

# 多模态模型精准医学图像数据标注规范

Specification for Multi modal model precision medical image data annotation

(征求意见稿)

XXXX - XX - XX 发布

XXXX - XX - XX 实施

# 目 次

前 言 .....	1
引 言 .....	2
1 范围 .....	3
2 规范性引用文件 .....	3
3 术语和定义 .....	3
4 缩略语 .....	4
5 基本原则 .....	4
6 标注流程 .....	5
7 数据基础要求 .....	6
8 标注技术要求 .....	7
9 质量控制与验证 .....	9
10 标注平台与工具要求 .....	11
11 安全、隐私与伦理要求 .....	11
附录 A（资料性）多模态数据标注任务描述要素说明及示例 .....	13
附录 B（资料性）常用数据结构字段说明及示例 .....	17
附录 C（资料性）多模态标注标签体系示例 .....	17
附录 D（资料性）标注质量评价指标与计算方法 .....	17
参 考 文 献 .....	23

# 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。

本文件由辽宁省数据局提出并归口。

本文件起草单位：沈阳大学、沈阳市妇婴医院、中国医科大学附属盛京医院、中国医科大学附属第一医院、中国医科大学、东北大学、沈阳市市场监管事务服务中心（沈阳标准化研究院）、沈阳工业大学、沈阳速影科技有限公司。

本文件主要起草人：邵一川、张志涛、孙洁、曹勇、尹红艳、刘柳、高扬、张丹、勾颖、王磊、崔笑宇、马腾、齐瑞群、孙海静、张乐、刘晓、孙妍、林艳、高媛、顾敏、李新星、郑继慧、周新佳、李勇男、赵蹇、冯时、杨晓东。

本文件发布实施后，任何单位和个人如有问题和意见建议，均可以通过来电和来函等方式进行反馈，我们将及时答复并认真处理，根据实际情况依法进行评估及复审。

归口管理部门通讯地址：辽宁省数据局（辽宁省沈阳市皇姑区崇山中路109号），联系电话：024-83988807。

标准起草单位通讯地址：辽宁省沈阳市沈阳大学（辽宁省沈阳市大东区望花南街21号），联系电话：024-62268721。

# 引 言

近年来，深度学习与大模型等人工智能技术在医学影像智能分析中的应用不断深化，多模态模型对医学影像数据及其相关文字描述数据的高质量标注提出了更高要求。标注数据的规范性、准确性与可追溯性，直接影响模型训练效果以及在临床应用中的安全性、可解释性与可重复性。为满足多模态模型在模型训练、科研分析、系统验证等场景的需求，有必要对多模态模型用精准医学图像数据标注活动进行统一规范。

本文件的制定旨在：

- 统一多模态模型医学图像数据标注的任务要素、标注粒度与跨模态对齐要求；
- 规范标注过程与质量控制方法，明确定量与定性评价指标；
- 明确标注组织管理、人员角色、工具平台及版本追溯等基本要求；
- 为辽宁省多模态医学人工智能应用提供可复现、可推广的标注标准化依据。

本文件可用于指导多模态模型医学图像数据标注项目的设计、实施、验收与监督管理，并可供医疗机构、科研机构、医学影像中心、数据服务机构、医学影像数据平台及相关企业在数据集建设与标注体系完善中参考。实施本文件有助于提升医学人工智能数据资源的规范化与共享水平，促进高质量数据集形成，支撑人工智能技术在医疗健康领域的安全、可靠应用与持续发展。

# 多模态模型精准医学图像数据标注规范

## 1 范围

本文件规定了多模态模型精准医学图像数据标注的基本原则、标注流程、技术要求、质量控制与验证等要求，明确了多模态关联标注与跨模态一致性评价的技术规则。

本文件适用于医疗机构、科研院所、人工智能企业等相关单位开展CT、MR、PET、PET-CT、超声、内窥镜图像、数字病理图像（WSI）等多模态医学影像的数据标注工作，可作为多模态医学 AI 模型训练、多模态大模型构建及临床智能辅助诊断系统开发的依据。

本文件不适用于医学影像设备质量控制参数的测量、医学诊断结论的出具以及非医学用途的图像标注。

## 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而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 42755 人工智能 面向机器学习的数据标注规程

YY/T 1833.1 人工智能医疗器械质量要求和评价 第1部分:术语

YY/T 1833.2 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

YY/T 1833.3 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

## 3 术语和定义

GB/T 42755、YY/T 1833.1、YY/T 1833.2、YY/T 1833.3界定的以及下列术语和定义适用于本文件。

### 3.1

**多模态模型** multi-modal model

以两种及以上模态数据为输入，进行联合表征、对齐或推理的模型。

注：本文件中的模态数据包括医学影像数据（如 CT、MR、PET、超声、数字病理等）及与其相关的文字描述数据（如检查报告、病理摘要、结构化临床字段、随访记录等）。

### 3.2

**精准医学图像数据标注** precision medical imaging data annotation

为满足多模态模型训练、验证、评价或应用需求，按照标注任务说明，对医学影像数据及其相关文字描述数据进行标记、结构化和关联，并经审核形成可交付标注结果的活动。

### 3.3

**跨模态对齐** multi-modal alignment

在同一病例（或同一检查）范围内，建立不同模态数据要素之间对应关系的过程或结果。

注：本文件中的跨模态对齐主要指影像标注对象（如 ROI/病灶/结构）与文字描述数据中的字段、片段或结构化条目之间的映射。

### 3.4

#### 模型辅助标注 model-assisted labeling

利用多模态模型或单模态模型对待标注数据生成建议标签、候选标注或一致性提示,并由人工审核、修正与确认后形成最终标注结果的标注方式。

## 4 缩略语

下列缩略语适用于本文件。

MR: 磁共振成像 (Magnetic Resonance Imaging)

PET: 正电子发射断层成像 (Positron Emission Tomography)

US: 超声 (Ultrasound)

WSI: 全景病理切片 (Whole Slide Image)

ROI: 感兴趣区域 (Region of Interest)

CT: 计算机断层成像 (Computed Tomography)

PACS: 医学影像归档与通信系统 (Picture Archiving and Communication System)

DICOM: 医学数字成像和通信 (Digital Imaging and Communications in Medicine)

NIfTI: 医学影像技术倡议格式 (Neuroimaging Informatics Technology Initiative)

