数据安全及个人信息保护等现行法律法规和标准规范要求。

系统应用场景涉及动物诊疗机构内部流程管理、报告生成、病历留存、处方关联、远程服务或第三方平台集成时，应与现行动物诊疗管理制度和宠物诊疗服务规范保持一致。

软件开发方不应作出超出证据支持的安全性、有效性或替代诊断能力声明。

4.2 尊重动物福利

系统设计与应用应以动物福利和检查安全为前提。

软件不应通过默认流程、界面提示或性能声明，诱导使用者实施不必要的重复采集、延长保定时间、追加非必要镇静/麻醉或其他可能增加动物风险的操作。

对需要镇静或麻醉配合完成的影像检查，系统应支持记录相关检查条件，并提示用户结合患者状态和临床风险进行综合判断。

当输入影像不满足分析要求时，系统宜优先提示人工复核、补充采集或改用其他检查方式，而非直接输出确定性排除性结论。

4.3 算法科学及可溯性原则

系统应公开、明确并可核查地说明其适用边界、主要功能、输出类型、已验证数据范围、已知薄弱场景及主要限制条件。

系统不应将未验证场景下的结果表述为已证实有效的输出。

对不同类型输出内容，应明确其证据等级及使用限制，不应将辅助提示、生成性描述与经充分验证的定量测量或检测结果混同表述。

4.4 不确定性与安全输出原则

当输入数据质量不足、信息不完整、超出适用边界，或模型证据不足以支持明确判断时，系统应输出不确定性说明、人工复核提示、补充检查建议或拒绝输出确定性结论。

系统不应将推测性内容、低置信内容或证据不足内容表述为明确诊断结论。

对排除性、否定性或“未见异常”类输出，应在满足相应适用边界和质量前提时方可提供。

4.5 有效性和临床导向原则

软件的核心功能应具有明确的临床价值，其有效性声明应有充分客观证据支持。

对于系统输出的不同类型内容，例如目标异常检测结果、非目标可疑异常提示、征象描述、量化测量、报告草拟建议等，应根据其证据水平分别开展验证，不应将不同证据等级的输出内容混同表述。

