ICS 35.240.80
CCS L67

团　　体　　标　　准

T/ISC 0167—2026

医疗行业专科高质量数据集建设指南 第1部分：呼吸内科

High-quality Datasets for Medical Specialty Construction Guidelines Part 1: Respiratory Medicine

（发布稿）

2026 - 10 - 28 发布　　2026 - 11 - 28 实施

中 国 互 联 网 协 会　　发 布

目　　次

1 范围 1
2 规范性引用文件 1
3 术语和定义 1
3.1 个人健康医疗数据 personal health data 1
3.2 健康医疗数据 health data 1
3.3 健康医疗数据控制者 health data controller 1
3.4 数据元 data element 2
3.5 数据标注 data labeling 2
3.6 标志 indicator 2
4 符号和缩略语 2
5 总体要求 2
6 数据需求 3
6.1 应用场景 3
6.2 数据范围 3
7 数据规划 3
7.1 数据架构 3
7.2 实施计划 3
8 数据采集 3
8.1 基本信息 3
8.2 健康史 3
8.3 主诉与现病史 4
8.4 体格检查数据 4
8.5 临床辅助检查 5
8.6 实验室检验 6
8.7 治疗相关信息 6
8.8 出院与随访 7
9 数据预处理 8
9.1 数据清洗 8
9.2 数据格式转换 8
10 数据标注 8
10.1 文本问答对构建 8
10.2 图像类数据标注 9
10.3 音频类数据标注 10
10.4 波形数据标注 10
附 录 A （资料性） 文本问答对详细表述示例 12
附 录 B （资料性） 呼吸内科常见病种图像类数据标注示例 13
附 录 C （资料性） 呼吸内科常见病种音频类数据标注示例 17
附 录 D （资料性） 呼吸内科常见病种波形类数据标注示例 21

参 考 文 献.....27

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。

本文件是《医疗行业专科高质量数据集建设指南》系列标准的第1部分。

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。

本文件由中国互联网协会归口。

本文件起草单位：广州医科大学附属第一医院、广州呼吸健康研究院、中国信息通信研究院、上海明品医学数据科技有限公司、上海中数智模科技有限公司、北京大学第一医院、内蒙古自治区人民医院、中国医科大学附属第一医院、吉林大学第二医院、复旦大学附属中山医院、浙江大学医学院附属第二医院、郑州大学第一附属医院、新疆维吾尔自治区人民医院、兰州大学第一医院、四川大学华西医院、云南省第一人民医院、复旦大学附属儿科医院、三明市第一医院、上海人工智能实验室、熵极涌现数据（天津）有限公司。

本文件主要起草人：钟南山、李时悦、黄超仪、陈智、陆殷、苏越明、简文华、王文熙、王柱锋、钟长镐、韩林月、桑新雨、武雅文、冯天宜、闵栋、姚娟娟、梁志龙、潘昊天、高柳村、王少伟、王晟晟、马靖、廖纪萍、高笑宇、马尚寅、李文扬、张捷、宋元林、李雯、王华启、魏雪梅、包海荣、刘丹、张云辉、张晓波、葛小玲、马傲、孙英明、陈曦、王超璜、徐捷、曲通。

医疗行业专科高质量数据集建设指南 第1部分：呼吸内科

1 范围

本文件规定了医疗行业呼吸内科高质量数据集建设的数据需求、数据规划、数据采集、数据预处理和数据标注等内容。

本文件适用于医疗机构、科研院所、医疗健康数据运营机构以及人工智能企业开展呼吸内科高质量数据集的设计、采集、处理、标注及维护等工作。

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 39725—2020 《信息安全技术 健康医疗数据安全指南》

GB/T 42755—2023 《人工智能 面向机器学习的数据标注规程》

GB/T 18391.1—2009 《信息技术 元数据注册系统（MDR） 第1部分：框架》

WS/T 363.1—2023 《卫生健康信息数据元目录 第1部分：总则》

WS/T 363.7—2023 《卫生健康信息数据元目录 第7部分：体格检查》

WS/T 364.5—2023 《卫生健康信息数据元值域代码 第5部分：健康危险因素》

3 术语和定义

下列术语和定义适用于本文件。

3.1

个人健康医疗数据

单独或者与其他信息结合后能够识别特定自然人或者反映特定自然人生理或心理健康的相关电子数据。

[来源：GB/T 39725—2020，3.1]

3.2

健康医疗数据

个人健康医疗数据以及由个人健康医疗数据加工处理之后得到的健康医疗相关电子数据。

[来源：GB/T 39725—2020，3.2]

3.3

健康医疗数据控制者

能够决定健康医疗数据处理目的、方式及范围等的组织或个人，包括提供健康医疗服务的组织、医保机构、政府机构、健康医疗科学研究机构、个体诊所等。

[来源：GB/T 39725—2020，3.5]

3.4

数据元

由一组属性规定其定义、标识、表示和允许值的数据单元。

[来源：GB/T 18391.1—2009，3.3.8]

3.5

数据标注

给数据样本指定目标变量和赋值的过程。

[来源：GB/T 42755-2023，3.1]

3.6

标志

指示符，两个且只有两个表明条件的值，其允许值通常被限定为真（T/True）或假（F/False）。

[来源：WS/T 363.1—2023，表 2，有修改]

4 符号和缩略语

下列符号和缩略语适用于本文件。

CT：计算机断层扫描（Computed Tomography）

COPD：慢性阻塞性肺疾病（Chronic Obstructive Pulmonary Disease）

CAT：COPD 评估测试（COPD Assessment Test）

DLCO：肺一氧化碳弥散量（Diffusing Capacity of the Lung for Carbon Monoxide）

EMR：电子病历系统（Electronic Medical Record System）

FVC：用力肺活量（Forced Vital Capacity）

FEV1：第1秒用力呼气容积（Forced Expiratory Volume in One Second）

GOLD：慢性阻塞性肺疾病全球倡议（Global Initiative for Chronic Obstructive Lung Disease）

LIS：实验室信息系统（Laboratory Information System）

MVV：最大自主通气量（Maximum Voluntary Ventilation）

PACS：影像归档和通信系统（Picture Archiving and Communication System）

PEF：呼气峰流速（Peak Expiratory Flow）

SAS：睡眠呼吸暂停综合征（Sleep Apnea Syndrome）

TLC：肺总量（Total Lung Capacity）

VC：肺活量（Vital Capacity）

5 总体要求

呼吸内科高质量数据集建设应按照生命周期有序展开，包括数据需求、数据规划、数据采集、数据预处理、数据标注等阶段。各阶段主要按以上顺序逐步开展，同时会对其他阶段进行反馈，或者会在其他阶段反馈下进行迭代优化。

6 数据需求

6.1 应用场景

呼吸内科高质量数据集的数据需求应根据预期人工智能应用确定，应用场景包括疾病诊断、治疗决策、检查解读、预后与随访、紧急情况识别。

6.2 数据范围

呼吸内科高质量数据集的数据范围包括基本信息、健康史、主诉与现病史、体格检查数据、临床辅助检查、实验室检验、治疗相关信息、出院与随访。

7 数据规划

7.1 数据架构

呼吸内科高质量数据集应以患者为中心构建数据架构，形成覆盖原始数据、预处理数据和标注或衍生数据的数据组织体系，包括：

a) 数据主题域：基本信息、健康史、主诉与现病史、体格检查数据、临床辅助检查、实验室检验、治疗相关信息、出院与随访；

b) 数据来源：EMR、LIS、PACS、肺功能检查系统、内窥镜系统、病理系统、多导睡眠监测系统、手术麻醉系统、随访管理系统及集成大数据平台等；

c) 数据层级：原始数据层、预处理数据层、标注或衍生数据层。

7.2 实施计划

呼吸内科高质量数据集构建实施计划应覆盖数据采集、数据预处理和数据标注等阶段，明确实施步骤、责任分工、工具平台、时间安排和交付物，涉及多机构数据协作时，宜明确数据提供、数据处理、数据审核、数据使用和结果反馈等协作机制。