JSON: 数据交换格式 (JavaScript Object Notation)

IoU: 相交并比 (Intersection over Union)

Dice: Dice相似系数 (Dice Similarity Coefficient)

HD: 豪斯多夫距离 (Hausdorff Distance)

ICC: 组内相关系数 (Intraclass Correlation Coefficient)

## 5 基本原则

### 5.1 合规性原则

应符合国家及地方相关法律法规、医疗行业规范及数据安全标准,严格保护患者隐私,确保数据采集、使用、存储等环节合法合规。

### 5.2 精准性原则

标注内容应严格遵循临床诊断标准、医学影像诊断规范及本文件要求,准确反映图像中的客观信息,误差控制在允许范围内。

### 5.3 一致性原则

同一标注任务的不同标注人员应遵循统一的标注规则,确保标注结果的一致性。

### 5.4 完整性原则

应全面覆盖多模态图像中与模型训练目标、临床应用需求相关的所有关键信息,对于多模态融合图像,应完整标注各模态间的关联信息。

### 5.5 安全性原则

严格遵守相关法律法规要求,对标注过程中的个人健康信息和敏感医学数据进行全程脱敏和安全管控,防止数据泄露。

### 5.6 可追溯性原则

应完整记录标注过程中的关键信息(如标注人员、标注时间、审核意见、修改记录等),确保标注结果可追溯、可核查。

## 6 标注流程

### 6.1 总体流程

数据标注总体流程符合图 1 所示,包括数据准备、任务定义、标注执行、质量控制、结果输出 5 个核心环节。

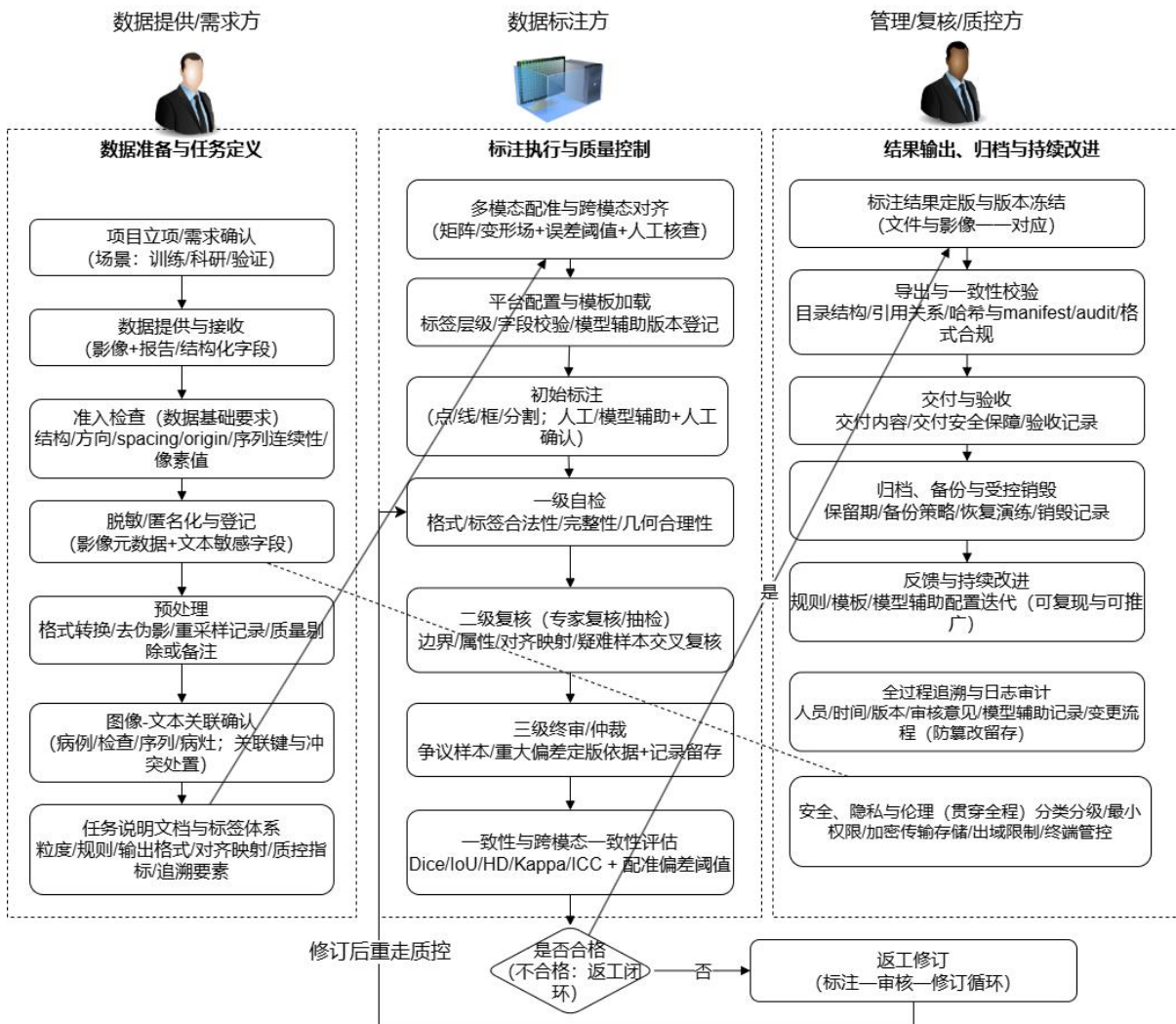


图 1 多模态医学图像数据标注流程图

### 6.2 数据准备

数据提供方应提供合法合规的多模态医学影像数据及关联文字描述数据,数据标注方应联合数据提

供方完成数据预处理，包括格式转换、去伪影、脱敏等，确保数据符合本文件第 7 章要求。

### 6.3 任务定义

6.3.1 数据标注方应根据应用场景（如结构标注、病灶标注、多模态关联标注、结构化报告标注等），明确标注范围、标签体系、标注粒度、空间要求、操作规则及预期输出格式；多模态数据还应明确跨模态配准方式、坐标体系选择及不同模态间的对应关系。

6.3.2 数据标注方应组织数据提供方共同完成医学影像数据与文字描述数据的关联确认，明确关联粒度（病例/检查/序列/病灶）与关联键规则（如病例号、影像 UID、检查时间等）；对缺失关联信息、关联信息冲突、时间不一致等情况，应制定处置规则并留存记录。

