ICS 35.020
CCS L70

团 体 标 准

T/ISC 0155—2026

医疗健康行业智能体 淋巴瘤辅助诊疗智能体技术要求

Technical requirements for Lymphoma diagnosis and treatment intelligent agent in the healthcare industry

（发布稿）

2026－10－28 发布 2026－11－28 实施

中国互联网协会 发 布

目　　次

前　　言 II
1 范围 1
2 规范性引用文件 1
3 术语和定义 1
4 缩略语 1
5 总体要求 1
6 通用技术指标要求 2
6.1 分类任务技术指标 2
6.2 图像分割任务技术指标 3
6.3 生成类任务技术指标 4
6.4 检索与排序技术指标 4
7 辅助诊断要求 5
7.1 临床表现分析 5
7.2 实验室检查分析 5
7.3 病理学检查分析 6
7.4 影像学检查分析 7
8 辅助治疗要求 7
8.1 综合分期分型 7
8.2 治疗方案选择 8
8.3 信息提醒 9
9 易用性要求 10
9.1 可理解性 11
9.2 易学性 11
9.3 易操作性 11
10 安全性要求 11
10.1 基础设施安全 11
10.2 数据安全 12
10.3 应用安全 12

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。

本文件由中国互联网协会归口。

本文件起草单位：上海交通大学医学院附属瑞金医院、上海罗氏制药有限公司、中国信息通信研究院、北京大学人民医院、重庆新桥医院、宁波大学附属人民医院、哈尔滨血液病肿瘤研究所、苏州大学附属第一医院、安徽省立医院、福建医科大学附属协和医院、南昌大学第一附属医院、中国医学科学院血液病医院、江苏省人民医院、四川大学华西医院、西安交通大学第一附属医院、广西医科大学第一附属医院、山东大学齐鲁医院、青岛大学附属医院、贵黔国际医院、河南省人民医院、郑州大学第一附属医院、NEJM AI、瑞金医院医学影像先进技术研究院、中国移动上海公司战略客户部、医渡云（北京）技术有限公司、北京深睿博联科技有限责任公司、卫宁健康科技集团股份有限公司、北京医准智能科技有限公司、数坤科技股份有限公司、沐心数智医疗科技(北京)有限公司、上海联影医疗高新技术研究院有限公司、工业互联网创新中心（上海）有限公司

本文件主要起草人：赵维莅、许彭鹏、王黎、顾于蓓、杜雨暄、任九选、贾斐、张卓然、相识、刘泊宁、黄涂半特、陈冠衣、张晓辉、张曦、陆滢、赵东陆、徐杨、朱小玉、徐兵、李菲、易树华、范磊、牛挺、贺鹏程、罗军、彭军、王伟、王季石、朱尊民、宋永平、赵剑飞、温宁、顾洪飞、徐济铭、李林峰、徐悦、李一鸣、余琛琳、张番栋、饶立伟、于克军、郝中华、赵英、胡阳、丁佳、李亚玲、齐立薇、方金武、王道洋、胡建颖。