8 数据采集

8.1 基本信息

健康医疗数据控制者应从EMR或集成大数据平台采集患者基本信息，包括：

a) 性别；

b) 年龄。

8.2 健康史

8.2.1 个人史

健康医疗数据控制者应从EMR或集成大数据平台采集个人史相关信息，宜重点关注患者个人生活习惯及有无烟、酒等嗜好，职业与工作条件及有无工业毒物、粉尘等物质接触史，包括：

a) 吸烟史：吸烟标志、吸烟状态、被动吸烟场所、吸食烟草种类（见WS/T 364.5—2023表1至表3）等；

b) 饮酒史：饮酒标志、饮酒频率、饮酒量、饮酒种类等；

c) 职业与工作条件：职业类别、职业病危害因素类别（见WS/T 364.5—2023表19）、暴露时长、防护措施使用情况等；

d) 生活环境：居住地空气质量、燃料类型（见WS/T 364.5—2023表23）、居住环境潮湿/霉菌暴露情况等。

8.2.2 过敏史

健康医疗数据控制者应从EMR或集成大数据平台采集过敏史相关信息，包括：

a) 过敏史标志；

b) 过敏原：灰尘、花粉、特定药物、食物等；

c) 过敏状态：可疑、明确等。

8.2.3 既往史

健康医疗数据控制者应从EMR或集成大数据平台采集既往史相关信息，宜重点关注儿童期病史，包括：

a) 儿童期病史：儿童期严重呼吸道感染史、百日咳病史、麻疹病史等；

b) 相关疾病史：哮喘、慢性阻塞性肺疾病、肺炎、黄疸性肝炎、糖尿病、高血压等；

c) 药品信息：沙丁胺醇气雾剂、布地奈德吸入剂、阿司匹林肠溶片、降糖药、降压药等。

8.3 主诉与现病史

健康医疗数据控制者应从EMR或集成大数据平台采集就诊时间与患者症状，宜重点关注呼吸道症状，包括：

a) 呼吸道症状：咳嗽、胸闷、鼻塞、流涕、喘息、晕厥、静息状态呼吸困难等；

b) 非呼吸道症状：烦躁不安、心悸、皮疹、头晕、肌痛、腹胀/腹泻、恶心/呕吐等。

8.4 体格检查数据

8.4.1 胸部体格检查

健康医疗数据控制者应从EMR或集成大数据平台采集体格检查数据，宜参照WS/T 363.7—2023等相关标准重点关注胸部体格检查信息，包括：

a) 呼吸运动：呼吸频率、呼吸节律、呼吸深度、呼吸方式等；

b) 胸廓体征：胸廓形态、肋间隙状态、皮下气肿标志、胸壁压痛等；

c) 气管位置：气管居中标志、气管偏移方向等；

d) 肺部震颤：语音震颤（语颤）增强/减弱、胸膜摩擦感等；

e) 肺部叩诊：叩诊音性质、肺下界位置、肺下界移动度等；

f) 呼吸音：正常呼吸音分布、异常呼吸音标志、呼吸音减弱/消失、呼吸音粗糙等；

g) 肺部啰音：干啰音（哮鸣音）标志、湿啰音（水泡音）标志、捻发音标志、啰音部位等；

h) 其他听诊：语音共振强度、胸膜摩擦音标志等。

8.4.2 其他体格检查

健康医疗数据控制者应从EMR或集成大数据平台采集其他体格检查数据，包括：

a) 生命体征：体温、脉率、收缩压、舒张压等；

b) 一般状态：意识状态、体位、面容表情等；

c) 皮肤黏膜：皮肤颜色、皮肤温湿度、皮疹、出血点、杵状指（趾）标志等；

d) 淋巴结：浅表淋巴结肿大标志、肿大部位、淋巴结质地等；

e) 心脏检查：心尖搏动位置、心率、心律、心脏杂音标志等；
f) 腹部检查：肝脏肿大标志、脾脏肿大标志、腹部压痛、移动性浊音等；
g) 四肢检查：下肢水肿标志、水肿程度、下肢静脉曲张等。

8.5 临床辅助检查

8.5.1 胸部影像学检查

健康医疗数据控制者应从PACS或集成大数据平台采集影像学检查数据，包括：
a) 检查名称：肺部CT普通扫描、肺部CT平扫加增强扫描、肺部CT高分辨率扫描、肺动脉CT血管造影等；
b) 设备信息：CT设备品牌、CT设备型号、扫描参数、造影剂使用情况等；
c) 检查结果：肺纹理改变、肺门阴影、纵隔轮廓、病灶位置、病灶大小、密度值、影像学征象等。

8.5.2 肺功能检查

健康医疗数据控制者应从肺功能仪、LIS或集成大数据平台采集肺功能检查数据，包括：
a) 检查时症状状态：症状稳定、症状加重、不详等；
b) 检查名称：肺量计检查、支气管舒张试验、支气管激发试验、肺容积测定、肺弥散功能检查、气道阻力测定、6分钟步行测试、运动心肺功能检查等；
c) 设备信息：肺功能仪品牌、肺功能仪型号等；
d) 检查参数数值：FVC、FEV1、TLC、RV等。

8.5.3 呼吸介入检查

健康医疗数据控制者应从内窥镜系统、病理系统或集成大数据平台采集呼吸介入检查数据，包括：
a) 检查名称：支气管镜检查、超声支气管镜检查、经皮肺穿刺活检术、支气管肺泡灌洗术等；
b) 设备信息：内镜品牌、内镜型号、穿刺针规格等；
c) 介入操作描述：活检、刷检、灌洗、针吸等；
d) 检查结果：气道通畅度、黏膜情况、腔内病变形态、镜下初步诊断、灌洗液性状、细胞学/组织病理学诊断结果等。

8.5.4 睡眠呼吸检查

健康医疗数据控制者应从多导睡眠监测系统、便携式睡眠监测设备或集成大数据平台采集睡眠呼吸检查数据，包括：
a) 检查名称：多导睡眠监测、便携式睡眠呼吸监测、无创正压通气智能压力滴定等；
b) 设备信息：监测设备品牌、监测导联数（通道数）、面罩类型等；
c) 检查参数数值：呼吸暂停低通气指数、血氧饱和度下降指数、最低血氧饱和度、平均血氧饱和度、氧减总时间、总睡眠时间、睡眠效率、睡眠分期占比等；
d) 检查结果：阻塞性/中枢性/混合性睡眠呼吸暂停判断、病情严重程度分级等。

8.5.5 其他辅助检查

健康医疗数据控制者应从心电生理系统、超声系统或集成大数据平台采集其他辅助检查数据，包括：
a) 检查名称：心电图检查、心脏彩色多普勒超声、呼出气一氧化氮测定等；
b) 设备信息：检查设备品牌、型号、探头/示踪剂参数等；
c) 检查结果：肺动脉收缩压、右室射血分数、FeNO浓度值等。