6.3.3 任务定义应以书面文档形式记录，应至少包含任务目标、数据模态、标注对象与输出、对齐映射规则、质控与追溯要素等内容，任务描述的要素与写法要求见附录 A。

### 6.4 标注执行

6.4.1 标注人员应依据任务定义和标签体系开展标注，标注方式包括点标注、线标注、框标注、多边形标注、像素级或体素级分割，具体方式根据标注精度需求确定；多模态影像应根据配准矩阵切换模态。

6.4.2 标注过程中，标注平台应自动记录标注操作等信息；采用模型辅助标注时，模型输出应仅作为参考，须完成人工校正并记录相关信息。

### 6.5 质量控制

6.5.1 审核分自动质控与人工审核：自动质控应检查标注几何合理性（如 mask 越界、切片错位）、标签合法性（符合标签体系）、格式规范性及跨模态一致性（配准误差）。人工审核应由具有医学背景的专家执行，重点检查病灶边界、器官轮廓、属性标注的正确性，排查漏标、误标或不一致情况。

6.5.2 不合格数据应退回标注员重新修改，形成“标注—审核—修订”循环，审核结果需书面记录并纳入标注档案。

### 6.6 结果输出

6.6.1 输出文件宜采用独立于影像文件的结构化标注文件形式（JSON、NIFTI 等），应包括标注数据文件、标注说明文档、质量控制记录等内容。

6.6.2 标注文件输出后，应与影像数据进行一致性校验，包括维度检查、坐标检查、方向矩阵匹配及版本关联验证。

## 7 数据基础要求

### 7.1 数据来源要求

多模态医学影像数据及相关文字描述数据应来源于合法合规的医疗机构或科研机构，数据采集过程应符合医学伦理要求，并遵守相关法律法规及数据隐私保护规定。

### 7.2 数据结构要求

输入影像数据应满足以下属性要求：

- a) 结构完整：影像数据应包含完整的 DICOM 或 NIFTI 文件结构，结构格式见附录 B；
- b) 方向矩阵(Direction)明确：影像方位必须以矩阵形式明确记录，方向矩阵必须能够唯一确定图像坐标系方向；

- c) 体素间距(Spacing)精确：提供三维体素间距，且单位明确。重采样操作应记录原始与重采样 Spacing；
- d) 图像原点(origin)记录完整：记录影像在物理空间中的原点坐标；
- e) 序列连续性：所有切片应按正确的空间位置排序，不得出现丢片、重复片、顺序错误或方向翻转；
- f) 像素值有效性：影像像素值应符合临床成像标准，如 CT 值单位(H)必须正确，MRI 图像应无严重偏场伪影或信号中断；
- g) 图像质量可接受：影像应具有足够分辨率、对比度及信噪比，以满足标注精度要求，严重伪影、运动模糊等应予以剔除或备注；
- h) 跨模态配准信息完整：多模态影像应附带配准矩阵或变形场及配准误差评价指标；如数据不满足上述要求，应在标注前进行预处理、校准或记录相应不符合项。

注：CT/MR/PET 等放射影像文件宜采用 DICOM；WSI 宜采用 OpenSlide 兼容格式或项目约定格式并提供金字塔与倍率元数据；内窥镜/超声视频或多帧数据宜采用项目约定的可解析格式，并提供帧索引与时间戳规则。

### 7.3 数据脱敏要求

文本数据脱敏应处理姓名、证件号、联系方式、住址等敏感字段，保留脱敏策略说明与抽检记录，如需保留脱敏后标识与原始标识的对应关系，应明确存储与访问规则并审计。影像元数据脱敏删除或匿名化患者身份信息，应不影响诊断和空间定位信息。

## 8 标注技术要求

### 8.1 标注类型分类

多模态医学影像标注根据任务目标和应用场景可分为五大类型，具体如下：

- a) 解剖结构标注：识别并标出人体器官、系统及其子结构，构建三维医学影像结构化基础，用于体积分析、手术规划及多模态结构对齐任务；
- b) 病灶标注：对异常区域进行像素级或体素级标注，用于表示不同病理类型的空间形态、范围和属性，是模型训练关键数据；
- c) 多模态融合标注：跨 CT、MRI、PET、超声等多模态影像之间建立结构对应关系，应用于多模态 AI 模型、多模态诊断系统、大模型训练；
- d) 结构化报告标注：从影像诊断报告或临床文本提取结构化标签，构建“影像—文本”对应关系的跨模态管理体系。用于弱监督训练、NLP 结构化抽取、影像报告自动生成等任务；
- e) 文本—图像跨模态标注：建立图像区域与文本描述的对对应关系，是大模型训练必要数据形式。

### 8.2 标注元素与标签体系

#### 8.2.1 一般要求

标注体系应基于医学专业知识建立统一、可扩展、可计算的标签系统，包括标签编码、标签命名、语义描述、结构属性及跨模态对照关系。

#### 8.2.2 核心元素

应遵循“标签统一、扩展标签可添加”的原则，至少包含以下核心元素：

- a) 标签编码：每个解剖结构与病灶标签必须具有唯一编码，可采用英文化命名（如 liver, lung\_upper\_lobe, brainstem, tumor\_hcc），并可根据层级形式构建树状结构；

- b) 语义定义：具有清晰的医学定义，如结构名称、组织学特征、功能意义、临床指征等，以确保标注人员理解一致；
- c) 结构属性：如器官类型（实质性器官/空腔器官）、解剖层级（系统 /器官/子结构）、边缘特征（清晰/模糊）等；
- d) 跨模态对应：记录跨模态标注的数据不同模态的特征表达，例如“MRI T2 高信号对应 CT 低密度区”；
- e) 扩展字段：如病灶强化特征、病理类型、分级（BI-RADS、NI-RADS）、时间序列特征（进展稳定）等。

注：附录 C 给出了多模态标注标签体系示例。

### 8.3 标注粒度与层级结构

标注体系应采用分层级设计，以适应不同标注任务的复杂度。层级结构一般分为三级或四级，可根据器官、病灶类型进行扩展：

- a) 一级（系统级）：如消化系统、神经系统、呼吸系统；
- b) 二级（器官级）：如肝脏、胰腺、脑、心脏、肺；
- c) 三级（子结构）：如肝段（S1-S8）、肺叶（左上叶/右中叶）、脑区；
- d) 四级（微结构或特征级）：如病灶坏死区、强化边缘区、软化区、动脉期强化区域。