临床应用中，系统输出应服务于执业兽医的判读、审核和决策支持，不应削弱必要的人工审查环节。

4.6 隐私保护和数据安全原则

应采取必要的数据安全与隐私保护措施，确保宠物相关信息、宠物主个人信息及诊疗数据在采集、传输、存储、处理、调用和共享过程中的安全。

对采用云端部署、远程推理、外部模型服务调用或第三方接口集成的系统，应明确数据脱敏、访问控制、传输加密、操作留痕、异常处置和责任边界要求。

5 总体要求

5.1 基本要求

宠物影像智能辅助诊断软件的设计、开发、部署、应用和评价，应围绕其声明用途开展，并与适用对象、输入条件、任务边界、输出形式和使用方式保持一致。

软件不应脱离声明用途被设计、宣传或使用为超出其预期范围的工具。

5.2 声明一致性要求

软件的功能设置、处理流程、结果展示、人工审核要求、记录要求和运行控制要求，应与其声明用途及相关说明保持一致。

软件不应通过默认设置、功能组合、附加模块或其他方式，形成超出声明范围的实际使用效果。

5.3 边界受控要求

软件应对其适用范围、适用条件、不适用情形和使用限制予以明确，并在声明范围内发挥辅助作用。

对于超出声明用途、超出适用范围、依据不足或高不确定性的情形，不应按声明范围内结果的方式输出或使用。

涉及边界外场景时，宜通过提示、限制、人工复核要求或其他适当方式降低误用风险。

5.4 算法与临床评价要求

软件输出应与输入信息、处理任务、形成依据和结果性质相一致。

不同类型的输出内容，如目标异常检测结果、非目标可疑异常提示、征象描述、量化测量、报告草案等，应根据其性质、证据水平和使用条件予以区分。

客观描述、分析性提示、机器生成草案和正式使用结果之间不应混同表述。

5.5 人工审核要求

软件输出应与其风险程度和声明要求相适应。

对声明需经人工审核、人工确认或人工复核后方可使用的内容，不应在未经审核确认的情况下直接进入正式使用环节。

软件输出不应替代执业医师的专业判断，不应脱离人工审核直接作为最终诊断、正式报告或其他具有独立临床后果的依据。

5.6 记录与追溯要求

软件应具备与其声明用途相适应的记录和追溯能力。

涉及结果形成、人工修订、最终确认及相关版本变化的重要信息，宜能够被记录、区分和追溯。

相关记录应支持结果复核、问题分析、责任界定和版本回顾。

5.7 验证衔接要求

本文件所提出的声明用途、任务边界、输出要求、人工审核要求、运行控制要求和记录要求，应能够转化为可验证、可测试、可复核的内容。

软件的设计与实现应便于对相关要求开展验证与测试，不应形成无法验证或难以复核的要求表述。

6 使用流程与输出要求

6.1 预期用途与目标用户

6.1.1 预期用途

a) 影像征象提取：对影像中的形态学特征、密度、信号、回声、边界、分布及其他可观察征象进行识别、提取或描述；

b) 异常检测/定位：对目标异常、可疑病灶或异常区域进行检测、分割、定位、标记或量化测量；

c) 鉴别诊断提示：基于影像表现，并结合经声明允许接入的相关临床信息，提供有限范围内的鉴别诊断线索、风险提示或进一步检查建议；

d) 结构化报告草案生成：依据分析结果生成结构化描述、初步结论草案或报告模板内容，供用户审核、修改和确认；

e) 多模态辅助判读：在声明支持的条件下，结合影像、病史摘要、实验室检查结果、既往检查结果或其他临床信息，提供综合辅助分析结果。

软件不应仅以“辅助诊断”作为笼统用途表述，而应明确其任务类型、输出形式、适用条件及使用限制。

6.1.2 目标用户

软件的目标用户应为执业兽医师，或在执业兽医师监督下开展相关操作的专业技术人员。

软件输出仅应用于辅助判读、辅助分析和辅助报告，不应替代执业兽医师作出最终诊断、处方、治疗决策或其他具有独立临床后果的决定。

输入影像的选择、提交及其是否适合作为本次辅助判读依据，应主要由具备相应专业能力的使用者结合诊疗目的作出判断。软件接收并处理输入影像，不应视为对该输入临床适用性的确认。

软件宜根据不同用户角色明确其功能使用范围、审核权限和责任边界。

6.1.3 用途边界与使用限制

软件应声明其支持的使用场景、任务边界和不适用情形。声明内容至少应包括：

- a) 支持的物种、影像模态、检查部位和任务类型；
- b) 允许接入的非影像临床信息类型；
- c) 允许输出的结果类型及其适用限制；
- d) 不属于预期用途的任务、场景或用户群体；
- e) 应由人工复核、不得直接采信或不得输出确定性结论的情形。

对于软件能够判断为超出声明用途、超出适用范围边界或不符合目标用户使用条件的情形，软件应限制相关功能调用，或输出“不适用”“需人工复核”“不能生成确定性结果”等提示。

对于虽然属于声明边界外、但软件不能判断的情形，相关控制责任主要由使用者结合软件声明和临床场景承担。

6.2 输入范围与任务边界

6.2.1 输入范围

宠物影像智能辅助诊断软件应明确声明其支持的输入类型、输入条件和输入质量要求。输入范围宜包括下列一种或多种：

- a) 影像数据，包括 X 线、超声、CT、MRI 及其他经声明支持的影像模态；
- b) 与影像配套的检查信息，包括检查部位、体位、序列、采集参数或其他必要检查说明；
- c) 经声明允许接入的非影像临床信息，包括病史摘要、体征信息、实验室检查结果、既往检查结果、既往报告或转诊信息；
- d) 经声明允许调用的知识支持信息，包括术语库、知识库或其他辅助检索信息。

软件应明确不同输入类型与相应功能之间的对应关系。未声明支持的输入类型，不应作为生成分析结果的依据。

输入影像与诊疗对象的一致性、检查目的相关性、来源可靠性及是否适合作为本次辅助判读依据，主要由使用者负责。

6.2.2 输入要求

软件应对输入的完整性、规范性和可用性提出要求。至少应明确下列内容：

- a) 支持的物种范围；
- b) 支持的影像模态、检查部位和任务类型；
- c) 影像格式、分辨率及其他技术处理要求；
- d) 非影像临床信息的类型、来源和结构化程度要求；
- e) 多模态输入时各类信息的组合条件和优先级。