医疗健康行业智能体 淋巴瘤辅助诊疗智能体技术要求

1 范围

本文件规定淋巴瘤辅助诊疗智能体的相关功能指标和应用规范，包括诊断辅助要求和治疗辅助要求两部分。

本文件适用于医疗健康行业智能体技术方开发淋巴瘤辅助诊疗智能体时进行系统建设和应用范式参考，以及医疗机构等应用方参照进行淋巴瘤辅助诊疗智能体选型。

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。

3 术语和定义

下列术语和定义适用于本文件。

3.1

智能体 ai agent

又称人工智能代理，是指驻留在某一环境下，能持续自主地发挥作用，具备驻留性、反应性、社会性、主动性等特征的计算实体。

3.2

医疗健康行业智能体 healthcare ai agent

在通用智能体的基础上，结合医疗健康行业特点设计的智能体，与医疗健康相关任务的适配度较高。

3.3

金标准 gold standard

当前临床医学界公认最可靠、最权威的诊断方法。

4 缩略语

下列缩略语适用于本文件。

AUC：曲线下面积（Area Under the Curve）

FN：假阴性（False Negative）

FP：假阳性（False Positive）

ICD-10：国际疾病分类第十版（International Classification of Disease, 10th Revision）

IPI：国际预后指数（International Prognostic Index）

IoU：交并比（Intersection over Union）

PET-CT：正电子发射计算机断层显像（Positron Emission Tomography – Computed Tomography）

ROC：接受者操作特性曲线（Receiver Operating Characteristic）

SNOMED CT：医学系统命名法——临床术语（Systematized Nomenclature of Medicine – Clinical Terms）

TN：真阴性（True Negative）

TP：真阳性（True Positive）

WHO：世界卫生组织（World Health Organization）

5 总体要求

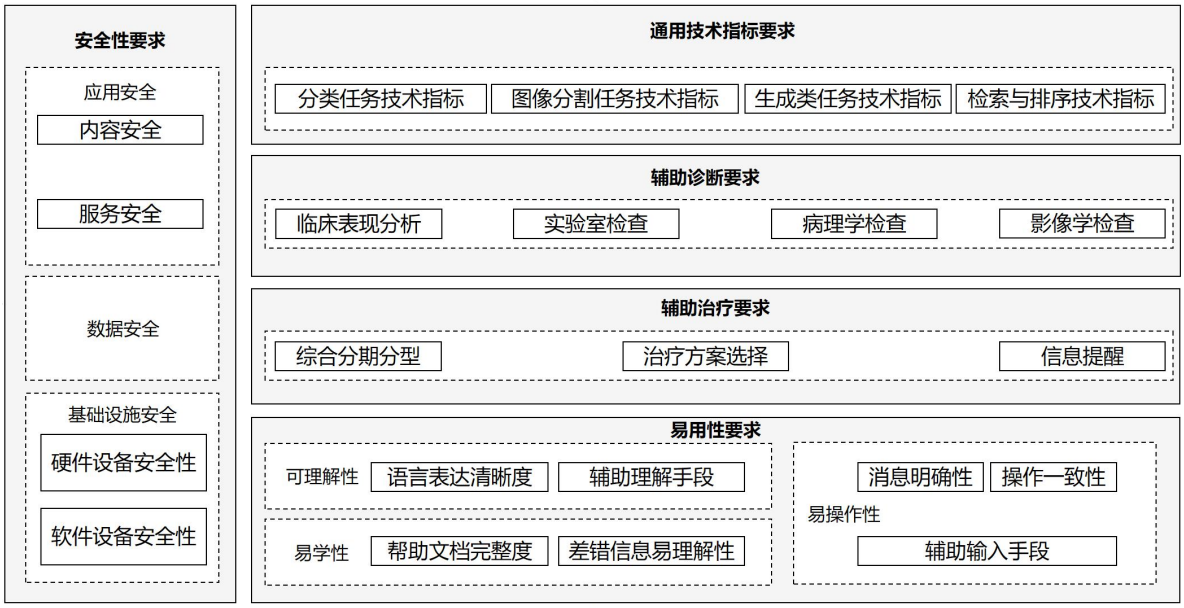

图 1　淋巴瘤辅助诊疗智能体的架构示意图

淋巴瘤辅助诊疗智能体应具备辅助诊断与辅助治疗两大核心能力。通过整合患者的临床表现、实验室检查、病理学检查与影像学检查等多模态数据，智能体应支持涵盖从综合分期分型、治疗方案选择到随访提醒的连续决策环节，为临床医生提供精准、高效的决策支持，提升淋巴瘤诊疗的规范性与个体化水平。