此外，病灶标注粒度可分为：图像级标签（weak label）、区域级（bbox/ROI）、边界级（polygon）、体素级分割（mask）。模型训练与临床场景一般要求体素级分割，需严格控制粒度一致性。

### 8.4 标注一致性要求

标注一致性要求如下：

- a) 空间一致性：标注在同一坐标体系，方向矩阵（direction）与体素间距（spacing）正确；
- b) 语义一致性：同一结构或病灶在不同图像、不同模态和不同病人中使用相同标签体系、相同编码和相同命名；
- c) 边界一致性：不同标注人员对同一器官或病灶的边界判断应一致，差异超过阈值（如 Dice<0.85）时需二次审核；
- d) 跨模态一致性：不同模态标签关系逻辑统一；
- e) 版本一致性：标注修改通过统一版本管理系统，保持标注文件与影像文件一一对应。

注：一致性要求可通过自动化工具与人工审核结合方式执行，并形成质控记录用于追溯。

### 8.5 多模态数据融合标注要求

#### 8.5.1 图像—图像配准要求

8.5.1.1 开展多模态标注前完成空间配准，优先采用刚体或仿射配准，存在明显组织弹性形变时可采用非线性配准（如 B-spline、Demon、SyN 等）。

8.5.1.2 配准过程应生成配准矩阵或变形场，并以可复现格式随标注数据一并保存；对 PET-CT 等配准影像，需检查是否存在旋转、平移偏移或设备异常造成的细微错位。

8.5.1.3 标注人员应在配准后的影像上进行标注，不得在未对齐的原始模态上分别标注同一结构；跨模态一致性检查应作为质控步骤，在不同模态中同步预览标注区域，偏差超过阈值（3-5 mm）时须重新调整。

8.5.1.4 最终输出应包含融合后的空间坐标体系、配准误差、配准算法版本及跨模态一致性报告。

#### 8.5.2 图像—文本对应关系标注要求

8.5.2.1 图像—文本对应标注应建立医学影像与诊断描述、病理特征、影像征象、临床语义之间的对应关系，标注内容包括影像中的解剖区域、病灶表现、信号特征、强化特征、影像征象与文本描述片段的映射关系（如报告中“右肺上叶见类圆形软组织密度结节，边界欠清”与图像中对应 ROI 或 segmentation mask 的链接）。

8.5.2.2 文本片段与影像区域的绑定应采用唯一 ID 编码，并记录语义类型（如征象、诊断、属性描述、建议）；标注系统应支持文本定位（text span selection）、句子级与片段级的对应关系记录。

8.5.2.3 跨模态标注应遵循语义一致性原则，图文描述冲突或错配时（如文本描述“强化明显”而影像对应区域无强化特征）需备注“不一致”，纳入质控流程。

8.5.2.4 最终生成的结构化标注文件应包括图像区域坐标、文本片段引用、语义类别标签及跨模态一致性状态。

### 8.5.3 图像—结构化信息融合要求

8.5.3.1 图像—结构化信息融合标注应将医学影像与临床结构化数据（如电子病历 EMR、检验检查数据、病理指标、患者基本特征）绑定（如乳腺 MRI 肿瘤区域与 BI-RADS 分类、HER2 状态、Ki-67 指标的融合，肝脏 CT 病灶与 AFP、ALT、肿瘤分期的关联）。

8.5.3.2 所有结构化临床信息应通过唯一键（如病例号、影像 UID、检查时间）与影像标注文件绑定；结构化信息需经过脱敏处理，明确字段定义（含数值类型、时间属性、单位、正常范围等）。

8.5.3.3 融合标注文件中应以键值对方式记录所有关联字段。

8.5.3.4 融合标注应保证时间一致性，影像采集时间与检验时间间隔超过 90 天时，需记录“时间偏差标签”，便于模型训练时过滤处理；最终融合文件应确保数据结构一致、字段定义规范、语义表达统一。

## 9 质量控制与验证

### 9.1 质量检查指标

#### 9.1.1 定量指标

定量指标宜根据任务类型选择，至少包括：

- a) 检测/定位任务：精确度、召回率等；
- b) 分割任务：Dice、IoU、Hausdorff 距离、Conformity 等；
- c) 分类/判别任务：灵敏度、特异度、准确率、ROC/AUC 等；
- d) 一致性评价：Kappa 等；
- e) 跨模态一致性：相关系数等。

评价指标将计算公示说明见附录 D。

注1：对于模型辅助标注，还必须评估AI预测与人工校正的偏差，包括AI初始mask的边界偏移、漏标率、过标率、区域错配等。

注2：具体指标及阈值应在标注任务说明或质控方案中明确，并与抽检比例、判定规则一并固化。

#### 9.1.2 定性指标

定性评价宜结合专家审查与规则检查开展，至少包括：

- a) 规范符合性：标签体系、命名规则、坐标系/空间参考、字段编码等是否符合本文件及任务要求；
- b) 完整性：标注对象、属性字段、跨模态映射关系是否缺失；
- c) 合理性：标注边界、属性取值与医学常识是否一致，是否存在明显逻辑冲突；

- d) 可追溯性：标注人员、审核意见、修改记录、版本信息、模型辅助标注记录等是否齐全；
- e) 可用性：标注结果是否满足模型训练、科研分析或系统验证等预期用途。

## 9.2 一致性评价方法

一致性评价用于衡量不同标注人员、不同版本和不同模态之间标注结果的一致性。本文件使用统计学指标与空间重叠指标结合的方法，主要包括：

- a) Dice：用于测量两个体素级掩码的重叠程度，是器官与病灶分割一致性评价的核心指标；
- b) IoU：用于检测类任务的bbox/ROI一致性评价；
- c) HD、HD95：用于衡量边界偏差，判断标注是否存在偏移或局部错误；
- d) Kappa系数：适用于分类类、属性类和文本类标注的一致性评价，可衡量语义标签之间的一致性；
- e) ICC：用于连续属性（肿瘤直径、体积、SUV值）的标注一致性分析。

一致性评价应在标注版本之间、标注员之间及模态之间计算，并记录在质控文件中。对于一致性较差的区域（如Dice<0.75），需进入专项审核流程，并根据错误类型重新标注或修正标签体系。