8.6 实验室检验

8.6.1 血气分析

健康医疗数据控制者应从血气分析仪或LIS采集数据，包括：

a) 检验项目：动脉血气分析；

b) 辅助信息：吸氧浓度、吸氧方式；

c) 检验结果：酸碱度、氧分压、二氧化碳分压、血氧饱和度、乳酸、剩余碱、肺泡-动脉氧分压差等。

8.6.2 呼吸道微生物检验

健康医疗数据控制者应从LIS采集病原学数据，包括：

a) 标本类型：痰液、肺泡灌洗液、咽拭子、胸水等；

b) 检验项目：细菌培养及药敏、呼吸道病毒核酸检测、结核分枝杆菌检测、真菌检测、宏基因组测序等；

c) 检验结果：病原体名称、菌落计数、抗生素最低抑菌浓度、阳性或阴性判定等。

8.6.3 胸部肿瘤标记物

健康医疗数据控制者应从LIS采集数据，包括：

a) 检验项目：癌胚抗原、神经元特异性烯醇化酶、细胞角蛋白19片段、鳞状细胞癌抗原、胃泌素释放肽前体等；

b) 检验结果：标志物浓度数值、参考范围等。

8.6.4 气道炎症与过敏检验

健康医疗数据控制者应从LIS采集数据，包括：

a) 检验项目：诱导痰细胞学分类、血清总IgE、特异性IgE、嗜酸性粒细胞计数等；

b) 检验结果：诱导痰嗜酸性粒细胞比例、诱导痰中性粒细胞比例、IgE浓度、过敏原分级等。

8.6.5 其他检验项目

健康医疗数据控制者应从LIS采集其他通用检验数据，包括：

a) 检验项目：血常规、感染炎症指标、D-二聚体、自身免疫抗体谱、生化全项等；

b) 检验结果：白细胞计数、C反应蛋白数值、D-二聚体定量值、抗核抗体滴度、钾离子浓度、白蛋白水平等。

8.7 治疗相关信息

8.7.1 医学诊断

健康医疗数据控制者应从EMR或集成大数据平台采集医学诊断信息，包括：

a) 诊断类型：门诊诊断、入院诊断/入院初诊、主要诊断、病理诊断、出院诊断等；

b) 标准化诊断名称：

c) ICD-10编码。

8.7.2 用药

健康医疗数据控制者应从EMR或集成大数据平台采集用药信息，包括：

a) 用药名称；

b) 用药开始时间；
c) 给药途径：吸入，口服，注射等；
d) 用药频次；
e) 用药天数；
f) 用法用量。

8.7.3 呼吸介入

健康医疗数据控制者应从EMR、手术麻醉系统或集成大数据平台采集呼吸介入信息，包括：
a) 治疗时间；
b) 介入类别：经支气管镜介入、经皮穿刺介入、肺血管介入等；
c) 手术名称：支气管动脉栓塞术、气道支架置入术、球囊扩张术、热消融治疗、冷冻治疗、放射性粒子植入术、光动力治疗等；
d) 治疗部位：气管、支气管、肺血管、肺部病灶位置等；
e) 关键耗材：支架型号、栓塞剂类型、消融针型号等；
f) 治疗结果：手术成功标志、血管再通标志、气道狭窄改善程度等。

8.7.4 其他干预措施

健康医疗数据控制者应从EMR或集成大数据平台采集其他干预措施信息，包括：
a) 干预类型：氧疗、机械通气、呼吸康复、物理治疗等；
b) 干预参数：氧流量、通气模式、通气参数等；
c) 干预结果：临床疗效评估、不良反应等。

8.8 出院与随访

8.8.1 出院评估与转归

健康医疗数据控制者应从EMR或集成大数据平台采集出院评估与转归信息，包括：
a) 出院时间；
b) 出院转归：治愈、好转、未愈、死亡、其他；
c) 出院诊断：出院主要诊断、次要诊断及相应的ICD-10编码等；
d) 出院时状态：呼吸困难分级、静息状态脱氧后血氧饱和度、出院前最后一次生命体征等。

8.8.2 出院医嘱

健康医疗数据控制者应从EMR或集成大数据平台采集出院医嘱信息，包括：
a) 出院带药：药品名称、单次剂量、给药途径、用药频次、预计疗程等；
b) 家庭呼吸支持：长期家庭氧疗处方、家庭无创呼吸机治疗参数等；
c) 康复与生活指导：呼吸功能锻炼方式、气道廓清技术指导、戒烟干预、疫苗接种建议等。

8.8.3 随访记录

健康医疗数据控制者应从随访管理系统或集成大数据平台采集随访数据，包括：
a) 随访时间：实际随访日期、下次计划随访日期；
b) 随访方式：门诊随访、电话随访、互联网医院/线上随访、入户随访；
c) 病情控制评估：自上次就诊以来的急性加重次数、急诊就诊次数、慢阻肺评估测试评分、哮喘控制测试评分等；

d) 复查与检验记录：复查肺功能参数变化、复查胸部影像学变化标志、复查血气分析指标等；
e) 依从性与不良反应：用药依从性、吸入装置使用正确率评估、药物相关不良反应标志。

9 数据预处理

9.1 数据清洗

健康医疗数据控制者应对采集的原始数据进行清洗，包括：
a) 缺失数据识别与标记；
b) 重复数据识别与处理；
c) 异常数据识别与标记；
d) 逻辑矛盾数据识别与处理；
e) 无效数据识别与处理。

9.2 数据格式转换

健康医疗数据控制者应对采集数据进行格式转换和一致性处理，包括：
a) 日期型数据采用“YYYY-MM-DD”格式；
b) 日期时间型数据采用“YYYY-MM-DD HH:MM:SS”格式；
c) 数值型数据采用统一单位；
d) 枚举型数据值域统一。

10 数据标注

10.1 文本问答对构建

健康医疗数据控制者宜基于医学知识（如临床指南、专家共识、医学教科书）及上述数据元，构建文本问答对。问答对的构建宜覆盖核心临床场景，包括：
a) 疾病诊断；
b) 治疗决策；
c) 检查解读；
d) 预后与随访；
e) 紧急情况识别。
每个场景的问答对构建宜包含表1所列的核心要素，详细表述的示例见附录A。

表1 文本问答对核心临床场景及表述要素

临床场景	问题表述核心要素	答案表述核心要素
疾病诊断	1. 患者基本信息（年龄、性别等）； 2. 临床症状体征； 3. 关键检查/检验结果。	1. 明确的诊断结论或疾病分期/分组/分级等； 2. 诊断依据； 3. 所依据的临床指南或共识名称。

治疗决策	1. 确诊疾病及当前分期/状态； 2. 既往治疗史或特殊合并症（如有）。	1. 推荐的具体治疗方案（药物、操作等）； 2. 方案选择的依据； 3. 备选方案或注意事项。
检查解读	1. 检查项目名称； 2. 检查结果：具体的检测数值或影像描述。	1．指标临床意义； 2．对该数值/描述的定性判断（正常/异常程度）； 3．异常结果解读：该结果提示的病理生理意义或疑似疾病； 4．后续检查建议。
预后与随访	1.确诊疾病及严重程度/分期； 2.关键治疗史或风险因素。	1. 预后评估（如风险分层、预期转归）； 2. 生活方式与自我管理建议（如戒烟、康复训练、疫苗接种）； 3. 随访计划（包括监测指标、频率）。
紧急情况识别	1. 患者基础疾病及突发的症状/体征； 2. 初始急救干预的反应（如有）。	1. 紧急情况判定标准； 2. 就医/住院指征； 3. 紧急处理原则。

10.2 图像类数据标注

健康医疗数据控制者宜基于医学知识（如临床指南、专家共识、医学教科书），对胸部 CT 等影像数据开展标注，明确标注对象、形式、规则、属性及其计量单位和推荐值域等核心要求。通用标注规范和要求如下，呼吸内科常见病种的图像类数据标注示例见附录 B。

10.2.1 标注形式

健康医疗数据控制者开展图像类数据标注时，标注形式主要分为分割标注和标框标注。

a) 分割标注：健康医疗数据控制者宜采用轮廓勾勒方式对“肺气肿”对应的低密度病变区域、“胸腔积液”的实际累及范围等目标病灶或特定解剖器官边缘进行精准的像素水平勾画，以获取精确的形态、面积或体积信息；

b) 标框标注：健康医疗数据控制者宜使用矩形边界框对所有可疑肺结节、肿块、斑片影等目标病变区域进行标注，若影像中存在多个独立病灶（多发病灶），应按病灶分布逐一进行标注。