当输入信息不完整、不规范、不满足声明输入条件或技术上无法处理时，软件应提示用户补充、修正、复核或限制相关处理；必要时应限制生成确定性结果。

对虽存在一定质量缺陷但未明显超出声明输入条件、且系统仍可完成处理的影像，软件可继续输出辅助结果；必要时宜提示相关局限信息。是否采信该结果，应由使用者结合原始影像和临床背景进行判断。

6.2.3 任务边界

软件应根据声明用途明确任务边界，并将任务边界与输出范围对应设置。任务边界至少应包括：

- a) 支持的物种边界；
- b) 支持的影像模态边界；
- c) 支持的检查部位边界；

- d) 支持的任务类型边界；
- e) 支持的临床信息接入边界；
- f) 支持的输出类型边界。

软件不应在超出声明边界的条件下输出确定性分析结论，不应将未声明支持的信息来源作为主要推理依据。

6.2.4 不适用情形

软件应明确声明不适用情形。下列情形宜列为不适用或限制使用情形：

- a) 依据软件声明用途判断，输入对应的物种、影像模态、检查部位或任务类型不属于支持范围；
- b) 关键影像缺失、关键体位不足或关键序列不完整，且不足以支持声明任务；
- c) 输入存在软件无法处理的情况，或影像质量严重受损，已明显影响系统完成声明任务；
- d) 非影像临床信息来源不明、内容冲突或无法验证，且已明显影响相关推理任务的可靠性；
- e) 任务目标超出声明用途，如要求直接输出最终诊断、治疗方案或替代人工出具正式报告；
- f) 其他经验证可能显著影响输出可靠性的情形。

对虽存在一定缺陷但未明显超出声明输入条件、且仍可由专业人员结合原始影像和临床背景进行判断的输入，不宜一概列为不适用情形。

对于属于声明不支持范围、但软件本身不能判断的情形，不宜要求软件自动识别并处置。

6.2.5 边界外处置要求

对于软件能够判断为超出声明用途、超出适用范围、不适用，或者无法处理的输入，软件应采取与风险相适应的处置措施，包括但不限于：

- a) 限制相关功能调用；
- b) 提示“不适用”“超出支持范围”或“无法处理”；
- c) 提示“需人工复核”或“信息不足”；
- d) 仅输出描述性信息，不输出确定性判断；
- e) 拒绝生成相关结果。

对于虽然实际上超出声明范围，但软件本身不能判断的输入，不宜要求软件自动识别并处置；该类情形主要由使用者根据软件声明用途和临床场景进行判断。

对虽存在一定质量缺陷但未明显超出声明输入条件、且系统仍可完成处理的输入，软件可不采取拒绝处置，但宜根据风险提示相关局限信息或降低输出确定性。

6.2.6 外部知识与附加信息使用要求

软件声明支持外部知识、术语库或其他附加信息时，应明确其用途范围和使用限制。

外部知识、附加信息或非影像文本内容可以用于辅助提示、辅助解释或辅助生成报告草案，但不应在缺乏相应影像依据时单独支持确定性影像结论。

对结合附加信息形成的输出，宜能够区分其信息来源，并与影像直接观察结果相区分。

6.3 输出分层

6.3.1 输出分层原则

宠物影像智能辅助诊断软件的输出应与其声明用途、任务边界和风险等级相适应，并宜按照输出内容性质划分为感知层输出、推理层输出和文书层输出。

不同层级输出的功能定位、使用方式、审核要求和可采信范围应予区分，不应将不同层级输出混同为单一“辅助诊断结论”。

6.3.2 感知层输出

感知层输出是指基于影像直接观察结果形成的客观分析结果，宜包括下列一种或多种：

- a) 异常检测结果；
- b) 目标分割结果；
- c) 病灶或异常区域定位结果；
- d) 长度、面积、体积、角度或其他量化测量结果；
- e) 影像征象标注或结构化描述结果。