## 9.3 质量控制流程

9.3.1 标注项目应建立“一级自检—二级复核—三级终审（仲裁）”的三级质量控制机制，覆盖标注、复核与终审（仲裁）环节：

- a) 一级质量控制（自检）：自动质控由系统执行，首先检查图像与掩码的空间维度、方向矩阵是一致，检查掩码越界、值异常、空洞、断层跳变等问题；
- b) 二级质量控制（复核）：由具备相应资质的复核人员执行，应对标注准确性、一致性、跨模态对齐关系及关键字段进行复核，对疑难或高风险样本宜实施交叉复核；
- c) 三级质量控制（终审/仲裁）：由项目负责人或医学专家组执行，应对争议样本、抽检不合格样本及重大偏差进行终审确认，终审结论应作为定版依据，并保留过程记录以便追溯。

9.3.2 GB/T 42755 中 6.2 所规定的通用要求适合于本文件。

## 9.4 验证方法

### 9.4.1 内部验证

内部验证应采用以下方式：

- a) 交叉验证：对数据量较大的项目，将标注数据分为多个子集，轮流以一个子集为验证集、其他为训练集（或参考集），计算不同轮次的质量评价指标，评估标注结果的稳定性和可靠性；
- b) 重复标注验证：选择部分数据，由不同标注人员按相同规则重复标注，计算一致性指标（如Kappa系数、ICC），评估标注一致性和可重复性；
- c) 专家验证：邀请医学专家抽样验证，专家依据专业知识评价标注准确性、合理性和完整性，提出验证意见，作为质量评估的重要参考。

### 9.4.2 外部验证

外部验证应采用以下方式：

- a) 参考标准比对：若存在权威的参考标准（如经病理诊断证实的病灶标注、由多名资深专家共同确定的标注结果等），可将标注结果与参考标准进行比对，计算两者之间的一致性指标和误差指标，评估标注结果的准确性；
- b) 第三方验证：委托独立的第三方机构对标注结果进行验证，第三方机构应具备相应的资质和专

业能力，按照本文件及标注任务要求，采用科学、公正的方法对标注质量进行评价，并出具验证报告。第三方验证结果可作为标注项目验收和标注质量认定的重要依据；

- c) 模型验证：将标注数据用于多模态模型训练，通过模型在测试集上的性能表现（如准确率、召回率、AUC 值等）间接验证标注数据的质量。若模型性能达到预期目标，说明标注数据质量较高，能够满足模型训练需求；若模型性能不理想，应分析原因，排查是否存在标注质量问题。

## 10 标注平台与工具要求

### 10.1 功能要求

10.1.1 标注平台或工具应支持多模态数据的接入、管理与标注，至少包括医学影像数据及与其配对的文字描述数据（报告文本、病理摘要、结构化字段等）的关联展示与引用管理。

10.1.2 平台应支持本文件涉及的主要标注任务类型，至少包括：

- a) 检测/定位标注（如bbox、关键点等）；
- b) 分割标注（如polygon、mask，支持二维/三维）；
- c) 属性标注（字段字典、枚举值、单位与范围校验）；
- d) 跨模态对齐标注（字段与影像标注对象的绑定、映射关系维护）。

10.1.3 平台应支持标签体系与模板管理，至少包括标签层级结构、属性字段定义、版本管理与变更记录；应支持按项目加载标准化标注模板。

10.1.4 平台应具备质控与审核功能，至少包括任务分发、进度管理、复核/仲裁流转、抽检管理、返工闭环与差异对比。

10.1.5 平台应支持数据导出与一致性校验，至少包括按规定目录结构导出影像与标注文件、输出结构化文本字段、生成校验信息（如哈希值、文件大小、引用关系），并支持导出前的格式与引用完整性检查。

10.1.6 平台宜支持多模态模型辅助标注配置与记录，至少包括模型版本登记、推理参数配置、辅助标注结果加载、人工修订记录与可追溯信息留存。

10.1.7 平台应具备身份认证、访问控制与日志审计能力，其审计范围、防篡改与留存要求应符合11.5。

### 10.2 性能与可靠性

10.2.1 平台应满足多模态数据标注的性能需求，至少包括影像加载、浏览、缩放/窗宽窗位调整、切片/序列切换及三维分割交互的响应要求；性能指标宜在项目实施方案中明确。