淋巴瘤辅助诊疗智能体的开发与应用应满足以下方面的要求：辅助诊断功能要求、辅助治疗功能要求、通用技术指标要求、易用性要求及安全性要求。

6　通用技术指标要求

6.1　分类任务技术指标

本类指标适用于所有判别性任务（如命名实体识别、异常项判定、病灶性质分类等），即模型将输入数据（如文本、图像等）划分到预定类别中的能力：

a）　准确率：针对所有分类任务，计算模型所有预测中正确预测的总体比例，计算公式如下：

$$Accuracy = \frac{TP+TN}{TP+TN+FP+FN}$$

式中：

Accuracy——准确率；

TP——真阳性数量；

TN——真阴性数量；

FP——假阳性数量；

FN——假阴性数量。

b）　精确率：针对二分类任务，计算模型预测为正例的结果中，实际为正例的比例，计算公式如下：

$$Precision = \frac{TP}{TP+FP}$$

式中：

Precision——精确率；

TP——真阳性数量；

FP——假阳性数量。

c）　召回率（灵敏度）：针对二分类任务，计算所有实际为正例的样本中，被模型正确识别为正例的比例，计算公式如下：

$$Recall = \frac{TP}{TP+FN}$$

式中：

Recall——召回率；

TP——真阳性数量；

FN——假阴性数量。

d） 特异度：针对二分类任务，计算模型正确预测为负例的样本数占实际负例总数的比例，计算公式如下：

$$Specificity = \frac{TN}{TN+FP}$$

式中：

Specificity——特异度；

TN——真阴性数量；

FP——假阳性数量。

e） F1 分数：针对二分类任务，计算精确率与召回率的调和平均数，计算公式如下：

$$F1 = 2 \times \frac{Precision \times Recall}{Precision + Recall}$$

F1——F1 分数；

Precision——精确率；

Recall——召回率。

f） ROC-AUC：针对二分类任务，计算 ROC 曲线下的面积 AUC，计算公式如下：

$$AUC = \int_0^1 TPR\, dFPR$$

式中：

TPR——真阳性率，即召回率；

FPR——假阳性率，表示所有实际为负的样本中，被错误预测为正的比例，计算公式如下：

$$FPR = \frac{FP}{FP+TN}$$

g） 宏精确率：针对多分类任务，计算每个类别的精确率再求算数平均，计算公式如下：

$$Macro - P = \frac{1}{K}\sum_{i=1}^{K} Precision_i$$

式中：

Macro-P——宏精确率；

K——分类任务总类别数；

$Precision_i$——第i类精准率。

h） 宏召回率：针对多分类任务，计算每个类别的召回率再求算数平均，计算公式如下：

$$Macro - R = \frac{1}{K}\sum_{i=1}^{K} Recall_i$$

式中：

Macro-R——宏召回率；

K——分类任务总类别数；

$Recall_i$——第i类召回率。

i） 宏 F1 分数：针对多分类任务，计算宏精确率与宏召回率的调和平均数，计算公式如下：

$$Macro - F1 = \frac{2 \times Macro-P \times Macro-R}{Macro-P + Macro-R}$$

式中：

Macro-F1——宏F1分数；

Macro-P——宏精确率；

Macro-R——宏召回率。

6.2 图像分割任务技术指标

本类指标适用于所有图像分割任务，即模型将图像中的每个像素点划分到特定类别的能力：

a） Dice 系数：模型预测的分割区域与金标准区域之间的重叠程度，计算公式如下：

$$Dice = \frac{2\times|X\cap Y|}{|X|+|Y|}$$

式中：

X——模型预测的区域像素集合；

Y——金标准区域像素集合。

b） IoU：模型预测的分割区域与金标准区域的交集相对于并集的比例，计算公式如下：

$$IoU = \frac{|X\cap Y|}{|X\cup Y|}$$

式中：

X——模型预测的区域像素集合；

Y——金标准区域像素集合。

6.3 生成类任务技术指标

本类指标适用于所有自然语言生成任务，即模型根据结构化数据或非结构化数据输入自动生成连贯文本的能力：

a） ROUGE-N：对生成任务，计算客观指标 ROUGE-N，其计算公式如下：

$$ROUGE-N = \frac{\sum_{S\in\{ReferenceSummaries\}}\sum_{gram_n\in S} Count_{match}(gram_n)}{\sum_{S\in\{ReferenceSummaries\}}\sum_{gram_n\in S} Count(gram_n)}$$

式中：

N——即n-gram，文本内容滑动窗口连续的字的个数，参考值为2；

$Count_{match}(gram_n)$——参考摘要和生成摘要中共有的n-gram的数量；

$Count(gram_n)$——参考摘要中n-gram的数量。

b） BERTScore：对生成任务，计算客观指标 BERTScore，计算公式如下：

$$sim(x_i, y_j) = \frac{Emb(x_i)\cdot Emb(y_j)}{\|Emb(x_i)\|\cdot\|Emb(y_j)\|}$$

$$Precision = \frac{1}{|x|}\sum_{x_i\in x}\max_{y_j\in y} sim\left(x_i, y_j\right)$$

$$Recall = \frac{1}{|y|}\sum_{y_j\in y}\max_{x_i\in x} sim\left(x_i, y_j\right)$$