10.2.2 标注规则

健康医疗数据控制者开展图像类数据标注时，标注规则包括：

a) 覆盖全面：病灶类标注应覆盖所有受累区域，包括散在、微小病灶（直径≥1mm）；

b) 解剖精确：结构类标注应按解剖分级/分区精准勾画，确保无遗漏、无过度标注；

c) 若病灶内包含空洞、钙化等亚结构，应在主病灶轮廓内进行嵌套式独立勾画。

10.2.3 标注属性规范

健康医疗数据控制者开展图像类数据标注时，标注属性规范包括：

a) 计量单位：长度单位为厘米（cm）或毫米（mm），面积单位为平方厘米（cm^2）或平方毫米（mm^2），体积单位为立方厘米（cm^3）或立方毫米（mm^3）；

b) 推荐值域：各类标注参数的推荐值域详见附录 B，对于无预设值域的参数宜标注实际测量值，并按临床指南推荐标准判定分级。

10.3 音频类数据标注

健康医疗数据控制者宜基于医学知识（如临床指南、专家共识、医学教科书），对肺部音频数据开展标注，明确标注对象、形式、规则、属性及其计量单位和推荐值域等核心要求。通用标注规范和要求如下，呼吸内科常见病种音频类数据的标注示例见附录 C。

10.3.1 标注形式

健康医疗数据控制者开展音频类数据标注时，标注形式由基本标注形式和核心标注维度两部分构成。

a) 基本标注形式：分为事件标注和片段标注，事件标注用于定位瞬时发生的音频事件，如湿啰音的爆裂点，片段标注应标注异常声音事件的起止时间，用于框定有明确持续时长的音频事件；

b) 核心标注维度：指赋予上述形式的具体属性标签，包括分类标注、分级标注、特征标注、参数标注。

10.3.2 标注规则

健康医疗数据控制者开展音频类数据标注时，标注规则包括：

a) 时间精度：所有时间戳（起点、终点、事件点）的标记误差应≤0.1s；

b) 独立性：连续或散在出现的多个异常音频事件，进行逐一标注；

c) 干扰排除：标注过程中需排除说话、翻身、梦话等非目标干扰音频，避免干扰标注结果。

10.3.3 标注属性规范

健康医疗数据控制者开展音频类数据标注时，标注属性规范包括：

a) 计量单位：时间单位为秒（s）或分钟（min），响度单位为分贝（dB），频率单位为赫兹（Hz）或次/分钟；

b) 推荐值域：各类标注参数的推荐值域详见附录 C，对于无预设值域的参数宜标注实际测量值，并按临床指南推荐标准判定分级。

10.4 波形数据标注

健康医疗数据控制者宜基于医学知识（如临床指南、专家共识、医学教科书），对呼吸功能与通气相关的波形数据开展标注，明确标注对象、形式、规则、属性及其计量单位和推荐值域等核心要求。通用标注规范和要求如下，呼吸内科常见病种的波形数据标注示例见附录 D。

10.4.1 标注形式

健康医疗数据控制者开展波形类数据标注时，标注形式包括特征点标注、片段标注、参数计算标注、周期标注、趋势标注、对比标注。

10.4.2 标注规则

健康医疗数据控制者开展波形类数据标注时，标注规则包括：

a) 精度要求：特征点标注需精准定位波形顶点、谷底、拐点等（坐标误差≤0.01），片段或周期的时间戳误差应≤0.1s；

b) 一致性要求：进行对比标注时（如支气管舒张试验前后），必须确保两次测试的波形在时间或

坐标轴上严格对齐；

c) 质控要求：不合格或受干扰的异常波形，必须明确标注其干扰类型或异常原因（如咳嗽、漏气）。

10.4.3 标注属性规范

健康医疗数据控制者开展波形类数据标注时，标注属性规范包括：

a) 计量单位：容积使用升（L）或毫升（mL），流量使用升/秒（L/s）或升/分钟（L/min），压力使用毫米汞柱（mmHg）或厘米水柱（cmH_2O），血氧饱和度（SpO_2）及其他百分比类指标使用百分号（%）；

b) 推荐值域：各类标注参数的推荐值域详见附录 D，对于无预设值域的参数宜标注实际测量值，并按临床指南推荐标准判定分级。

附 录 A
（资料性）
文本问答对详细表述示例

A.1 以慢性阻塞性肺疾病为例，问答对具体描述见表A.1。

表 A.1 慢性阻塞性肺疾病（COPD）问答对

临床场景	问题表述	答案表述
疾病诊断	患者，男，68 岁，吸烟 40 年，CAT 评分 22 分，mMRC 3 级，过去一年有 2 次因急性加重住院，该患者的 GOLD 分组是什么？	根据 GOLD 2024 指南，该患者症状负担重(CAT≥10，mMRC≥2)，急性加重风险高(过去一年≥2 次中度急性加重或≥1 次住院事件)，属于 GOLD E 组。
治疗决策	对于评估为 GOLD E 组的 COPD 患者，推荐何种初始药物治疗方案？	E 组患者推荐的初始治疗采用双支扩剂（LAMA+LABA）。若患者血嗜酸性粒细胞计数≥300 cells/μL，可考虑初始使用 ICS+LAMA+LABA 三联疗法。方案选择需结合患者意愿、合并症及药物可及性等综合决定。
检查解读	肺功能报告中的“FEV1/FVC=0.65”和“FEV1 占预计值 45%”分别代表什么？	1．FEV1/FVC=0.65（<0.70）：是确诊气流受限的金标准，表明存在阻塞性通气功能障碍； 2．FEV1 占预计值 45%：表示患者第一秒用力呼气容积仅为健康同龄同体型人群平均值的 45%，属于 GOLD 3 级（重度）。
预后与随访	对于确诊为重度 COPD 的患者，经双支扩剂治疗后病情稳定，在生活管理和常规随访方面有哪些核心建议？	该患者预后需重点关注气流受限进展及急性加重风险，生活管理及随访方面建议患者： 1．立即并完全戒烟，这是延缓疾病进展最有效的措施； 2．定期接种流感和肺炎疫苗； 3．积极参与肺康复训练以改善活动耐力和生活质量； 4．遵医嘱规律用药，并定期随访，评估症状变化、急性加重风险和肺功能。
紧急情况识别	COPD 患者居家期间出现静息状态下呼吸困难突然加剧、口唇发绀，是否需立即就医，判定依据是什么？	1. 患者出现症状显著加重（静息呼吸困难加剧）、新的阳性体征（发绀），符合 COPD 急性加重的紧急就医标准； 2. 需立即就医并评估住院指征，若初始急诊治疗反应不佳则需进一步住院治疗

附 录 B
（资料性）
呼吸内科常见病种图像类数据标注示例

B.1 慢性气道疾病图像标注对象、形式、规则与属性规范，见表B.1。

表 B.1 慢性气道疾病图像标注对象、形式、规则及属性规范

标注对象	标注形式	标注规则	标注属性
肺气肿	分割标注	1. 标注范围为所有受累的肺实质区域； 2. 勾画密度低于特定阈值且无明显壁结构的低密度区域轮廓； 3. 基于分割结果计算整体肺气肿指数（EI）。	1. 病理类型： 小叶中心型、全小叶型、间隔旁型； 2. 优势分布： 上肺优势、下肺优势、弥漫性分布。
肺大疱	标框标注或分割标注	1. 标注直径大于 1cm、壁薄于 1mm 的局灶性含气囊腔； 2. 标注每个肺大疱，若多个肺大疱融合成片，按“融合性肺大疱”整体标注，不宜拆分； 3. 仅需定位时，宜采用标框标注，边界框包围囊腔。病灶形态清晰、边界明确时，宜采用分割标注囊壁轮廓。	1. 解剖定位： 遵循肺段解剖学命名规范，明确标注所在肺叶及肺段； 2. 大小：记录最大径（单位：mm）； 3. 类型：单个、多个、融合性； 4. 数量： 在病例级别统计肺大疱总数量。
气道壁增厚	分割标注	1. 标注范围为 3-5 级支气管； 2. 在选定支气管的横断面上，勾画气道内壁和外壁轮廓； 3. 基于分割结果计算管壁面积百分比（WA%）及管壁厚度（WT）。	1. 气道名称： 标注被测量支气管的解剖学名称； 2. 增厚程度： 根据 WA% 及 WT 数值，结合临床标准标注为“轻度”、“中度”、“重度”。
空气潴留	分割标注	1. 标注同一患者呼气相-吸气相配对的胸部 CT 图像，以吸气相 CT 作为参照标准； 2. 勾画相较于周围正常衰减的肺实质，密度衰减更少、呈现为马赛克样或片状的低密度区域。	1. 累及范围：以肺叶或肺段为单位，标注受影响的解剖范围（如“双下肺叶”、“左舌叶”）； 2. 严重程度：基于受累范围占相应肺叶/全肺的体积百分比，分为轻度、中度、重度。