10.2.2 平台应具备稳定运行能力，支持并发用户访问与任务协同；应采取措施避免数据丢失、任务状态异常或重复提交等问题。

10.2.3 平台应具备容错与恢复能力，至少包括异常告警、自动保存、断点续作、数据备份与恢复；发生故障时应可定位原因并恢复至一致状态。

10.2.4 平台应保证数据与结果的完整性，至少包括：上传/导出校验、版本冻结与只读保护、关键文件的校验与一致性检查。

## 11 安全、隐私与伦理要求

### 11.1 安全管控与分级控制

11.1.1 对数据分类分级，明确个人信息、敏感数据控制要求，不同安全级别数据采取差异化措施（网络隔离、访问等级、脱敏强度等）。

11.1.2 标注平台可按最小权限配置账户，角色分数据接收/预处理、标注、复核、质控、管理员等。

11.1.3 外包或多方协作需明确数据出域条件、二次分发限制与责任追究机制。

## 11.2 隐私保护与脱敏

11.2.1 数据进入标注环境前需脱敏，覆盖影像元数据与文本敏感信息，记录脱敏规则与工具版本，脱敏映射表与输出包隔离存储。

11.2.2 文本数据脱敏需处理姓名、证件号等敏感字段，保留策略说明与抽检记录；如需保留脱敏后与原始标识对应关系，需明确存储与访问规则并审计。

## 11.3 访问控制与日志审计

11.3.1 标注平台宜采用多因素认证，禁止账户共享，关键操作（登录、数据导入导出、标注修改等）需日志审计，日志防篡改并按规定留存。

11.3.2 标注终端需准入控制、病毒防护、屏幕录制/截屏管控、外设管控及补丁管理。

## 11.4 数据传输、存储与销毁

11.4.1 数据传输采用加密通道，跨域传输需审批与白名单，留存传输清单；存储需访问控制与加密，高敏感数据采用专用隔离存储。

11.4.2 数据备份需制定策略（频次、介质、保留期），定期恢复演练；项目结束或超保留期需受控销毁或归档，记录销毁对象、方式与执行人。

## 11.5 全过程追溯与版本管理

11.5.1 追溯信息覆盖数据来源、标注过程、模型辅助信息、质控与交付信息，与病例/序列/标注对象关联。

11.5.2 标注结果定版后冻结版本并只读保护，修订需走变更流程并保留旧版本，追溯信息通过输出包体现。

## 附录 A

### (资料性)

#### 多模态数据标注任务描述要素说明及示例

##### A.1 总则

本附录给出“任务定义文档”的一种写法示例，用于说明任务要素的组织方式；具体字段与阈值由项目任务定义确定。

##### A.2 任务概述

###### A.2.1 任务目标

面向多模态模型训练/验证，完成甲状腺结节在超声影像中的定位/勾画，并与同次检查报告的关键描述建立图像—文本对应关系。

###### A.2.2 数据模态

内容包含如下：

- a) 医学影像：甲状腺超声静态图像（必要时含关键帧序列）；
- b) 文字描述：对应检查报告文本及结构化字段（如部位、大小、征象描述、结论/分级字段等）。

###### A.2.3 标注方式

模型辅助（半自动）+ 人工确认/修订。

###### A.2.4 交付内容

内包含如下：

- a) 影像端：结节 ROI 轮廓/掩膜；
- b) 文本端：报告字段抽取或实体/片段标注（按任务约定）；
- c) 跨模态端：ROI 与报告字段/片段的对齐映射；
- d) 分级端：按既定规则体系形成分级标签（示例：ACR TI-RADS 1~5 类）。

###### A.2.5 结果组织

影像标注文件 + 结构化字段（JSON/表格）+ 对齐映射 + 清单/校验信息（manifest），并与样本 ID、版本号、人员信息关联。

##### A.3 关键标注规则

###### A.3.1 勾画规则

模型输出仅作候选提示（如中心点/初始轮廓/掩膜），最终轮廓由标注人员修订确认；多发结节按任务约定编号并分别输出；边界不清/伪影遮挡等情况按任务约定策略处理，并记录异常标记。

### A.3.2 分级规则

在勾画确认后执行；分级取值范围与依据应明确；分级仅用于记录既有分级结论，不用于生成新的医学诊断结论。

### A.3.3 文本标注规则

字段定义、取值集合、缺失/不确定编码方式应明确；结构化字段与自由文本不一致时，按任务约定优先级处理并记录依据。

### A.3.4 跨模态对齐规则

对齐限定在同一病例同一次检查范围内；明确允许的一对一/一对多/多对一映射；图像—文本冲突样本触发复核/仲裁并记录处理结论。

### A.3.5 一致性监控

可采用抽样复标/交叉复核方式监控人员间一致性与自身一致性，并记录差异与整改结果（频次与比例由任务定义约定）。

### A.3.6 角色与职责

内容包含如下：

- a) 角色配置：至少包括勾画标注、勾画审核、分级标注、分级仲裁；涉及文本与对齐时设置文本标注/审核（可兼任但职责应分离描述）；
- b) 能力要求：勾画标注人员具备相应判读能力并完成规则与工具培训；审核/仲裁人员具备更高资质与复核能力，负责确认与争议裁决；
- c) 上岗要求：可设置岗前考核与复测机制（指标与阈值由任务定义约定）。

## A.4 工具与环境

### A.4.1 工具能力

支持超声图像读取显示、半自动勾画、人工编辑、审核流转、导出与版本记录；支持分级录入、文本字段标注/抽取、对齐映射建立、冲突提示与导出；支持权限与日志留存。

### A.4.2 环境要求

终端显示与交互满足任务需求（示例：分辨率不低于 1920×1080）；在受控网络与分级授权条件下访问数据，满足数据安全与保密要求。

## A.5 数据准备与样本约束

内容包含如下：

- a) 采集与合规信息：记录采集时间范围、采集机构/设备与合规性信息，并明确数据来源人群范围；
- b) 清洗与剔除：剔除不完整、严重伪影遮挡、分辨率显著不足、带有影响标注的文字测量标记等样本，并记录剔除原因；
- c) 去重与关联：开展去重核验（维度由任务定义约定）；确保影像与报告在检查级（或约定粒度）可正确关联，关联失败样本隔离并记录处理结果。

## 附录 B

### (资料性)

#### 常用数据结构字段说明及示例

##### B.1 DICOM格式

在实际使用中, DICOM 文件应保留完整 Header 信息, 包括 ImagePosition(Patient)、ImageOrientation (Patient)、Pixel Spacing、Slice Thickness、Frame of Reference UID 等字段。若 DICOM 文件被错误压缩或头信息被删除, 将导致影像无法重建正确三维结构。以下为一个正确的 DICOM 头信息示例(节选), 展示关键结构字段:

```
(0020,0032)Image Position(Patient):[-158.0, -158.0, -240.0]
(0020,0037)Image Orientation(Patient):[1.0,0.0,0.0,0.0,1.0,0.0]
(0028,0030)Pixel Spacing:[0.742,0.742]
(0018,0050)Slice Thickness:1.000
(0020,1041)Slice Location:-240.0
(0020,000E)Series Instance UID: 1.2.840.113619.2.55.3.604688234.652.159976
(0020,0052)Frame of Reference UID: 1.2.840.113619.2.55.2.1234567890
```

该示例包含完整的空间矩阵信息, 可保证影像在三维物理空间中正确定位。

##### B.2 NIfTI格式

对于科研数据或 AI 模型训练, 常使用 NIfTI 格式(.nii 或.nii.gz)。NIfTI 文件将三维影像存储为一个整体, 同时在 Header 中包含 affine 矩阵, 用于描述坐标系统、体素间距与方向信息。为了确保模型训练端与标注端的一致性, affine 矩阵必须正确记录, 尤其是在重采样、图像旋转、切割等操作后必须同步更新。以下为一个正确的 NIfTI affine 矩阵示例:

```
Affine:
[[0.742, 0.000, 0.000, -158.0],
 [0.000, 0.742, 0.000, -158.0],
 [0.000,0.000, 1.000, -240.0],
 [0.000,0.000, 0.000, 1.000]]
```

其中

- a) 左上角 3x3 矩阵表示方向(与 DICOM Orientation 对应);
- b) 最后一列前三项为 origin;
- c) 对角线表示体素间距(spacing);