感知层输出应尽可能保持与影像可观察信息的一致性，并宜采用标注、框选、分割掩膜、测量值、结构化字段或其他便于核查的形式呈现。

6.3.3 推理层输出

推理层输出是指在感知层输出基础上，结合经声明允许接入的非影像临床信息所形成的分析性结果，宜包括下列一种或多种：

- a) 征象归纳；
- b) 鉴别诊断提示；
- c) 风险提示；
- d) 进一步检查建议；

e) 其他经声明支持的辅助分析结果。

推理层输出应与其依据相对应，并应与直接观察结果相区分。

推理层输出不应表述为未经审核的最终诊断结论，不应替代执业兽医师的专业判断。

6.3.4 文书层输出

文书层输出是指依据感知层输出和推理层输出生成的报告性文本结果，宜包括下列一种或多种：

- a) 结构化描述文本；
- b) 影像所见草案；
- c) 初步结论草案；
- d) 报告模板内容；
- e) 其他经声明支持的报告辅助内容。

文书层输出应明确标识其机器生成属性，并应供用户审核、修改和确认后后方可用于正式报告或病历记录。

文书层输出不应在未经人工审核的情况下直接作为正式诊断报告、正式病历结论或其他具有独立临床后果的文书内容。

6.3.5 分层输出对应关系

软件宜明确不同任务类型对应的输出层级和输出形式。至少宜满足下列要求：

- a) 以异常检测、分割、定位和测量为主要功能的，输出应以感知层输出为主；
- b) 以征象归纳、鉴别诊断提示和进一步检查建议为主要功能的，输出应以推理层输出为主；
- c) 以报告辅助生成为主要功能的，输出应以文书层输出为主，并宜同时保留相应的感知层或推理层依据；
- d) 涉及多层输出的，应明确各层输出之间的生成关系和审核要求。

6.3.6 不同层级输出的使用要求

不同层级输出应实施差异化使用要求：

- a) 感知层输出可作为影像核查和辅助判读参考；
- b) 推理层输出应作为提示性信息使用，并结合影像依据和临床背景进行人工复核；
- c) 文书层输出应作为待审核草案使用，不应直接替代正式报告撰写；

d) 对高风险场景、高不确定性场景或边界外场景，软件宜降低输出层级，必要时仅输出感知层结果或拒绝输出推理层、文书层结果。

6.3.7 输出呈现要求

软件应根据输出层级采用与其性质相适应的呈现方式，并宜满足下列要求：

- a) 感知层输出宜与原始影像关联展示；
- b) 推理层输出宜与其主要依据相对应，并与客观描述相区分；
- c) 文书层输出宜与正式报告内容在界面、标识或流程上予以区分；
- d) 不同层级输出不应通过界面呈现或文案表述造成其法律效力、临床效力或审核状态的混淆。

6.4 证据锚定与不确定性表达

6.4.1 基本要求

宠物影像智能辅助诊断软件的输出应与其依据相对应，并应能够反映结果形成所依托的影像信息、非影像临床信息及其他声明允许使用的信息来源。

输出内容不应仅给出脱离依据的结论性表述。涉及征象描述、异常提示、鉴别诊断提示、进一步检查建议或报告草案生成的，均应满足可核查、可区分、可复核的要求。

6.4.2 影像依据锚定要求

基于影像形成的输出，应与相应影像依据建立对应关系。至少宜满足下列要求：

- a) 关键征象描述应能够对应至相关影像、切面、序列、帧或区域；
- b) 异常检测、分割、定位和测量结果应能够与原始影像关联展示；
- c) 征象归纳或提示性结论宜能够对应至支撑该结果的主要影像表现；
- d) 多张影像、多个序列或多部位联合分析时，宜能够区分各项结论的主要依据来源。

影像依据锚定不应因结果生成方式不同而缺失。对无法建立明确影像对应关系的输出，不应表述为确定性影像结论。

6.4.3 非影像信息依据表达要求

基于病史、体征、实验室检查结果、既往检查结果或其他非影像临床信息形成的输出，应明确其信息来源类别，并与影像直接观察结果相区分。

涉及非影像信息参与推理的，宜满足下列要求：

- a) 说明所使用的非影像信息类型；

- b) 区分影像直接观察结果与综合分析结果；
- c) 存在信息缺失、冲突或来源不明时，应提示其对结果可靠性的影响；
- d) 非影像信息不应在缺乏相应影像支撑时单独形成确定性影像判断。

