

# T/CAB

## 中国产学研合作促进会团体标准

T/ CAB 0161-2022

### 肺结节 CT 影像采集和数据处理技术规范

Technical Specification for Collection and Processing of Pulmonary  
Nodule CT Images

2022-07-19 发布

2022-07-19 实施

中国产学研合作促进会

发布

## 目 次

前言 .....	2
1 范围 .....	3
2 规范性引用文件 .....	3
3 术语和定义 .....	3
4 影像采集和处理标准 .....	4



## 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这些专利的责任。

本文件由中国产学研合作促进会提出并归口。

本文件起草单位：上海长征医院影像科、中国食品药品检定研究院医疗器械检定所光机电室、国家卫生健康委能力建设和继续教育中心、西安交通大学第一附属医院医学影像科、大连大学附属中山医院影像科、空军军医大学西京医院、中华放射学杂志、卫生部北京医院放射科、上海交通大学医学院附属瑞金医院、苏州大学附属第一医院放射科、天津医科大学肿瘤医院、解放军总医院第一医学中心、宁夏医科大学总医院、北京协和医院放射科、中国医学科学院肿瘤医院影像诊断科、华中科技大学同济医学院附属同济医院放射科、上海市胸科医院、上海市肺科医院、上海市公共卫生临床中心、上海交通大学医学院附属仁济医院、复旦大学附属华山医院、复旦大学附属华东医院、复旦大学附属肿瘤医院、广州医科大学附属第一医院、浙江中医药大学附属第一医院、中国医科大学附属第一医院、兰州大学第二医院、吉林大学白求恩第一医院、南昌大学第二附属医院、浙江大学医学院附属第二医院、首都医科大学附属北京朝阳医院、四川大学华西医院、上海九州智慧医学影像研究院、推想医疗科技股份有限公司、上海杏脉信息科技有限公司、数坤（上海）医疗科技有限公司、北京深睿博联科技有限责任公司、上海联影智能医疗科技有限公司、慧影医疗科技（北京）股份有限公司、北京医准智能科技有限公司、强联智创（北京）科技有限公司、医智源健康科技（杭州）有限公司、通用电气医疗系统（中国）有限公司、飞利浦（中国）投资有限公司、西门子医疗系统有限公司。

本文件主要起草人：刘士远、杨爱平、王浩、萧毅、张琳琳、郭佑民、伍建林、郑敏文、刘凯、马兆毅、陈起航、胡春洪、叶兆祥、赵绍宏、朱力、宋伟、于红、吴宁、夏黎明、范丽、史景云、施裕新、吴华伟、李铭、曾庆思、许茂盛、张立娜、张惠茅、严福华、周俊林、龚良庚、张敏鸣、杨旗、宋彬、姚振威、宋少莉、陈海泉、姜丽岩、王升平、蒋薇、孟祥峰、李佳戈、王少康、何川、马春娥、柴象飞、李一鸣、郑颖、杨光明、王平、赵磊、黄毅、詹翊强、魏国亮。

# 肺结节 CT 影像采集和数据处理技术规范

## 1 范围

本文件确立了肺结节CT影像数据采集和处理标准框架，规定相关数据采集设备、采集过程、数据脱敏、数据处理的具体规程。

本文件适用于面向人工智能、临床科研、医学教育、医药研发等场景，需要实施肺结节CT影像采集和数据处理的企业、高校、科研院所、政府机构等用户。

## 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 25000.12-2017 系统与软件工程 系统与软件质量要求和评价 (SQuaRE)

GB/T 35295-2017 信息技术大数据

YY/T 1833.1-2022 人工智能医疗器械质量要求和评价 第1部分：术语

YY/T 1833.2-2022 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

## 3 术语和定义

下列术语和定义适用于本文件。

### 3.1 数据 data

信息的可再解释的形式化表示，以适用于通信、解释或处理。

[来源：GB/T 25000.12-2017，4.2]

### 3.2 特征 features

能表达模式本质的功能或结构特点的可度量属性，如大小、纹理、形状等。好的特征应能使同类模式聚类、不同类模式分离。

[来源：ISBN 978-7-03-059487-7，08.0386]

### 3.3 数据质量 data quality

在指定条件下使用时，数据的特性满足明确的、隐含的要求的程度。

[来源：GB/T 25000.12-2017，4.3]

### 3.4 数据集 data set

数据记录汇聚的数据形式。

[来源：GB/T 35295-2017，2.1.46]

### 3.5 数据清洗 data cleaning

检测和修正数据集中错误数据项的过程。

[来源: ISBN 978-7-03-059487-7, 07.0392, 有修改]