##### B.3 JSON格式

所有标注结果应独立存为结构化标注文件, 推荐使用JSON格式。JSON中必须包含影像引用(如UID或文件名)、标签编码(如liver、tumor)、标注类型(mask、polygon、bbox)、空间坐标(pixel或voxel

坐标)、属性信息(如形态描述、组织学特征)以及版本号、标注者信息等元数据。以下提供一个完整的JSON标注文件示例:

```
{
  "image_id": "CT_00123",
  "spacing": [0.742, 0.742, 1.0],
  "origin": [-158.0, -158.0, -240.0],
  "direction": [[1,0,0],[0,1,0],[0,0,1]],
  "annotations": [
    {
      "label": "liver",
      "type": "mask",
      "mask_file": "CT_00123_liver.nii.gz",
      "attributes": {
        "organ_system": "digestive",
        "level": "organ"
      }
    },
    {
      "label": "tumor",
      "type": "mask",
      "mask_file": "CT_00123_tumor_01.nii.gz",
      "attributes": {
        "lesion_type": "HCC",
        "shape": "irregular",
        "enhancement": "arterial hyperenhancement"
      }
    }
  ],
  "annotator": "doctor_A",
  "version": "1.0",
  "timestamp": "2025-03-12T10:25:00"
}
```

此示例展示了完整的影像空间信息、标注文件路径、标签体系信息及属性标签,是AI模型训练与标注审核系统所需要的标准化结构。

## 附录 C

### (资料性)

#### 多模态标注标签体系示例

##### C.1 标签体系示例（节选）

本文件采用“类别（category）—标签编码（label\_code）—标注类型（type）—属性（attributes）”的结构组织标注信息。标签体系可按器官、病灶、病例级属性等维度分层扩展：

- a) 器官类（organ）：如 LUNG\_LEFT、LUNG\_RIGHT；
- b) 病灶类（lesion）：如 PNEUMONIA\_GGO、PNEUMONIA\_CONSOLIDATION；
- c) 属性类（attribute）：如 CASE\_LEVEL（病例级诊断、分级、评分等）。

注：标签命名宜采用英文大写/下划线形式，保持跨机构一致性；如需本地化显示，可在工具层做映射，不影响机器可读编码。

##### C.2 多模态标注数据 JSON Schema（结构要点）

多模态标注数据建议采用 JSON（或 JSON Lines）组织，基本结构如下：

- a) task\_id：任务唯一标识；
- b) image：影像对象信息（至少包括 uid、modality、shape、spacing、origin、direction 等）；
- c) annotations[]：标注对象列表（器官/病灶/属性等）；
- d) metadata：标注、复核、软件、时间戳、质控、版本与审计信息。

##### C.3 annotations 对象字段（最小字段集）

每条标注对象宜至少包含：

- a) id：标注对象唯一 ID；
- b) category：类别（organ / lesion / attribute / …）；
- c) label\_code：标签编码；
- d) type：标注类型（mask / bbox / polygon / point / attribute）；
- e) data：几何或属性载荷；
  - mask\_file（mask 类型）
  - bbox（bbox 类型）
  - polygon + slice\_index（2D polygon）
  - attributes（属性类/补充字段）
  - cross\_modal\_links[]（可选）：跨模态关联 ID（用于图像—图像、图像—文本、图像—结构化字段的对应）

##### C.4 metadata（追溯与质控字段要点）

为满足可追溯性，建议在 metadata 中记录：

- a) annotator / reviewer / institution；

- b) software.name / software.version / software.ai\_model（可选）；
- c) timestamps.start / timestamps.end；
- d) qc.status / qc.issues（可选）；
- e) version: 数据包版本号；
- f) audit[]: 关键操作流水（创建、模型建议、人工修订、复核确认等）。

### C.5 示例（片段）

以下为示例片段（仅展示结构，不代表唯一实现）：

```
{
  "task_id": "TASK_CT_LUNG_000001",
  "image": {"uid": "...", "modality": "CT", "shape": [512, 512, 320], "spacing": [0.7, 0.7, 1.0]},
  "annotations": [

{"id": "ORG_LUNG_LEFT", "category": "organ", "label_code": "LUNG_LEFT", "type": "mask", "mask_file": "masks/xxx.nii.gz"},
{"id": "LESION_01", "category": "lesion", "label_code": "PNEUMONIA_GGO", "type": "bbox", "bbox": [...], "cross_modal_links": ["TXT_SPAN_01"]}
  ],
  "metadata": {"annotator": "...", "reviewer": "...", "version": "v1.0"}
}
```

## 附录 D (资料性) 标注质量评价指标与计算方法

### D.1 总则

本附录给出多模态模型精准医学图像数据标注中常用的准确性评价指标与计算方法。对影像目标检测/定位类标注宜采用对象级指标；对分割类标注宜采用像素/体素级指标；对分类或结构化文本字段标注宜采用样本级指标。

### D.2 召回率

被正确检测出的目标数量占有所有目标数量的比例，计算方法见式 (D.1)

$$Rec = \frac{TP_{object}}{TP_{object} + FN_{object}} \dots\dots\dots (D.1)$$

式中：

Rec —— 召回率；

TP\_object —— 正确检测目标数；

FN\_object —— 漏检目标数。

### D.3 精确度

被正确检测出的目标数量占有所有被检出对象的比例，计算方法见式 (D.2)

$$Pre = \frac{TP_{object}}{TP_{object} + FP_{object}} \dots\dots\dots (D.2)$$

式中：

Pre —— 精确度；

FP\_object —— 误检目标数。

### D.4 Dice 系数

用于计算两个分割区域的重合度，计算方法见式 (D.3)

$$Dice = 2 \times \frac{|A \cap B|}{|A| + |B|} \dots\dots\dots (D.3)$$

式中：

A —— 真实分割区域；

B —— 预测分割区域。

### D.5 Conformity 系数

错误分割的像素数与正确分割目标区域像素数之间的比例关系，计算方法见式 (D.4)

$$Conformity = 1 - \frac{FP_{pixel}}{TP_{pixel}} \dots\dots\dots (D.4)$$

式中：

FP\_pixel —— 错误分割像素数；

TP\_pixel —— 正确分割像素数。

#### D.6 交并比

评价预测的分割区域与真实分割区域的重合程度，计算方法见式 (D.5)

$$IOU = \frac{|A \cap B|}{|A \cup B|} \dots\dots\dots (D.5)$$

式中：

IOU —— 交并比。

#### D.7 Hausdorff 距离

描述两个分割区域轮廓的距离，双向 Hausdorff 距离计算方法见式 (D.6)

$$d_H(X, Y) = \max \left\{ \max_{x \in X} \min_{y \in Y} d(x, y), \max_{y \in Y} \min_{x \in X} d(x, y) \right\} \dots\dots\dots (D.6)$$