$$BERTScore = \frac{2\times Precision\times Recall}{Precision+Recall}$$

式中：

$Emb(x_i)$——句子x中第i个词元在经过编码器后的嵌入向量；

$Emb(y_i)$——句子y中第i个词元在经过编码器后的嵌入向量；

$sim(x_i, y_i)$——两词语嵌入向量的余弦相似度；

Precision——精确率；

Recall——召回率；

6.4 检索与排序技术指标

本类指标适用于所有信息检索任务，即系统从大规模数据库中查找并排序最相关项目的能力：

a） Recall@K：衡量在所有相关条目中，有多少个在召回的 K 个召回条目中出现，计算公式如下：

$$Recall@K = \frac{N_{R-K}}{N_R}$$

式中：

N_{R-K}——K个召回条目中相关条目数量；

N_R——总相关条目数量。

b） 余弦相似度：计算各个向量在空间中的夹角的余弦值，来衡量向量间相似度，通常使用召回结果中最低的余弦相似度衡量总体召回效果，计算公式如下：

$$Cosine\ Similarity = \frac{Emb（x）\cdot Emb（y）}{\|Emb（x）\|\cdot\|Emb（y）\|}$$

式中：

$Emb(x_i)$——句子x经过编码器后[cls]位置的嵌入向量；

$Emb(y_i)$——句子y经过编码器后[cls]位置的嵌入向量。

7 辅助诊断要求

7.1 临床表现分析

7.1.1 功能要求

7.1.1.1 命名实体识别

淋巴瘤辅助诊疗智能体应支持对患者主诉、现病史、既往史、体格检查等非结构化文本描述进行理解和关键信息提取的能力：

a) 应支持从非结构化的临床文本中识别淋巴瘤相关核心医学实体，包括但不限于：症状、主诉、现病史、既往史、个人史、家族史、过敏史、体征、解剖部位、疾病名称等；
b) 应支持识别与区分文本对症状的否定描述（如“无盗汗”）与不确定描述（如“疑似淋巴结肿大”）。

7.1.1.2 信息结构化

淋巴瘤辅助诊疗智能体应支持对识别出的命名实体结构化的能力：

a) 应支持将识别出的命名实体及其关系进行结构化处理，生成标准化的数据字段与结构化或半结构化数据，例如，将“左侧颈部可触及一约 2cm x 3cm 肿大淋巴结”解析为｛部位：左侧颈部，体征：淋巴结肿大，尺寸：2x3cm｝；
b) 应支持对未出现特征的预填写（如“未出现 B 症状”）；
c) 应支持对识别出的实体根据标准医学术语体系（如 ICD10，SNOMED CT 等）对识别出的实体进行编码；
d) 应支持基于历史临床表现记录进行纵向分析。

7.1.2 准确度要求

本功能涉及指标如下，通用技术指标计算方法均应符合第6章要求：

a) 准确率（计算方法应符合 6.1 节）；
b) 精确率（计算方法应符合 6.1 节）；
c) 召回率（计算方法应符合 6.1 节）；
d) F1 分数（计算方法应符合 6.1 节）；
e) ROC-AUC（计算方法应符合 6.1 节）；
f) 宏精确率（计算方法应符合 6.1 节）；
g) 宏召回率（计算方法应符合 6.1 节）；
h) 宏 F1 分数（计算方法应符合 6.1 节）；
i) ROUGE-N（计算方法应符合 6.3 节）；
j) BERTScore（计算方法应符合 6.3 节）。

7.2 实验室检查分析

7.2.1 功能要求

7.2.1.1 异常项识别

淋巴瘤辅助诊疗智能体应支持对血常规、生化全项等单次实验室检查结果进行异常判定的能力：

a) 应支持对单个异常项目进行识别；
b) 应支持对具有特定临床意义的异常项目组合进行识别；
c) 应支持对检查子项目异常进行识别；
d) 应支持对异常项、异常项组合、异常子项进行标记；

e） 应支持 LIS 系统对接与实验室检查报告自动处理；
f） 宜支持结合患者当前治疗方案，提示可能由药物引起的实验室检查异常。

7.2.2 趋势分析

淋巴瘤辅助诊疗智能体应支持对血常规、生化全项等实验室检查结果基于患者基线的时序数据趋势分析：

a） 应支持对患者历史实验室检查结果进行纵向趋势分析并针对快速恶化指标进行提示；
b） 应支持基于患者基线识别数值处于非异常范围但相对基线恶化的检查项目识别。

7.2.3 准确度要求

本功能涉及指标如下，通用技术指标均应符合第6章要求：

a） 准确率（计算方法应符合 6.1 节）；
b） 精确率（计算方法应符合 6.1 节）；
c） 召回率（计算方法应符合 6.1 节）；
d） F1 分数（计算方法应符合 6.1 节）；
e） ROC-AUC（计算方法应符合 6.1 节）。