6.4.4 证据不足情形处置要求

当存在下列情形之一时，软件应降低输出确定性，或限制相关输出：

- a) 关键影像依据缺失；
- b) 影像质量不足以支持相应判断；
- c) 必要的非影像临床信息缺失、冲突或不可验证；
- d) 输入虽可处理，但现有依据不足以支撑相应推理层或文书层输出；
- e) 现有依据不足以支持明确判断。

在证据不足情形下，软件宜采取下列一种或多种处置方式：

- a) 提示“依据不足”“需结合临床进一步判断”或类似信息；
- b) 仅输出感知层结果，不输出推理层或文书层结论；
- c) 将输出限定为提示性、可能性或待核实性表述；
- d) 拒绝生成确定性判断。

6.4.5 不确定性表达要求

软件对存在不确定性的输出，应采用与风险相适应的方式进行表达，不应以确定性、排他性或误导性表述替代真实不确定性状态。

不确定性表达宜考虑下列因素：

- a) 输入质量及完整性；
- b) 影像表现是否典型；
- c) 不同依据之间是否一致；
- d) 任务是否处于声明用途和适用范围内；
- e) 结果是否依赖附加信息或多步骤推理。

对高风险场景、高不确定性场景或边界外场景，软件应加强不确定性提示，必要时不输出确定性结论。

6.4.6 输出表述要求

输出表述应与其证据强度和审核状态相一致，并宜符合下列要求：

- a) 客观影像描述与分析性推断应分开表述；
- b) 提示性意见与确定性结论应分开表述；
- c) 机器生成草案与经人工确认内容应分开表述；
- d) 不能由现有依据直接支持的内容，不应以既成事实方式表述；
- e) 存在多种可能解释的，宜避免单一路径的绝对化表达。

6.4.7 文书层输出特别要求

结构化报告草案、影像所见草案、初步结论草案等文书层输出，应明确标识其机器生成属性，并应满足下列要求：

- a) 应供用户审核、修改和确认；
- b) 应能够与正式报告状态区分；
- c) 关键结论宜保留相应依据线索或核查入口；
- d) 对依据不足、边界外或高不确定性内容，不应生成与正式确诊等同的表述。

6.5 人工审核、结果确认与记录要求

6.5.1 基本要求

宠物影像智能辅助诊断软件的使用应坚持人工审核原则。软件输出应作为辅助判读、辅助分析和辅助报告的参考信息，不应替代执业兽医师的专业判断，不应脱离人工审核直接作为最终诊断、正式报告或其他具有独立临床后果的依据。

涉及推理层输出和文书层输出的，均应纳入人工审核流程。涉及高风险场景、高不确定性场景或边界外场景的，宜加强人工复核要求。

6.5.2 审核责任要求

软件应用过程中应明确人工审核责任主体和审核环节。审核责任宜与用户角色、功能权限和使用场景相对应，并满足下列要求：

- a) 执业兽医师应对最终采信结果负责；
- b) 在执业兽医师监督下开展操作的专业技术人员，其操作范围应与授权范围相一致；
- c) 不同岗位用户的功能权限、修改权限和确认权限宜予区分；

d) 未经具备相应职责的人员审核确认，软件输出不应进入正式诊疗文书或正式报告流程。

6.5.3 人工审核要求

人工审核应结合原始影像、相关临床信息和软件输出结果开展，并宜包括下列内容：

- a) 核查输入信息是否完整、适用和符合声明范围；
- b) 核查感知层输出与原始影像是否一致；
- c) 核查推理层输出是否与影像依据及临床背景相符；
- d) 核查文书层输出是否存在遗漏、误述、夸大或不当确定性表述；
- e) 核查边界外、依据不足或高不确定性内容是否已作出适当提示或限制。

对于发现存在明显错误、依据不足、超出适用范围或存在误导风险的输出，应进行修订、降级使用或不予采信。

6.5.4 修订要求

软件输出经人工审核后，宜允许进行修改、补充、删除或重写，并应满足下列要求：

- a) 感知层输出的修订宜能够对应至具体影像依据；
- b) 推理层输出的修订宜结合影像表现和临床背景进行；
- c) 文书层输出的修订不应掩盖机器生成属性和审核过程；
- d) 对重大修订内容，宜能够反映其修订后状态；
- e) 对不予采信的输出，宜能够明确标识或终止后续引用。