B.2 肺部感染性疾病图像标注对象、形式、规则与属性规范，见表B.2。

表 B.2 肺部感染性疾病图像标注对象、形式、规则及属性规范

标注对象	标注形式	标注规则	标注属性

感染病灶	分割标注	1. 勾画病灶轮廓； 2. 区分单一病灶与多发病灶，多发病灶宜逐一独立标注，不宜遗漏散在或微小病灶； 3. 若病灶内包含空洞、钙化、支气管充气征等亚结构，宜在病灶轮廓内勾画亚结构范围。	1. 病灶类型：实变影、磨玻璃影、渗出影、结节状影、条索状影； 2. 解剖定位：遵循肺段解剖学命名规范，明确标注所在肺叶及肺段； 3. 密度特征：均匀密度、混杂密度； 4. 伴随征象：有无空洞（记录洞壁厚度）、支气管充气征、胸膜牵拉、胸腔积液。
空洞/空腔	分割标注	1. 标注病灶内出现的气液平面或无肺纹理的透亮区，并勾画空洞壁； 2. 对于厚壁空洞，宜勾画空洞腔和洞壁。	1. 空洞特征：厚壁、薄壁； 2. 内壁特征：光滑、不规则、结节状； 3. 内容物：是否存在液平、球状物（如真菌球）。
胸腔积液	分割标注	1. 勾画积液与肺组织、胸膜的分界轮廓； 2. 区分游离性积液与包裹性积液，包裹性积液宜单独标注包裹范围及胸膜增厚情况； 3. 基于分割结果计算积液的体积。	1. 分侧：左侧、右侧、双侧。 2. 积液量分级：少量、中量、大量； 3. 积液性质：游离性、包裹性； 4. 伴随改变：胸膜增厚、胸膜粘连、脓胸形成（若有）。

B.3 间质性肺病图像标注对象、形式、规则与属性规范，见表B.3。

表 B.3 间质性肺病图像标注对象、形式、规则及属性规范

标注对象	标注形式	标注规则	标注属性
网格影/蜂窝影	分割标注	1. 勾画呈网格状改变的肺间质区域及蜂窝样囊状气腔轮廓； 2. 分区勾画弥漫性网格影或多发蜂窝影。	1. 分布特征：胸膜下、基底部分布、弥漫性； 2. 囊腔大小：标注囊腔直径范围或平均直径（单位：mm）； 3. 严重程度：根据受累范围分为轻、中、重度。
牵拉性支气管扩张	分割标注	1. 勾画因周围肺实质纤维化牵拉而形态不规则、管径增粗的支气管轮廓； 2. 勾画宜沿扩张支气管的长轴进行，显示其形态不规则改变。	1. 解剖定位：标注所在肺段； 2. 扩张形态：串珠样、扭曲状、囊状； 3. 伴随征象：是否伴有蜂窝影、网格影。
磨玻璃影	分割标注	1. 勾画肺实质密度轻度增高但血管和支气管轮廓仍可辨认的区域。	1. 分布特征：胸膜下、沿支气管血管束、弥漫性、基底优势、上肺优势； 2. 伴随征象：小叶间隔增厚（铺路石征）、网格影、牵拉性支气

			管扩张。

B.4 肺结节与肿块图像标注对象、形式、规则与属性规范，见表B.4。

表 B.4 肺结节与肿块图像标注对象、形式、规则及属性规范

标注对象	标注形式	标注规则	标注属性
肺结节	分割标注	1. 标注直径≤3cm 的局灶性、类圆形、密度增高的肺部阴影，勾画结节的完整三维轮廓； 2. 对于部分实性结节，勾画实性成分和磨玻璃成分的轮廓； 3. 基于分割结果计算结节最大径、平均径、体积、球形度等参数。	1. 病灶类型：实性结节、部分实性结节、纯磨玻璃结节； 2. 解剖定位：遵循肺段解剖学命名规范，明确标注所在肺叶、肺段，并标注距胸膜、肺门的垂直距离（单位：mm）； 3. 形态与结构特征 •边缘状态：光滑、分叶状、毛刺状； •病灶形态：圆形、类圆形、不规则形； •内部结构：空洞、空泡征、支气管充气征、坏死区，钙化（需注明形态及分布）
肺肿块	分割标注	1. 标注直径>3cm 的肺部占位性病变，勾画肿块的实性成分、坏死囊变区； 2. 标注肿块是否侵犯邻近结构。	1. 大小与体积：记录三维径线并计算肿瘤体积； 2. 侵犯范围：胸膜、胸壁、纵隔、大血管等。
纵隔及肺门淋巴结	标框标注或分割标注	1. 依据解剖分区标注短径≥1cm、形态不规则、密度不均、强化异常或与病灶存在解剖引流关联的可疑淋巴结； 2. 若淋巴结融合成团，按单一融合病灶整体标注，明确融合范围及涉及的淋巴结分区，计算融合病灶最大径。	1. 淋巴结分区：按 IASLC 标准标注具体分区（如 1R、2L、4R、7 区等）； 2. 大小：记录淋巴结短径（单位：mm），融合淋巴结需记录最大径及涉及分区数量； 3. 形态特征：形态是否规则、有无坏死、钙化。

B.5 睡眠呼吸障碍图像标注对象、形式、规则与属性规范，见表B.5。

表 B.5 睡眠呼吸障碍图像标注对象、形式、规则及属性规范

标注对象	标注形式	标注规则	标注属性

上气道腔	分割标注	1. 沿上气道内壁（鼻咽、口咽、喉咽）进行勾画，勾画气道内壁轮廓； 2. 测量狭窄部位的前后径、左右径及截面积，测量对应层面的气道壁厚度； 3. 若存在多部位狭窄，分别标注各狭窄部位，明确主次狭窄部位。	1. 狭窄部位：鼻腔、鼻咽部、口咽部（软腭后区、舌后区）、喉咽部、混合型狭窄； 2. 狭窄程度：按气道截面积或前后径/左右径比值分为轻度、中度、重度。
上气道周围软组织	分割标注	1. 勾画与气道塌陷相关的关键软组织轮廓； 2. 基于分割结果计算软组织的体积及评估对气道的侵占程度。	1. 软组织名称：软腭、舌体、悬雍垂、咽旁脂肪垫； 2. 体积参数：软组织体积（单位：cm^3）及占颈部总体积的百分比； 3. 压迫程度：无压迫、轻度压迫、中度压迫、重度压迫。