7.3 病理学检查分析

7.3.1 功能要求

7.3.1.1 病理切片分析

淋巴瘤辅助诊疗智能体应支持对淋巴瘤病理切片图像进行自动化处理，辅助病理医生进行定量分析的能力：

a） 应支持对病理切片图像中的细胞区域、组织区域、背景区域进行有效分割；
b） 应支持对肿瘤细胞富集区域进行识别和轮廓勾画；
c） 应支持对特定免疫组化染色切片进行阳性表达区域（着色）与阴性区域（未着色）的定量分割；
d） 应支持对疑难或非典型病理切片提供鉴别诊断建议与潜在亚型分类的参考，并给出相应的置信度。

7.3.1.2 分子病理学分析

淋巴瘤辅助诊疗智能体应支持对细胞测序数据进行分析的能力：

a） 应支持识别并结构化提取分子病理学报告中的关键生物标志物信息；
b） 应支持对测序数据的预处理；
c） 应支持对测序数据的分析（如过表达基因筛选，通路富集分析等）与图像化展示（如火山图、桑基图等）；
d） 应支持对细胞与组织进行分类；
e） 宜支持基因扰动等高阶功能。

7.3.2 准确度要求

本功能涉及指标如下，通用技术指标均应符合第6章要求：

a） 准确率（计算方法应符合 6.1 节）；
b） 精确率（计算方法应符合 6.1 节）；
c） 召回率（计算方法应符合 6.1 节）；
d） F1 分数（计算方法应符合 6.1 节）；
e） ROC-AUC（计算方法应符合 6.1 节）；
f） 宏精确率（计算方法应符合 6.1 节）；
g） 宏召回率（计算方法应符合 6.1 节）；
h） 宏 F1 分数（计算方法应符合 6.1 节）；
i） Dice 系数（计算方法应符合 6.2 节）；
j） IoU（计算方法应符合 6.2 节）；

k) ROUGE-N（计算方法应符合 6.3 节）；
l) BERTScore（计算方法应符合 6.3 节）。

7.4 影像学检查分析

7.4.1 功能要求

7.4.1.1 图像分割

淋巴瘤辅助诊疗智能体应支持对目标病灶的精确识别与轮廓勾画：

a) 应支持对影像图像中的可疑病灶（包括淋巴结肿大、结外侵犯及脏器受累病灶）进行自动检测、识别与精确分割；
b) 应支持对多模态影像（如 PET-CT）进行空间配准，实现解剖与代谢信息的信息融合与协同分割；

7.4.1.2 图像分类

淋巴瘤辅助诊疗智能体应支持对分割后的病灶或图像特征进行定性分析，辅助判断其性质：

a) 应支持对分割后的病灶进行多维度分类，包括但不限于按解剖部位（如颈部、纵隔、腹膜后）、良恶性风险进行分类；
b) 应支持基于系列影像比较与 Lugano 评价标准，对淋巴瘤状态进行评估分类（如完全缓解、部分缓解、疾病稳定、疾病进展），并辅助进行 Deauville 五分法评分。

7.4.1.3 报告生成

淋巴瘤辅助诊疗智能体应支持自动整合分析结果，生成包含关键量化指标与临床辅助建议的结构化报告：

a) 应支持自动提取并量化每个目标病灶的核心影像学特征（如长径/短径、最大标准摄取值（SUVmax）、平均标准摄取值（SUVmean）、肿瘤代谢体积（MTV）等）；
b) 应支持对核心影像学特征提供置信度与潜在原因辅助医生判断（如 SUV 值异常可能为炎症导致等）；
c) 应支持自动生成结构化的影像学报告，内容包括目标病灶列表（附基线及当前次检查的量化对比）、疗效评估结论、以及基于影像学的 Ann Arbor 分期建议；
d) 宜支持可交互报告生成，如点击报告中的病灶可直接定位至图像相应位置。

7.4.2 准确度要求

本功能涉及指标如下，通用技术指标均应符合第6章要求：

a) 准确率（计算方法应符合 6.1 节）；
b) 精确率（计算方法应符合 6.1 节）；
c) 召回率（计算方法应符合 6.1 节）；
d) F1 分数（计算方法应符合 6.1 节）；
e) ROC-AUC（计算方法应符合 6.1 节）；
f) 宏精确率（计算方法应符合 6.1 节）；
g) 宏召回率（计算方法应符合 6.1 节）；
h) 宏 F1 分数（计算方法应符合 6.1 节）；
i) Dice 系数（计算方法应符合 6.2 节）；
j) IoU（计算方法应符合 6.2 节）；
k) ROUGE-N（计算方法应符合 6.3 节）；
l) BERTScore（计算方法应符合 6.3 节）。

