

# 浙江省肿瘤医院

## 卵巢癌多中心专病数据库建设项目需求

### 一、项目概述

本项目聚焦卵巢癌诊疗难点旨在建设全国多中心专病库。项目立足我院丰富的诊疗、科研及专病库基础，联合国内多家意愿医院，旨在建成全国领先的卵巢癌多中心多模态大数据标准平台，构建智能诊疗与预后评估大模型，打造全国卵巢癌规范化诊疗与质控网络，破解卵巢癌晚期就诊、耐药复发等问题，提升患者生存率，为卵巢癌精准诊疗提供中国经验。

### 二、建设内容

项目分期推进，首先定标准并接入首批协作单位完成第一批数据治理，后续逐步扩大规模、训练模型并嵌入临床流程，最终建成多中心协作平台。建设中制定统一数据标准，涵盖患者信息、病理检查、生物样本库等多类字段；搭建多层次数据汇集体系，除了通过 NLP（自然语言处理）技术实现临床数据结构化治理之外，还需建立跨模态数据的融合机制；建立组织架构与工作流程保障项目推进，强化数据存储与网络安全；打造分级数据共享机制，推动产学研用联动，实现科研、临床、行业规范多维成果落地。

### 三、采购清单

序号	采购内容	数量	核心要求
1	卵巢癌多中心专病数据库基础平台	1 套	包含患者数据管理、临床数据结构化治理、多中心权限管控、脱敏导出等基础软件模块。
2	跨平台系统数据接口开发服务，历史及增量数据采集、治理等	1 项	必须包含与各中心院内已建专病数据库、科研数据库或 HIS、Lis、Pacs 等业务系统、生物样本库系统的数据对接，及数据抽取、汇集、清洗、治理、人工质控审核等工作。
3	卵巢癌 AI 分子分型与辅助决策系统	1 套	基于多模态融合数据的算法模型构建、训练及可视化展示模块定制开发。

#### 四、建设内容与范围

序号	一级模块	二级模块	三级模块	功能要求
1	卵巢癌专病管理系统	卵巢癌患者管理	检索	系统支持在患者所有结构化数据及非结构化文本（含病历、检查报告、病理报告）中进行关键字检索，支持中英文及数字混合检索，支持模糊匹配，并对匹配内容高亮显示； 系统支持根据按任意字段组合进行病例搜索； 系统支持条件树进行联合搜索。
2			患者管理	系统支持患者数据的新增、修改、软删除、合并及批量导入导出； 所有操作受多中心权限控制，自动记录操作日志； 关键操作（删除、合并）需双人审核或二次确认。
3		卵巢癌专病数据分析	各维度分析	系统支持按人口学、临床、诊疗、多中心等维度进行统计分析； 系统提供柱状图、饼图、折线图、生存曲线等可视化图表； 系统支持多中心数据对比和图表导出； 系统支持点击图表中的数据点或图例，自动跳转至对应的明细患者列表，并支持导出。
4		卵巢癌专病后台管理	专病库数据集管理	支持查看及定义专病库数据模型（表名、字段名、类型）； 支持字段控件配置（单选/多选/日期/文本框等）； 系统支持病人基本信息的唯一性判断设置； 系统提供表单预览； 系统支持数据指标项之间的逻辑关系设置、设置是否为必填项、设置字段显示隐藏逻辑、表单预览； 系统支持统一数据字典与值域代码的可视化配置，支持标准化映射规则的设置与版本管理。
5			专病库纳排管理	系统支持对专病库的数据进行纳排条件进行管理，并进行可视化配置（AND/OR 嵌套、时间窗、多中心）； 系统支持对纳入专病库的患者进行审核，排除系统支持对专病库的数据更新进行配置，支持自动更新，包括更新频次、更新时间，支持手工更新，提供更新日志； 系统支持对专病库的字段填充内容进行人工审核修改及记录保留。