6.5.5 结果确认要求

经人工审核和必要修订后的结果，方可进入正式使用环节。结果确认宜符合下列要求：

- a) 确认前应完成必要审核；
- b) 确认内容应与最终采信结果一致；
- c) 确认后的正式报告、正式记录或其他正式输出，应与未审核草案状态相区分；
- d) 未完成确认的草案内容，不应视为正式结论。

6.5.6 记录要求

软件应用过程宜保留满足追溯和复核需要的记录。记录内容至少宜包括：

- a) 原始输入信息；
- b) 软件原始输出结果；

- c) 人工修订内容;
- d) 最终确认结果;
- e) 所使用的软件版本, 以及相关模型版本、知识库版本、提示模板版本或规则版本;
- f) 必要的时间信息和操作信息。

记录应能够支持结果追溯、质量复核、问题分析和版本回顾。

6.5.7 记录管理要求

记录管理应与软件风险水平、使用场景和质量管理要求相适应, 并宜满足下列要求:

- a) 能够区分原始输出、修订过程和最终结果;
- b) 能够识别关键版本变更对结果的影响;
- c) 能够支持错误案例复盘和问题定位;
- d) 能够为软件更新后的回归评价提供基础;
- e) 涉及正式诊疗记录的, 应与相关管理要求保持一致。

6.5.8 特殊情形要求

对下列情形, 宜提出更严格的人工审核和记录要求:

- a) 高风险病例;
- b) 边界外或疑似边界外输入;
- c) 依据不足或高不确定性输出;
- d) 多模态综合分析输出;
- e) 经版本更新后首次投入使用阶段;
- f) 其他可能显著影响结果可靠性的情形。

对上述情形, 必要时宜限制自动生成结论性内容, 或仅保留提示性、待核实性输出。

7 系统实现与运行控制要求

7.1 基本要求

7.1.1 总体要求

软件应根据声明用途、任务边界、输出要求和人工审核要求进行设计与实现, 并形成与使用场景相适应的运行控制机制。

软件的功能设置、处理流程、结果呈现、审核支持和记录留痕, 应能够支持其按声明方式使用。

对于能够通过系统机制实现的控制要求，宜通过技术控制、流程控制或二者结合的方式予以落实。

7.1.2 一致性要求

软件实现、功能设置、界面展示和运行流程应与声明内容一致，不应突破已声明的使用边界。

软件不应通过默认设置、组合调用、附加功能或其他实现方式，形成超出声明用途、任务边界或输出范围的实际使用效果。

软件的实现方式不应因界面设计、流程安排或状态标识不清，导致用户对输出性质、审核状态或使用条件产生误解。

7.2 输入与功能调用控制

7.2.1 输入控制要求

软件应对声明支持的输入类型、输入条件和技术处理要求进行控制，并明确不同输入与相应功能的对应关系。

未声明支持的输入类型，不应作为生成相应分析结果的依据。

当软件支持多种输入类型、多种任务模式或多模态输入时，宜明确各类输入与相应功能、输出类型之间的对应关系。

7.2.2 功能调用控制要求

对软件能够判断的超出声明用途、超出适用范围、不适用或无法处理情形，应按声明要求采取提示、限制、降级或拒绝等措施。

当输入信息不完整、不规范或不满足声明输入条件，但系统仍可完成部分处理时，软件可根据风险采取限制生成确定性结果、提示局限性或要求人工复核等措施。

对不同用户角色、不同使用场景或不同任务模式，软件宜提供与其职责边界相适应的功能调用控制。

7.2.3 边界说明

对于实际上属于声明边界外、但软件未声明具备判断能力的情形，不要求其自动识别和拦截。

该类情形相关控制主要由使用者依据软件声明用途、适用范围和具体临床场景进行判断。

7.3 输出与展示控制

7.3.1 输出控制要求

软件输出应与声明用途、任务类型和输出范围相对应，并与其依据、确定性和审核状态保持一致。

不属于声明输出范围的内容，不应以相应结果形式输出。

对依据不足、超出适用范围或需人工复核的情形，软件应按声明要求对输出进行限制，或作出相应提示。

对于软件能够识别且声明要求控制的边界外情形、依据不足情形或其他需限制输出的情形，软件宜根据声明要求和风险控制需要，对输出层级、输出内容或输出方式进行限制；必要时可仅输出感知层结果，或不输出推理层、文书层结果。