8 辅助治疗要求

8.1 综合分期分型

8.1.1 功能要求

8.1.1.1 多模态数据嵌入

淋巴瘤辅助诊疗智能体应支持将多模态的诊疗数据映射到统一的特征空间：

a) 应支持将多源异构数据（包括临床表现、实验室检查、影像学检查、病理学检查结果）映射到统一的特征空间生成患者综合特征向量；
b) 宜支持将多模态数据进行特征对齐至自然语言特征空间。

8.1.1.2 相似病人检索

淋巴瘤辅助诊疗智能体应支持寻找相似病人作为决策参考：

a) 应支持基于患者综合特征向量，从向量数据库中检索相似病例；
b) 宜支持对初步检索相似病例进行重排序，提高检索质量；
c) 宜支持对检索病例与当前病例的差异进行显式标化与提示；
d) 宜支持基于近似最邻近算法（如 NSW，HNSW 等）进行快速检索召回；
e) 宜支持对所有检索病例相似度低的情况开启循证医学证据（如寻找近期出版刊物中的相似病例）匹配。

8.1.1.3 患者分类

淋巴瘤辅助诊疗智能体应支持基于病人特征进行亚型分类与新亚型探索：

a) 应支持辅助进行 Ann Arbor 分期、CA 分期、Rai 分期、Binet 分期等；
b) 应支持辅助进行 WHO 分类及国际预后指数（IPI）等预后评分；
c) 宜支持基于降维算法或非监督机器学习算法对病人特征进行聚类并发现潜在的患者亚组，为临床研究提供线索。

8.1.2 准确度要求

本功能涉及指标如下，通用技术指标均应符合第6章要求：

a) 准确率（计算方法应符合 6.1 节）；
b) 精确率（计算方法应符合 6.1 节）；
c) 召回率（计算方法应符合 6.1 节）；
d) F1 分数（计算方法应符合 6.1 节）；
e) ROC-AUC（计算方法应符合 6.1 节）；
f) 宏精确率（计算方法应符合 6.1 节）；
g) 宏召回率（计算方法应符合 6.1 节）；
h) 宏 F1 分数（计算方法应符合 6.1 节）；
i) Recall@K（计算方法应符合 6.4 节）；
j) 余弦相似度（计算方法应符合 6.4 节）；
k) 伪 R^2：此指标计算潜类别模型或分类模型结果对观测数据变异的解释程度，计算公式如下：

$$R^2_{pseudo} = 1 - \frac{lnL(M)}{lnL(M_0)}$$

式中：

lnL(M)——当前模型的似然函数对数值；

$lnL(M_0)$——仅包含截距项模型的似然函数对数值。

8.2 治疗方案选择

8.2.1 功能要求

8.2.1.1 治疗方案推荐

淋巴瘤辅助诊疗智能体应支持依据指南与患者特征生成个体化治疗建议：

a) 应支持根据权威诊疗指南，推荐多个治疗方案供医生选择；
b) 应支持在推荐方案中标注关键药物、剂量及疗程信息；

c） 应支持考虑患者禁忌症、经济状况（如 CAR-T 疗法价格等）、药物可及性、医保政策覆盖情况、患者保险覆盖情况、关键合并症、治疗目标（如根治、姑息缓解等）等限制条件对方案进行适应性过滤；
d） 应支持综合考虑各治疗方案与医疗保险的适配情况并进行保险与治疗方案组合推荐；
e） 应支持疑难病例提示开启多学科会诊，探索个体化疗法；
f） 应支持对与指南或患者条件不符的推荐方案，明确标注禁忌点或排除原因；
g） 应支持基于疗效评估与不良反应监测数据动态调整治疗方案。

8.2.1.2 预后对比

淋巴瘤辅助诊疗智能体应支持量化展示不同治疗方案的潜在疗效与风险：
a） 应支持基于相似病例对不同方案（如 R-CHOP、R-CHOP-X 等）的预期疗效（如无进展生存期、两年复发率、五年存活率等）进行对比展示；
b） 应支持基于相似病例对不同方案的不良反应进行预测与对比。