### 3.6 数据采集 data acquisition

数据由生成装置按照数据采集规范生成, 以数字化格式存储并传输到目标系统的过程。

[来源: YY/T 1833.1-2022, 3.3.25]

### 3.7 数据脱敏 data masking

对敏感信息通过去标识化或匿名化, 实现敏感隐私数据的可靠保护。

[来源: YY/T 1833.1-2022, 3.3.26]

### 3.8 数据查重 data duplicate checking

检查和比对数据集中的数据项以及对重复数据进行排除, 确保数据的唯一性。

## 4 影像采集和数据处理要求

### 4.1 采集设备

#### 4.1.1 兼容性要求

设备方面应考虑CT制造商、设备型号、同一台设备的不同成像参数配置(如层厚、层间距、剂量设置、重建算法等)。从操作层面来说, 可以依据表1参数范围进行选择。

表1 采集设备兼容性要求

层厚与层间距	推荐图像重建层厚不超过 1mm。为较为准确地进行计算机辅助检测及容积分析, 层厚 $\leq 1$ mm 者, 可以无间隔连续重建; 层厚 $>1$ mm 者, 重建间隔选择准直层厚的 50%~80%。
影像数据	应符合 DICOM3.0 协议标准数据。
探测器	16 排以上

#### 4.1.2 采集要求

##### 4.1.2.1 采集方式

为了在诊断过程中充分识别病灶, 胸部检查采用螺旋CT平扫或增强扫描模式。

##### 4.1.2.2 采集协议

符合DICOM 3.0协议的标准数据。

##### 4.1.2.3 扫描及图像输出参数

扫描参数	管电压（70 kV~140kV），管电流（10 mA~400mA）组合；增强扫描：静脉注射含碘对比剂 60ml~70ml，注射速率 2ml/s~5ml/s。
辐射剂量	0.4mSv~8mSv；也可以使用 mGy 等其他单位描述剂量。扫描时应开启“剂量报告（dose report）”功能，以便将机器自动生成的剂量报告进行常规存储。
窗宽窗位	推荐肺窗（窗位，-700HU~-500HU；窗宽，1500HU~2000HU），纵隔窗（窗位，30HU~70HU；窗宽，350HU~400HU）；其他窗宽、窗位可以用于辅助阅片，例如骨窗、中间窗。
重建方式	采用高空间分辨率算法和标准算法（或软组织算法）同时进行重建。注：在 DICOM 字段中，不同设备对重建算法的表述是多种多样的，数据集制造责任方宜与设备制造商确认其含义。
显示方式	应采用横断面结合靶重建、多平面重组（MPR）及最大密度投影（MIP）等后处理技术辅助阅片，多方位显示肺结节的形态学特征。

## 4.2 采集过程

### 4.2.1 采集人员要求

#### 4.2.1.1 人员资质要求

数据采集技师应在二甲及以上医院从事CT图像采集工作5年以上。

#### 4.2.1.2 培训

根据数据采集要求，对参与数据采集的人员进行培训。包括但不限于：设备操作培训、对比剂使用培训、图像处理培训。

#### 4.2.1.3 考核

采集人员考核标准，要求熟悉胸部平扫及增强CT扫描相关技术要点，能根据临床情况调整合适的参数进行扫描，以获得符合要求的数据。

设备操作考核内容主要有：设备参数设定和操作、增强扫描、对比剂注入、窗宽窗位设定。考核标准从操作规范性、操作娴熟度、操作稳定度三方面进行综合评价。

### 4.2.2 采集过程要求

每个病例的扫描范围至少从肺尖到肋膈角（包括全肺），训练受检者呼吸屏气，不相关的部位可以删除。图像视野应当覆盖整个胸部横断面。受检者吸气末一次屏气完成扫描，避免运动伪影、误操作等。成像过程使用的设备应当在有效期内，并符合《放射科管理规范与质控标准（2017版）》要求。

### 4.2.3 数据质量要求

数据应采集具有标准DICOM格式的真实、原始临床数据，图像矩阵不低于512×512。数据集制造责任方应确保保脱敏后的原始数据的准确性，避免有损压缩、图像篡改等操作。每个病例的图像应当保持连续完整，不得出现缺层、错层等情况。DICOM头文件记录的信息应保持完备，满足肺结节数据分层抽样的需要。不同医院的数据质量、成像参数配置应保持一致。