7.3.2 展示控制要求

软件应通过界面、标识或流程区分不同性质和不同状态的输出，避免将客观描述、分析性提示、机器生成草案和正式结果相混淆。

输出展示不应因呈现方式、文案表述或流程设置，造成其审核状态、采信条件或使用性质的误解。

7.3.3 文书输出要求

文书层输出应明确标识其机器生成属性，并与正式报告、正式病历记录或其他正式文书状态相区分。未经人工审核、修改和确认的文书层输出，不应直接作为正式使用内容。

7.4 人工审核与结果确认支持

7.4.1 审核支持要求

对声明需经人工审核、人工确认或人工复核的输出，软件应提供相应流程支持。

软件不应使相关结果在未经审核的情况下直接进入正式使用环节。

7.4.2 权限与修订要求

软件宜根据用户角色设置查看、修改、确认等权限，并支持对原始输出进行修改、补充、删除或不予采信处理。

不同用户角色的操作范围和确认权限宜与其职责边界相适应。

7.4.3 结果确认要求

未经人工审核确认的结果，不应进入正式使用环节。

经确认后的结果，应与未审核草案状态相区分。

7.5 记录与版本留痕

7.5.1 记录要求

软件应具备记录原始输入、原始输出、人工修订、最终确认结果及相关操作信息的能力。

相关记录应能够反映结果形成、修改和确认的主要过程。

7.5.2 版本留痕要求

软件应能够记录与结果形成相关的软件版本，以及相关模型版本、知识库版本、提示模板版本或规则版本等信息。

涉及多个版本共同影响输出结果的，宜能够反映其对应关系。

7.5.3 追溯要求

相关记录应能够支持结果追溯、质量复核、问题分析和版本回顾。

对原始输出、修订过程和最终结果，宜予以区分，避免不同状态内容相混淆。

7.6 变更控制

7.6.1 变更控制要求

涉及模型、规则、知识库、提示模板、输出方式或审核流程等重要变更时，应进行相应控制。

相关变更不应在未完成必要评估和确认的情况下直接投入正式使用。

7.6.2 一致性要求

变更后软件仍应与声明用途、任务边界、输出要求和人工审核要求保持一致。

变更不应导致软件实际功能、输出形式或使用方式突破原有声明范围；确需调整声明内容的，应按相关要求同步更新。

8 验证与测试要求

8.1 总则

8.1.1 验证目的

本章规定了宠物影像智能辅助诊断软件的验证与测试要求，用于评价软件是否符合声明用途和相关控制要求，并评价模型输出是否具有实际辅助应用价值。

8.1.2 验证依据

验证与测试应依据软件声明的适用范围、任务类型、输入条件、输出类型和使用限制开展。

对未声明支持的能力，不作为常规验证内容。

8.1.3 验证方法

验证与测试宜采用文档核查、场景测试、专家评审和记录抽查等简单、可实施的方法。

能够合并验证的内容，宜在同一场景中一并完成。

8.1.4 结果记录

验证与测试结果应形成记录，并支持复核、追溯和更新后再验证。

8.2 测试资料与参考答案要求

8.2.1 测试资料

测试资料应与软件声明用途一致，并覆盖主要使用场景。

8.2.2 场景类型

测试资料宜包括下列场景：

- a) 声明范围内的典型应用场景；
- b) 软件能够判断的边界外、不适用或无法处理场景；
- c) 高风险或高不确定性场景。

8.2.3 参考答案

模型有效性验证应有参考答案。

参考答案可采用专家一致性判读结果、经审核确认的标准报告，或结合临床信息形成的参考结论。

8.2.4 参考答案形成

难以获得单一标准答案时，可由多名专家共同评议形成参考答案。

8.2.5 资料与答案对应

测试资料、参考答案和验证结果之间应能对应，并便于复核。

8.3 声明符合性与边界控制验证

8.3.1 声明内容核查

应核查软件声明的适用范围、任务类型、输出类型和使用限制是否明确。

8.3.2 声明一致性核查

应核查软件说明、界面和实际功能是否与声明内容一致。

8.3.3 边界控制验证

对软件能够判断的边界外、不适用或无法处理情形，应验证其是否能够提示、限制、降级或拒绝。

8.3.4 输出控制验证