8.2.2 准确度要求

本功能涉及指标如下，通用技术指标均应符合第6章要求：
a） 准确率（计算方法应符合 6.1 节）；
b） 精确率（计算方法应符合 6.1 节）；
c） 召回率（计算方法应符合 6.1 节）；
d） F1 分数（计算方法应符合 6.1 节）；
e） ROC-AUC（计算方法应符合 6.1 节）；
f） 宏精确率（计算方法应符合 6.1 节）；
g） 宏召回率（计算方法应符合 6.1 节）；
h） 宏 F1 分数（计算方法应符合 6.1 节）；
i） C 指数：即 C-index，指在所有可比较样本对中，预测风险与实际事件一致样本对所占的比例，计算公式如下：

$$C(\mathrm{t}) = \frac{\sum_{i\neq j} I(T_i<D_i)I(M_i>M_j)I(T_i<X_j)\,\widehat{G}(T_i)^{-2}}{\sum_{i\neq j} I(T_i<D_i)I(T_i<X_j)\,\widehat{G}(T_i)^{-2}}$$

式中：
T_i——个体i发生事件的时间；
D_i——个体i删失时间；
M_i——模型对个体i预测的风险值；
X_j——个体j的观测时间，取T_j与D_j中的小值；
$\widehat{G}(T_i)$——删失生存函数估计。
j） Brier 分数：模型在特定时间点预测事件发生概率的准确程度，计算公式如下：

$$BS(t) = \frac{1}{N}\sum_{i=1}^{N} W_i(t)[I(X_i > t) - \hat{S}_i(t)]^2$$

式中：
N——总样本数；
X_i——个体i的观测时间；
$\hat{S}_i(t)$ ——模型预测的个体i在t时间的生存概率；
$W_i(t)$——逆概率删失加权权重。

8.3 信息提醒

8.3.1 功能要求

8.3.1.1 患者随访

淋巴瘤辅助诊疗智能体应支持对患者端进行随访推送并收集信息：

a) 应支持基于治疗方案与患者分期分型，自动生成并管理结构化的个体化随访计划（包括随访时间、复查项目、症状筛查清单等）；
b) 应支持通过多种渠道（如医院 APP、短信、公众号等）向患者端自动推送随访提醒和随访问卷；
c) 宜支持基于患者上报的数据，自动推荐个性化的健康教育内容或进行进一步询问。

8.3.1.2 信息预警

淋巴瘤辅助诊疗智能体应支持对随访异常信息、医保控费信息进行医生端预警：
a) 应支持设定预警规则，对患者上报的严重症状（如高热、剧烈疼痛、新发肿块）、异常体征或复查结果中的关键异常值进行自动识别；
b) 应支持分级预警机制，并将预警信息（包括患者信息、预警内容、风险等级）实时推送至医生工作站；
c) 应支持建立预警处理闭环，通过日志记录并追踪预警信息的确认状态与处理措施；
d) 应支持基于医保知识库对高风险治疗方案进行显式提醒。

8.3.1.3 信息推送

淋巴瘤辅助诊疗智能体应支持对最近循证医学知识推送：
a) 应支持接入与维护淋巴瘤专业知识库，内容涵盖权威临床指南、重要临床研究文献、荟萃分析、新药进展等；
b) 应支持诊疗决策步骤中检索循证医学证据并引用以辅助理解；
c) 应支持基于当前患者的具体情况（如分型、分期、当前治疗线数），从知识库中智能匹配并筛选相关的循证医学知识；
d) 应支持在诊疗关键节点（如制定初始方案、调整方案、处理难治复发时）向医生主动、精准地推送相关的知识摘要；
e) 宜支持医生通过自然语言提问的方式，主动获取特定临床问题的循证答案；
f) 宜支持基于循证医学证据等级（如荟萃分析高于案例分析等）分级汇总循证医学结果。

8.3.2 准确度要求

本功能涉及指标如下，通用技术指标均应符合第6章要求：
a) 准确率（计算方法应符合 6.1 节）；
b) 精确率（计算方法应符合 6.1 节）；
c) 召回率（计算方法应符合 6.1 节）；
d) F1 分数（计算方法应符合 6.1 节）；
e) ROC-AUC（计算方法应符合 6.1 节）；
f) 宏精确率（计算方法应符合 6.1 节）；
g) 宏召回率（计算方法应符合 6.1 节）；
h) 宏 F1 分数（计算方法应符合 6.1 节）；
i) ROUGE-N（计算方法应符合 6.3 节）；
j) BERTScore（计算方法应符合 6.3 节）；
k) Recall@K（计算方法应符合 6.4 节）；
l) 余弦相似度（计算方法应符合 6.4 节）；
m) 忠实度：衡量模型生成的内容是否忠实于给定的检索证据，将生成文本分解为多个独立论点后进行计算，计算公式如下：