附 录 C
（资料性）
呼吸内科常见病种音频类数据标注示例

C.1 慢性气道疾病音频标注对象、形式、规则与属性规范，见表C.1。

表 C.1 慢性阻塞性肺疾病（COPD）音频标注对象、形式、规则及属性规范

标注对象	标注形式	标注规则	标注属性
呼气延长音	片段标注 + 参数计算标注	1. 标注呼气阶段持续时间超过正常参考范围（成人平静呼吸呼气时长≥2秒）的音频片段，明确起止时间戳； 2. 结合吸气音片段，计算呼气/吸气时间比（E/I 比）； 3. 多段呼气延长音宜分别独立标注，标注时需排除咳嗽、说话等干扰音频。	1. 持续时长； 2. 呼气/吸气时间比； 3. 强度分级（1-4 级）； 4. 听诊部位（前胸/侧胸/后背分区）。
哮鸣音	事件标注 + 分级标注	1. 标注音频中出现的高调、连续性乐音样音频事件，精准标记起止时刻； 2. 按响度分为 1-4 级（1 级仅安静环境可闻，4 级无需听诊器即可闻及）； 3. 区分单频哮鸣音与多频哮鸣音，多频哮鸣音需标注主频率区间。	1. 频率范围（Hz）； 2. 强度分级； 3. 持续时长； 4. 分布特征（单侧/双侧、局部/弥漫）。
湿啰音	事件标注 + 分类标注	1. 标注断续性水泡音或爆裂音事件，按出现时相分为吸气相湿啰音、呼气相湿啰音及双相湿啰音；2. 区分粗湿啰音（气管/主支气管来源）、中湿啰音（中等支气管来源）、细湿啰音（细支气管来源）；3. 散在湿啰音需逐一标注，密集湿啰音可标注核心区域及范围覆盖度	1. 啰音类型； 2. 出现时相； 3. 强度； 4. 听诊部位； 5. 密集程度。
咳嗽音	片段标注 + 特征标注	1. 标注完整咳嗽周期（吸气 - 爆发性呼气 - 收尾）的音频片段，排除清嗓、喷嚏等类似声音； 2. 区分干咳与湿咳（湿咳音频含气泡音或痰鸣音特征）； 3. 记录单位时间内（如 1 分钟）咳嗽次数及单次咳嗽持续时长。	1. 咳嗽类型（干咳/湿咳）； 2. 单次持续时长； 3. 频率（次/分钟）； 4. 强度分级。

C.2 哮喘音频标注对象、形式、规则与属性规范，见表C.2。

表 C.2 哮喘音频标注对象、形式、规则及属性规范

标注对象	标注形式	标注规则	标注属性
哮鸣音	事件标注 + 分类标注	1. 区分吸气相哮鸣音、呼气相哮鸣音及双相哮鸣音，优先标注呼气相主导哮鸣音； 2. 按音调高低分为高调哮鸣音（>400Hz）与低调哮鸣音（≤400Hz）； 3. 发作期与缓解期哮鸣音需分别标注，记录缓解前后变化。	1. 时相分布； 2. 音调分类； 3. 强度； 4. 出现场景（静息/活动/诱发试验）； 5. 缓解响应（对支气管扩张剂是否敏感）。
咳嗽音	片段标注 + 关联标注	1. 标注哮喘相关咳嗽片段，区分刺激性干咳与夜间阵发性咳嗽； 2. 标注咳嗽与喘息、哮鸣音的时间关联性（如咳嗽后出现喘息）； 3. 记录咳嗽诱发因素（如冷空气、运动、异味）对应的音频片段。	1. 咳嗽类型； 2. 诱发因素； 3. 与喘息/哮鸣音的关联关系； 4. 频率（次/小时）； 5. 持续时长。
呼吸困难相关音频	片段标注 + 特征标注	1. 标注呼吸频率加快（成人静息状态 ＞20 次/分钟）、呼吸加深对应的音频片段； 2. 识别并标注 “三凹征” 相关呼吸音（胸骨上窝、锁骨上窝、肋间隙凹陷时的异常呼吸音）； 3. 结合呼吸节律，标注节律不齐、叹息样呼吸等异常模式	1. 呼吸频率（次/分钟）； 2. 呼吸模式（浅快/深慢/节律不齐）； 3. 异常呼吸征（有无三凹征）； 4. 持续时长； 5. 严重程度分级。

C.3 睡眠呼吸暂停综合征（SAS）音频标注对象、形式、规则与属性规范，见表C.3。

表 C.3 睡眠呼吸暂停综合征（SAS）音频标注对象、形式、规则及属性规范

标注对象	标注形式	标注规则	标注属性
呼吸暂停段	片段标注 + 验证标注	1. 标注睡眠中呼吸音完全消失且持续时间≥10 秒的音频片段，明确起止时间戳； 2. 结合胸腹运动信号（若有）辅助验证，区分中枢性（无胸腹运动）与阻塞性（有胸腹运动）呼吸暂停； 3. 连续出现的呼吸暂停段需逐一独立标注，记录间隔时间。	1. 类型（中枢性/阻塞性/混合性）； 2. 持续时长； 3. 发生时段（浅睡眠/深睡眠/快速眼动睡眠）； 4. 间隔时间（与前一次暂停的时间差）。
低通气段	片段标注 + 参数计算标注	1. 标注呼吸音强度较基线降低≥50%、持续时间≥10 秒的音频片段； 2. 计算低通气段的呼吸强度降低比例及持续时长； 3. 区分单一低通气与连续低通气，连续低通气需标注总时长及片段数	1. 呼吸强度降低比例（%）； 2. 持续时长； 3. 发生时段； 4. 连续片段数（若为连续低通气）

鼾声	事件标注 + 分类标注	1. 标注睡眠中出现的规律性鼾声事件，标记起止时刻，排除梦话、翻身等干扰声音； 2. 按响度分为轻度（≤60dB）、中度（61-80dB）、重度（>80dB）； 3. 区分规则鼾声与不规则鼾声，不规则鼾声需标注频率波动特征。	1. 响度（dB）； 2. 类型（规则/不规则）； 3. 频率范围（Hz）； 4. 持续时长； 5. 发生时段。
呼吸努力相关音频	事件标注 + 特征标注	1. 标注呼吸暂停/低通气前后出现的用力呼吸音、喉鸣音等音频事件； 2. 识别并标注觉醒相关音频（如突然的呼吸加深、肢体活动伴随音）； 3. 记录呼吸努力音频与呼吸暂停/低通气的时间关联（如暂停后即刻出现）。	1. 音频类型（用力呼吸音/喉鸣音/觉醒相关音）； 2. 持续时长； 3. 与呼吸暂停/低通气的关联关系； 4. 强度分级。

C.4 肺部感染性疾病音频标注对象、形式、规则与属性规范，见表C.4。

表 C.4 肺部感染性疾病音频标注对象、形式、规则及属性规范

标注对象	标注形式	标注规则	标注属性
湿啰音	事件标注 + 分类标注	1. 标注感染相关的渗出性湿啰音，按出现时相分为吸气早期、吸气中期、吸气晚期湿啰音； 2. 区分粗、中、细湿啰音，细湿啰音（捻发音）需重点标注（提示肺泡内渗出）； 3. 弥漫性湿啰音需标注覆盖区域，局限性湿啰音明确解剖定位。	1. 啰音分级（粗/中/细）； 2. 出现时相； 3. 分布范围（弥漫性/局限性）； 4. 听诊部位； 5. 强度。
干啰音	事件标注 + 特征标注	1. 标注气道炎症或分泌物阻塞导致的干啰音，区分高调干啰音（哮鸣音）与低调干啰音（鼾音）； 2. 记录干啰音的连续性（持续/间断），间断性干啰音需标注出现频率； 3. 排除哮喘、COPD 等慢性疾病基础干啰音，重点标注感染新发或加重的干啰音。	1. 音调分类（高调/低调）； 2. 连续性； 3. 出现频率（次/分钟）； 4. 分布部位； 5. 与感染的关联（新发/加重）。
胸膜摩擦音	事件标注 + 验证标注	1. 标注吸气相和呼气相均可闻及的粗糙、搔抓样音频事件，明确起止时刻；2. 结合胸痛相关临床信息（若有）辅助验证，标注摩擦音与呼吸运动的同步性；3. 区分局限性与弥漫性胸膜摩擦音，明确听诊部位。	1. 出现时相（吸气相/呼气相/双相）；2. 分布范围（局限性/弥漫性）；3. 听诊部位；4. 强度；5. 与呼吸运动的同步性。