对边界外、不适用或无法处理情形，应验证软件是否避免输出与声明用途不符的确定性结果。

8.3.5 验证方式

本节宜采用文档核查和场景测试相结合的方式开展。

8.4 模型有效性验证

8.4.1 验证目的

应验证模型输出是否具有实际辅助应用价值。

8.4.2 验证方式

模型有效性验证宜采用专家对照评审方式开展。

8.4.3 评审内容

应评审模型输出与参考答案的一致程度，以及是否存在明显错误、明显误导或关键遗漏，并评价其是否具有可采纳性。

8.4.4 结果表示

验证结果宜采用专家认可率、基本一致率、可采纳率、明显错误率和关键遗漏率表示。

8.4.5 验证记录

模型有效性验证过程和结果应形成记录，并便于复核。

8.5 场景化综合验证

8.5.1 验证目的

应结合实际使用场景，对软件输入、模型输出、结果控制、人工审核和记录留痕等进行综合验证。

场景化综合验证应以软件声明用途和声明边界为依据开展。

8.5.2 典型场景验证

应对声明范围内的典型应用场景开展验证，确认软件能够按声明用途完成处理并输出结果。

8.5.3 边界场景验证

对软件声明能够判断的边界外、不适用或无法处理场景，应开展验证，确认其相关控制措施有效。

对软件未声明具备该类判断能力的，不要求验证其相应控制能力。

8.5.4 验证结果

场景化综合验证应能反映软件在实际使用中的整体运行情况，并形成相应记录。

8.6 人工审核与结果控制验证

8.6.1 验证目的

应对软件声明的人机配合方式、人工审核要求和结果控制要求进行验证。

8.6.2 人工审核要求验证

软件声明设置人工审核、人工确认或人工复核环节的，应验证相关提示、流程和操作要求是否明确，并能够实际执行。

软件未声明设置该类环节的，不将其作为常规验证内容。

8.6.3 结果控制验证

应验证结果展示、提示信息、使用限制和相关控制措施是否与软件声明一致。

8.6.4 审核与结果衔接验证

对声明需经人工审核后方可使用的结果，应验证审核前后结果状态、使用条件和记录要求是否明确。

8.6.5 验证记录

人工审核与结果控制验证过程和结果应形成记录，并便于复核和追溯。

8.7 验证结果记录与结论

8.7.1 记录要求

验证过程、验证结果和相关结论应形成记录，并便于复核、追溯和必要时更新。

8.7.2 记录内容

验证记录宜包括验证项目、验证资料、验证方法、验证过程、验证结果、发现问题及其处置情况。

8.7.3 结果分析

应结合各项验证结果，对软件是否符合声明用途、声明边界和相关控制要求进行综合分析。

8.7.4 验证结论

应根据验证结果形成明确结论。

验证结论宜包括验证通过、经整改后可接受，或不符合相关要求等情形。

8.7.5 结论使用

验证结论应作为软件上线、变更控制、持续改进和后续评价的依据之一。

参考文献

- [1] 国家互联网信息办公室令第15号《生成式人工智能服务管理暂行办法》，2023年7月10日公布，2023年8月15日起施行
- [2] 中国电子技术标准化研究院.《人工智能标准化白皮书（2018版）》
- [3] 《兽用处方药和非处方药管理办法》（农业部令2013年第2号）
- [4] 《中华人民共和国动物防疫法》（2021年修订）
- [5] 中华人民共和国农业农村部令2022年第5号《动物诊疗机构管理办法》
- [6] 中华人民共和国农业农村部令2022年第6号《执业兽医和乡村兽医管理办法》
- [7] 中华人民共和国农业农村部公告第734号附件1《动物诊疗病历管理规范》，2023年12月12日发布，2024年5月1日起施行
- [8] 中华人民共和国农业农村部公告第734号附件2《兽医处方格式及应用规范》，2023年12月12日修订发布，2024年5月1日起施行
- [9] DICOM Standard (Digital Imaging and Communications in Medicine), relevant parts of the current edition
-