$$Faithfulness = \frac{N_v}{N}$$

式中：
N_v——可被上下文信息证实的论点；
N——总生成论点。

9 易用性要求

9.1 可理解性

9.1.1 语言表达清晰程度

医疗健康行业智能体界面文字、提示及交互内容应简洁准确并以标准医学术语表示，避免口语化内容导致的歧义。

9.1.2 辅助理解手段

医疗健康行业智能体涉及召回信息时应显式引用相关来源辅助医师理解。

9.2 易学性

9.2.1 帮助文档完整性

医疗健康行业智能体应配备结构化帮助文档，含功能说明、操作指南及常见问题解答，支持关键词检索，内容随平台更新同步修订。

9.2.2 差错信息易理解性

医疗健康行业智能体操作错误或系统异常时，差错信息应明确原因并提供解决方案，不应以技术代码表述。

9.3 易操作性

9.3.1 操作一致性

医疗健康行业智能体各功能模块操作逻辑、交互样式应保持统一，降低用户学习成本。

9.3.2 消息明确性

a) 医疗健康行业智能体向用户推送的各类消息，如检查提醒、复诊通知、用药提示等，内容应明确具体，包含关键信息，如时间、地点、注意事项等。
b) 医疗健康行业智能体向用户推送的各类消息的标题和正文应简洁明了，不应使用冗长复杂的表述。
c) 医疗健康行业智能体消息推送应具备合理的频率和时机，不应过度打扰用户。

9.3.3 辅助输入手段

医疗健康行业智能体应支持智能联想、语音、手写等多种输入方式。

10 安全性要求

10.1 基础设施安全

10.1.1 硬件设备安全性

医疗健康行业智能体涉及的硬件设备（如网络设备、存储设备、计算设备等）的安全防护能力应包含：

a) 通用安全要求：
   1) 应满足物理安全保障要求，包含防火、防雷、防水、灾备、授权等；
   2) 应满足功能安全保障要求，包含设备标签、硬件接口安全、固件安全、驱动程序安全等；
   3) 应满足管理安全保障要求，包含管理机制、管理人员等；
b) 计算设备安全专用要求：
   1) 应具备保障人工智能加速芯片应具备通用安全保障能力，包含 AI 加速芯片信息窃取防护、架构安全漏洞防护等；
   2) 应具备保障人工智能加速芯片在异构场景下应具备稳定运行的能力，包含 CPU 与 GPU 相结合的场景；
   3) 应具备保障人工智能加速芯片运行环境安全的能力。

10.1.2 软件系统安全性

医疗健康行业智能体所依赖的软件系统，如依赖库、AI框架等，应具备安全防护能力，包含：

a） 漏洞管理：软件设施应定期进行漏洞扫描和修复，具备完善的漏洞响应机制；

b） 安全更新：软件设施应及时更新安全补丁，以防止新出现的安全威胁。

10.2 数据安全

医疗健康行业智能体应支持数据采集、数据预处理、数据使用等数据相关内容具备安全防护能力，包含：

a） 基础数据安全：数据管理系统应具备保障存储安全、隐私保护、过程安全、销毁安全的能力；

b） 医疗数据安全：数据管理系统应具备病历分级审阅权限配置、最小必要原则核查及敏感个人信息单独同意核验的能力。

10.3 应用安全

10.3.1 内容安全

医疗健康行业智能体输出内容（含生成内容、决策内容）应符合全人类普适的道德伦理及医学伦理要求。

a） 应支持尊重人权，包括医疗健康行业智能体输出内容（含生成内容、决策内容）应遵循人权的普遍性和不可侵犯性的原则，尊重人类平等、尊严和自由的权利；

b） 应支持无偏见歧视性，包括医疗健康行业智能体输出内容（含生成内容、决策内容）避免产生偏见及歧视性结果的程度；

c） 应符合科技伦理原则，包括增进人类福祉、坚持公平公正、推动透明可释、确保可控可信等；

d） 应遵循科技伦理指标，包括公平性、透明可释性、数据隐私、可控可靠性、内容向善、责任可追溯、可持续性等。

10.3.2 服务安全

医疗健康行业智能体应支持服务安全可信、内容安全可信等应用相关内容具备安全防护能力，包含：

服务安全：医疗健康行业智能体涉及的模型安全性应满足模型安全保障要求，包含MTTF、服务安全性、服务合规性、反馈处置机制等。