式中：

$d_H(X, Y)$  —— 双向 Hausdorff 距离；

X —— 预测的分割区域；

Y —— 人工标注的分割区域；

$d(x, y)$  —— X、Y 两个区域任意两点之间的距离。

#### D.8 Pearson 相关系数

两个变量的协方差除以其标准差乘积

$$\rho(X, Y) = \frac{cov(X, Y)}{\sigma_X \sigma_Y} = E[(X - \mu_X)(Y - \mu_Y)] / (\sigma_X \sigma_Y) \dots\dots\dots (D.7)$$

式中：

$\rho(X, Y)$  —— Pearson 相关系数；

$\mu_X$  —— X 的平均值；

$\mu_Y$  —— Y 的平均值；

$\sigma_X$  —— X 的标准差；

$\sigma_Y$  —— Y 的标准差；

E —— 期望（即平均值）。

Pearson 相关系数的绝对值越大，相关性越强；

相关系数越接近于 1 或 -1，相关度越强；

相关系数越接近于 0，相关度越弱。

#### D.9 灵敏度

$$Sen = \frac{TP_{sample}}{TP_{sample} + FN_{sample}} \times 100\% \quad \dots\dots\dots (D.8)$$

式中：

Sen —— 灵敏度；

TP\_sample —— 真阳性样本数量；

FN\_sample —— 假阴性样本数量。

#### D.10 特异度

$$Spe = \frac{TN_{sample}}{FP_{sample} + TN_{sample}} \times 100\% \quad \dots\dots\dots (D.9)$$

式中：

Spe —— 特异度；

TN\_sample —— 真阴性样本数量；

FP\_sample —— 假阳性样本数量。

#### D.11 准确率

准确率用 Acc 表示，计算方法见公式 (D.10)：

$$Acc = \frac{\sum_{j=1}^m N_{jj}}{\sum_{i=1}^m \sum_{j=1}^m N_{ij}} \quad \dots\dots\dots (D.10)$$

式中：

$N_{ij}$  —— 泛指混淆矩阵第 i 行、第 j 列的元素；

$N_{ji}$  —— 泛指混淆矩阵第 j 行、第 i 列的元素；

#### D.12 Kappa 系数

$$k = \frac{Acc - p_e}{1 - p_e} \quad \dots\dots\dots (D.11)$$

其中:

$$p_e = \frac{\sum_{i=1}^m (\sum_j N_{ij})(\sum_j N_{ji})}{(\sum_{i=1}^m \sum_{j=1}^m N_{ij})^2} \dots\dots\dots (D.12)$$

式中:

$N_{ij}$  —— 泛指混淆矩阵第  $i$  行、第  $j$  列的元素;

$N_{ji}$  —— 泛指混淆矩阵第  $j$  行、第  $i$  列的元素;

Acc —— 准确率。

## 参 考 文 献

- [1] GB/T 42755—2023. 人工智能 面向机器学习的数据标注规程[S]. 北京: 中国标准出版社, 2023.
- [2] GB/T 35273—2020. 信息安全技术 个人信息安全规范[S]. 北京: 中国标准出版社, 2020.
- [3] GB/T 1.1—2020. 标准化工作导则 第1部分: 标准的结构和编写规则[S]. 北京: 中国标准出版社, 2020.
- [4] GB/T 43697—2024. 数据分类分级规则[S]. 北京: 中国标准出版社, 2024.
- [5] GB/T 22239—2019. 信息安全技术 网络安全等级保护基本要求[S]. 北京: 中国标准出版社, 2019.
- [6] 国家卫生和计划生育委员会. WS 445—2014 医学影像诊断中心基本标准[S]. 北京: 中国标准出版社, 2014.
- [7] 国家卫生健康委员会. WS 519—2018 医学影像诊断中心管理规范[S]. 北京: 中国标准出版社, 2018.
- [8] National Electrical Manufacturers Association (NEMA). Digital Imaging and Communications in Medicine (DICOM) Standard[S]. Rosslyn, VA: NEMA, 现行版本.
- [9] Neuroimaging Informatics Technology Initiative (NIFTI). NIFTI-1 Data Format Specification[S]. Bethesda, MD: National Institutes of Health, 现行版本.
- [10] Radiological Society of North America (RSNA). RadLex Radiology Lexicon[EB/OL]. <https://www.rsna.org/radlex>.
- [11] International Health Terminology Standards Development Organisation (IHTSDO). SNOMED CT: Systematized Nomenclature of Medicine—Clinical Terms[S]. London: IHTSDO, 现行版本.
- [12] Regenstrief Institute. LOINC: Logical Observation Identifiers Names and Codes[S]. Indianapolis: Regenstrief Institute, 现行版本.
- [13] Health Level Seven International (HL7). FHIR: Fast Healthcare Interoperability Resources[S]. Ann Arbor, MI: HL7, 现行版本.
- [14] 全国人民代表大会常务委员会. 中华人民共和国数据安全法[S]. 北京: 人民出版社, 2021.
- [15] 全国人民代表大会常务委员会. 中华人民共和国个人信息保护法[S]. 北京: 人民出版社, 2021.
- [16] 全国人民代表大会常务委员会. 中华人民共和国网络安全法[S]. 北京: 人民出版社, 2017.
- [17] 国家卫生健康委员会. 电子病历管理规范[S]. 北京: 国家卫生健康委员会, 现行版本.
- [18] 国家卫生健康委员会. 医疗质量管理办法[S]. 北京: 国家卫生健康委员会, 现行版本.
- [19] 国家卫生健康委员会. 国家智慧医疗标准体系建设指南[S]. 北京: 国家卫生健康委员会, 现行版本.
- [20] 国家医学影像数据库(NIMI)建设工作组. 国家医学影像数据库(NIMI)技术体系与数据规范[S]. 北京: 国家卫生健康委员会, 现行版本.