C.5 间质性肺病音频标注对象、形式、规则与属性规范，见表C.5。

表 C.5 间质性肺病音频标注对象、形式、规则及属性规范

标注对象	标注形式	标注规则	标注属性
Velcro 啰音	事件标注 + 分类标注	1. 标注吸气晚期出现的细湿啰音，呈捻发音或 “Velcro 尼龙粘扣撕开” 样特征，精准标记起止时刻； 2. 区分局限性与弥漫性 Velcro 啰音，弥漫性需标注肺叶分区分布； 3. 按强度分为 1-3 级（1 级仅深吸气末可闻，3 级平静吸气即可闻及）。	1. 出现时相（吸气早/中/晚期）； 2. 分布范围（局限性/弥漫性）； 3. 听诊部位（肺叶分区）； 4. 强度分级； 5. 持续时长。
干啰音	事件标注 + 特征标注	1. 标注气道纤维化或狭窄导致的高调干性啰音，排除感染、哮喘等其他原因所致干啰音； 2. 记录干啰音的频率稳定性（固定/波动），固定频率干啰音需标注主频率； 3. 多部位干性啰音需按听诊区域分别标注，明确主导部位。	1. 主频率（Hz）； 2. 频率稳定性（固定/波动）； 3. 分布部位； 4. 强度； 5. 持续时长。
限制性呼吸相关音频	片段标注 + 特征标注	1. 标注呼吸浅快（成人静息呼吸频率 > 20 次/分钟、潮气量降低）对应的音频片段； 2. 识别并标注叹息样呼吸、呼吸节律不齐等限制性呼吸模式； 3. 记录单位时间内呼吸次数及单次呼吸持续时长，计算呼吸频率变异系数。	1. 呼吸频率（次/分钟）； 2. 呼吸模式（浅快/叹息样/节律不齐）； 3. 单次呼吸持续时长； 4. 频率变异系数； 5. 严重程度分级。

附 录 D
（资料性）
呼吸内科常见病种波形类数据标注示例

D.1 慢性气道疾病波形数据标注对象、形式、规则与属性规范，见表D.1。

表 D.1 慢性阻塞性肺疾病（COPD）波形数据标注对象、形式、规则及属性规范

标注对象	标注形式	标注规则	标注属性
呼气流量 - 容积（F-V）曲线	特征点标注 + 参数计算标注	1. 标注曲线中峰值流量（PEF）对应的波形顶点、呼气 1 秒末流量（FEV1）对应时间点及曲线下降支形态； 2. 需确保波形为合格测试（呼气时间≥6 秒、无明显漏气或咳嗽干扰），不合格波形需标注干扰类型； 3. 基于波形计算 FEV1、用力肺活量（FVC）、FEV1/FVC 比值，若 FEV1/FVC＜0.7 则标注为阻塞性特征波形。	1. FEV1 数值（L）及占预计值百分比（% pred）； 2. FVC 数值（L）及占预计值百分比（% pred）； 3. FEV1/FVC 比值； 4. 曲线形态（低平型/平台型/正常衰减型）； 5. GOLD 分级（1-4 级）。
肺活量（VC）波形	片段标注 + 参数标注	1. 标注平静吸气末、深吸气末、深呼气末对应的波形节点，区分慢肺活量（SVC）与用力肺活量（FVC）波形； 2. 排除吸气中断、呼气不完全导致的异常波形，标注有效波形的容积变化范围； 3. 计算肺活量（VC）、补吸气量（IRV）、补呼气量（ERV）等参数。	1. VC 数值（L）及 % pred； 2. 波形完整性（完全/不完全）； 3. 通气功能类型（单纯阻塞性/混合性）； 4. 测试有效性（合格/不合格及原因）。
最大通气量（MVV）波形	周期标注 + 趋势标注	1. 标注 12 秒或 15 秒连续呼吸周期的波形片段，识别呼吸频率、潮气量的稳定性； 2. 标注波形中是否存在呼吸节律不齐、流量波动过大等异常特征； 3. 基于波形计算 MVV 数值，若＜80% pred 则标注为通气储备功能下降。	1. MVV 数值（L/min）及 % pred； 2. 呼吸频率（次/分钟）； 3. 平均潮气量（mL）； 4. 波形稳定性（稳定/不稳定）。

支气管舒张试验前后 F-V 曲线	对比标注 + 参数计算标注	1. 分别标注舒张前（基础状态）和舒张后（吸入支气管扩张剂 15-20 分钟后）的 F-V 曲线，明确两次测试的波形对应关系； 2. 计算舒张后 FEV1 改善率=[（舒张后 FEV1 - 舒张前 FEV1）/ 舒张前 FEV1]×100%，改善率≥12% 且绝对值增加≥200mL 则标注为阳性波形； 3. 标注曲线中 PEF、FEV1 的变化趋势，排除测试干扰因素（如用药不规范、咳嗽）。	1. 舒张前/后 FEV1 数值及 % pred； 2. FEV1 改善率（%）； 3. 试验结果（阳性/阴性/可疑）； 4. 波形变化特征（流量提升/曲线形态改善程度）。

D.2 哮喘波形数据标注对象、形式、规则与属性规范，见表D.2。

表 D.2 哮喘波形标注对象、形式、规则及属性规范

标注对象	标注形式	标注规则	标注属性
支气管舒张试验前后 F-V 曲线	对比标注 + 参数计算标注	1. 分别标注舒张前（基础状态）和舒张后（吸入支气管扩张剂 15-20 分钟后）的 F-V 曲线，明确两次测试的波形对应关系； 2. 计算舒张后 FEV1 改善率=[（舒张后 FEV1 - 舒张前 FEV1）/舒张前 FEV1]×100%，改善率≥12% 且绝对值增加≥200mL 则标注为阳性波形； 3. 标注曲线中 PEF、FEV1 的变化趋势，排除测试干扰因素（如用药不规范、咳嗽）。	1. 舒张前/后 FEV1 数值及 % pred； 2. FEV1 改善率（%）； 3. 试验结果（阳性/阴性/可疑）； 4. 波形变化特征（流量提升/曲线形态改善程度）。
支气管激发试验剂量 - 反应波形	分级标注 + 拐点标注	1. 按激发剂浓度梯度（如乙酰甲胆碱 0.03-16mg/mL）标注各浓度下的 F-V 曲线，识别 FEV1 下降节点； 2. 当 FEV1 较基础值下降≥20% 时，标注对应的激发剂浓度（PC20）及波形拐点； 3. 标注试验过程中是否出现喘息、咳嗽对应的异常波形片段。	1. 基础 FEV1 及各浓度下 FEV1 数值； 2. PC20 值（mg/mL）； 3. 试验结果（阳性/阴性/临界）； 4. 不良反应相关波形（有/无及类型）。
峰流速（PEF）日变异率波形	趋势标注 + 区间标注	1. 标注 24 小时内多次测量的 PEF 波形，明确每日最高值、最低值对应的时间点； 2. 计算 PEF 日变异率 =[（每日最高 PEF - 每日最低 PEF）/（每日最高 PEF + 每日最低 PEF）/2]×100%，变异率≥20% 则标注为哮喘特征性波形； 3. 标注波形中是否存在晨起/夜间低谷的节律特征。	1. 每日最高/最低 PEF 数值（L/min）； 2. PEF 日变异率（%）； 3. 节律特征（晨起低谷型/夜间低谷型/无规律）； 4. 控制水平（良好/不佳）。