### 4.3 数据处理要求

#### 4.3.1 数据合规性

##### 4.3.1.1 要求

伦理批件或其他官方批件应当真实合规。从志愿者、体检中心、社区筛查项目、科研项目等其他途径获取的图像应当通过同等效力的批准程序保证数据脱敏、受检者隐私安全和受检者利益。

数据脱敏是保护受检者隐私的技术手段。图像本身、头文件、附属文件，以及数据集的元数据中均不应包含与受检者隐私有关的任何信息。

受检者个人信息、医疗保险信息、活动轨迹、经济状况、家庭情况、财务信息、社会信息等个人隐私不得搜集和显示；图像头文件中涉及受检者隐私的字段必须进行脱敏处理。

##### 4.3.1.2 内容

数据集制造责任方应解析序列DICOM头文件中的标签（tag），检查各个字段并去除敏感信息或隐私信息，尤其是一些主要的共有字段，主要字段包括：

- a) 机构名称 (Institution Name) 0x00080080;
- b) 机构地址 (Institution Address) 0x00080081;
- c) 受检者姓名 (Patient Name) 0x00100010;
- d) 受检者生日 (Patient Birth Date) 0x00100030;
- e) 受检者地址 (Patient Address) 0x00101040。

##### 4.3.1.3 方法

读取和检查DICOM头文件里与受检者个人信息有关的字段。如果该字段含有敏感信息，应采取数据脱敏或数据匿名化措施，防止隐私泄露。

#### 4.3.2 数据清洗

##### 4.3.2.1 要求

对数据进行检查和确认，去除不合格数据，筛选满足质量要求的数据。

图像的质量要求一般包括格式的有效性、单个文件的完整性、序列的完整性、图像内容的合理性等。比如不符合DICOM标准的图像、破损或无法读取的图像、不连续的图像序列、与肺部无关的序列都是应去除的对象。

数据集制造责任方可根据需要，制定更具体的图像质量要求，比如峰值信噪比等指标。图像清洗应当建立规程，形成操作记录。

##### 4.3.2.2 方法

比较医院原始数据与数据集制造责任方采集的数据，对数据的有效性、合理性和完整性进行检查。对各个影像序列的连续性、内部一致性进行人工分析，检查是否出现数据缺层、错层、序列缺失、超范围扫描、伪影遮挡等问题。伪影遮挡问题除了程序自动化判断外，需要人工复核。

#### 4.3.3 数据查重

##### 4.3.3.1 要求

图像数据不应与其他公开数据集发生重合；同一批数据内部不得出现图像重合；所有病例应当真实唯一。数据集制造责任方应明确数据查重的范围。

#### 4.3.3.2 方法

在原始数据入库处理的过程中，数据集制造责任方可针对DICOM文件解析得到序列号（Series Instance UID）和图像实例编号（Instance Number）的tag值，与数据集内部和其他公开数据集进行比对，也可以采用消息摘要算法等其他信息进行比对。若发现重复，则对此部分数据进行记录和封存，从数据集中排除。

## 参 考 文 献

- [1]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[Z]. 北京: 国家药品监督管理局医疗器械技术审评中心, 2019.
- [2]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)[Z]. 北京: 国家药品监督管理局医疗器械技术审评中心, 2020.
- [3]国家药品监督管理局. 人工智能医疗器械注册技术审查指导原则(征求意见稿)[Z]. 北京: 国家药品监督管理局, 2021.
- [4] 中国食品药品检定研究院, 中华医学会放射学分会心胸学组. 胸部CT肺结节数据标注与质量控制专家共识(2018)[J]. 中华放射学杂志, 2019, 53(1):9-15.
- [5]中华医学会放射学分会, 中国食品药品检定研究院, 国家卫生健康委能力建设与继续教育中心, 等. 胸部CT肺结节数据集构建及质量控制专家共识[J]. 中华放射学杂志, 2021, 55(2):104-110.
- [6]中华医学会放射学分会. 放射科管理规范与质控标准(2017版)[M]. 徐克, 刘士远等. 北京: 人民卫生出版社, 2017.
- [7] Digital Imaging and Communications in Medicine(DICOM). NEMA PS 3.1-PS 3.12. Rosslyn, VA, The National Electrical Manufacturers Association 1992-1995.
- [8]中华医学会放射学分会. 放射科管理规范与质控标准(2017版)[M]. 徐克, 刘士远等. 北京: 人民卫生出版社, 2017.
- [9]石明国, 王鸣鹏, 余建明. 放射师临床工作指南[J]. 人民卫生出版社, 2013.
- [10]计算机科学技术名词审定委员会. 计算机科学技术名词[M]. 科学出版社, 2002.