D.3 呼吸治疗（呼吸机/无创通气）波形数据标注对象、形式、规则与属性规范，见表D.3。

表 D.3 呼吸治疗（呼吸机/无创通气）波形标注对象、形式、规则及属性规范

标注对象	标注形式	标注规则	标注属性
吸气气流波形	形态标注 + 参数计算标注	1. 标注吸气相气流的模式（方波/递减波/递增波）、峰值点及持续时间； 2. 计算吸气峰流速（PIF），若 PIF ＞60L/min（成人），标注为高流速风险波形（可能导致气道损伤）； 3. 区分呼吸机控制气流与患者自主吸气气流叠加特征，排除咳嗽导致的气流骤升。	1. 气流模式（方波/递减波/递增波/混合型）； 2. PIF（L/min）； 3. 吸气气流持续时长（秒）； 4. 流速稳定性（稳定/波动）； 5. 高流速风险（是/否）。
呼气气流波形	片段标注 + 完整性标注	1. 标注呼气相气流下降段、基线回归点，判断呼气是否完全（呼气末气流回归基线且持续≥0.3 秒）； 2. 若呼气末气流未回归基线（残留气流＞5L/min），标注为呼气不完全波形，提示 Auto-PEEP 风险； 3. 区分被动呼气与主动呼气（主动呼气表现为气流骤升），标注主动呼气相关片段。	1. 呼气峰流速（PEF，L/min）； 2. 呼气持续时长（秒）； 3. 呼气完整性（完全/不完全）； 4. 残留气流速度（L/min）； 5. 呼气类型（被动/主动）。
漏气气流波形	异常段标注 + 量化标注	1. 标注通气周期中未参与肺内气体交换的气流片段（如无创通气面罩边缘漏气、气管插管气囊漏气）； 2. 计算漏气量（Leak）= 吸入潮气量（VTin）- 呼出潮气量（VTexp），若 Leak＞30% VTin，标注为中度漏气；＞50% VTin 为重度漏气； 3. 区分持续性漏气与间歇性漏气，间歇性漏气需标注发作频率及持续时长。	1. 漏气量（mL）； 2. 漏气占比（% VTin）； 3. 漏气类型（持续/间歇）； 4. 漏气发作频率（次/分钟）； 5. 通气影响（无影响/轻度影响/严重影响）。

吸气容积-时间曲线	形态标注+导数标注	1. 标注曲线的上升斜率（反映吸气流速变化）、平台期（若有），区分线性上升（方波气流）与非线性上升（递减波气流）； 2. 计算曲线上升斜率（ΔV/Δt），若斜率＜100mL/秒，提示吸气流速不足；＞300mL/秒，提示流速过快； 3. 识别曲线中是否存在“停滞段”（提示吸气中断或人机不同步）。	1. 上升斜率（mL/秒）； 2. 曲线形态（线性/非线性/停滞型）； 3. 吸气中断情况（有/无及持续时长）； 4. 流速适配性（适配/不足/过快）。
呼吸周期同步波形	对比标注+相位标注	1. 同步标注患者自主呼吸节律（如胸腹运动波形、膈肌电信号波形）与呼吸机通气周期波形，识别吸气相、呼气相的时间匹配度； 2. 若呼吸机吸气相启动早于患者自主吸气（强制通气），标注为“通气超前”；晚于患者自主吸气＞200ms，标注为“通气滞后”； 3. 计算呼吸周期匹配率（同步周期数/总周期数×100%），＜80%为同步不良。	1. 呼吸周期匹配率（%）； 2. 同步性分级（良好/一般/不良）； 3. 同步异常类型（超前/滞后/无同步）； 4. 异常周期占比（%）； 5. 患者耐受性（良好/轻度不适/严重不适）。

D.4 间质性肺病波形数据标注对象、形式、规则与属性规范，见表D.4。

表 D.4 肺纤维化（间质性肺病）波形标注对象、形式、规则及属性规范

标注对象	标注形式	标注规则	标注属性
肺总量（TLC）容积曲线	特征标注+参数计算标注	1. 标注深吸气至肺总量、深呼气至残气量（RV）对应的波形顶点与谷底，明确曲线的容积变化范围； 2. 计算TLC、RV、功能残气量（FRC）数值，若TLC＜80% pred则标注为限制性通气功能障碍波形； 3. 标注曲线形态是否呈现“陡升陡降”（肺弹性减退特征），排除呼气不完全干扰。	1. TLC数值（L）及%pred； 2. RV/TLC比值； 3. 通气功能类型（单纯限制性/混合性）； 4. 曲线形态特征（弹性减退型/正常型）。

弥散功能（DLCO）波形	峰值标注 + 斜率标注	1. 标注一氧化碳弥散过程中肺泡气 - 肺毛细血管血液气体交换对应的波形，识别弥散峰值及曲线下降斜率； 2. 计算 DLCO 数值及占预计值百分比，若 DLCO＜80% pred 则标注为弥散功能降低波形； 3. 结合肺泡容积（VA），计算 DLCO/VA 比值，排除 VA 降低对 DLCO 的影响。	1. DLCO 数值（mL/min/mmHg）及 % pred； 2. VA 数值（L）及 % pred； 3. DLCO/VA 比值； 4. 弥散功能分级（正常/轻度降低/中度降低/重度降低）。

D.5 睡眠呼吸暂停综合征（SAS）波形数据标注对象、形式、规则与属性规范，见表D.5。

表 D.5 睡眠呼吸暂停综合征（SAS）波形标注对象、形式、规则及属性规范

标注对象	标注形式	标注规则	标注属性
呼吸努力波形（胸/腹运动）	片段标注 + 类型标注	1. 标注睡眠中呼吸努力消失（中枢性）或存在（阻塞性）的波形片段，明确起止时间戳； 2. 呼吸努力消失且持续≥10 秒，标注为中枢性呼吸暂停波形；呼吸努力存在但气道气流消失≥10 秒，标注为阻塞性呼吸暂停波形； 3. 标注波形中呼吸努力强度（振幅）的变化，区分轻度、中度、重度努力增强。	1. 事件类型（中枢性暂停/阻塞性暂停/混合性暂停）； 2. 持续时长（秒）； 3. 呼吸努力强度（轻度/中度/重度）； 4. 发生睡眠分期（N1/N2/N3/REM）。
血氧饱和度（SpO_2）波形	下降段标注 + 参数标注	1. 标注 SpO_2较基线下降≥3% 且持续≥10 秒的波形片段，明确下降起始点、最低点及恢复点； 2. 计算血氧饱和度下降幅度（ΔSpO_2）及最低 SpO_2值，若最低 SpO_2＜90% 则标注为低氧血症相关波形； 3. 关联呼吸暂停/低通气事件，标注 SpO_2下降与呼吸事件的时间对应关系。	1. 基线 SpO_2（%）； 2. 最低 SpO_2（%）； 3. 下降幅度（ΔSpO_2，%）； 4. 低氧持续时长（秒）； 5. 关联呼吸事件类型。

气道压力/气流波形	异常段标注 + 特征标注	1. 标注气道气流完全消失（暂停）或幅度降低≥50%（低通气）的波形片段，持续≥10 秒为有效事件； 2. 区分规则气流波形与不规则气流波形（如打鼾相关的锯齿波），标注打鼾对应的气流波动特征； 3. 若为 CPAP 治疗中监测，标注治疗压力下的气流稳定性（是否仍存在暂停/低通气）。	1. 事件类型（呼吸暂停/低通气）； 2. 持续时长（秒）； 3. 气流形态（规则/不规则/锯齿波）； 4. 治疗响应（CPAP 治疗时有效/无效）。

参 考 文 献

[1] 基于深度学习的电子病历多模态数据融合研究进展

[2] 面向医学文本处理的医学实体标注规范

[3] 面向专病电子病历的实体语料库构建方法

[4] 面向呼吸内科智能诊断模型研究

[5] 慢性阻塞性肺疾病胸部 CT 检查及评价中国专家共识

[6] 临床实践指南制订过程中临床问题的构建—基于病例引导式方法

[7] 面向医疗领域智能问答系统的问题文本聚类研究

[8] Diagnosis and Treatment of Adults with Community-acquired Pneumonia. An Official Clinical Practice Guideline of the American Thoracic Society and Infectious Diseases Society of America

[9] Diagnosis of Idiopathic Pulmonary Fibrosis. An Official ATS/ERS/JRS/ALAT Clinical Practice Guideline

[10] Guidelines for Management of Incidental Pulmonary Nodules Detected on CT Images

[11] Three-dimensional imaging of the upper airway anatomy in obstructive sleep apnea: a systematic review