6			专病库多中心权限管理	<p>系统支持设置不同机构组织, 并支持树形机构管理;</p> <p>系统支持专病库角色创建、模板及继承;</p> <p>系统支持设置数据范围的权限, 如查看本机构的患者</p> <p>系统支持设置数据集的权限</p> <p>系统支持设置字段的权限</p> <p>功能权限按模块控制;</p> <p>数据范围按机构/项目/队列/时间控制;</p> <p>字段级权限 (可见/脱敏/隐藏);</p> <p>支持跨中心数据访问申请与审批流程 (多级审批、临时授权);</p> <p>权限变更全审计。</p>
7			多中心数据协同	<p>系统支持分中心数据脱敏加密上传、断点续传、增量同步;</p> <p>系统支持数据冲突处理规则配置;</p> <p>提供传输状态监控仪表盘。</p>
8			日志审计	<p>记录用户登录、查询、导出、删除等操作 (含时间、IP、内容);</p> <p>日志不可篡改;</p> <p>敏感操作 (批量导出) 需双人审批。</p>
9			数据安全与脱敏	<p>系统支持动态脱敏规则配置 (按角色/字段);</p> <p>系统支持数据水印;</p> <p>导出需审批流程。</p>
10	跨平台系统数据接口开发服务, 历史及增量数据采集、治理等	卵巢癌专病数据导出	<p>系统支持导出申请与审批流程 (用途、数据范围);</p> <p>敏感字段可配置脱敏规则;</p> <p>系统支持用户自定义导出字段及保存模板; 支持 Excel、CSV、JSON、PDF 格式;</p> <p>单次导出上限可配置;</p> <p>导出文件自动添加水印 (含用户+时间+用途);</p> <p>所有导出操作记录审计日志。</p>	
11		卵巢癌专病科研管理	科研患者管理	<p>用户自定义创建条件进行患者筛选 (字段+操作符+逻辑嵌套、时间窗、全文检索), 也可快速引用已有条件进行检索, 支持保存或分享条件模板;</p> <p>系统支持检索结果可视化展示和实时统计图表辅助;</p> <p>系统支持发起数据使用申请 (用途、字段、期限), 审批通过后获得全量人群访问权限;</p> <p>系统支持科研课题管理 (创建、成员、期限、人群关联);</p> <p>系统支持检索结果支持同步到一个或多个科研课题。</p>
12			科研数据	系统支持针对科研课题人群进行描述性统计、

			分析	组间比较（t 检验、卡方等，自动推荐）、生存分析（KM 曲线+Cox 回归，可自定义事件/时间）、相关性分析（热力图）、亚组分层分析； 系统支持生成交互式图表（散点图、箱线图、森林图等）； 系统支持分析结果保存到课题，一键生成统计报告（Word/PDF）。
13			科研后台管理	系统支持科研项目全周期管理（创建、审批、状态变更、归档）； 系统支持项目成员与角色配置（多机构），细粒度权限（数据范围、导出等）； 系统支持临时授权与到期自动回收； 系统支持项目内数据操作审计日志； 项目终止后数据转为只读或归档，删除需审批； 系统支持多中心机构管理员配置。
14			科研数据导出	系统支持自定义导出字段及保存模板； 格式支持 Excel、CSV、JSON、PDF； 系统支持自动脱敏/匿名化； 导出文件添加水印（用户、课题、时间）； 所有导出记录审计并提供数据使用凭证。
15		卵巢癌患者全景视图	全景临床信息	△系统支持以 P 维度查看患者的专病信息，及原始病历信息； 以患者为中心集成概览卡片及诊疗信息时间轴（诊断、用药、检查、检验、病理、手术）； 提供 NLP 实体-关系图谱可视化（网络图或时间线图）； 支持分屏对照； 系统支持结构化字段反查原始病历原文并高亮，支持查看专病信息与原始信息的关系，支持原始信息对应的实体识别，关系挂载，病历等可视化显示。
16	生物样本库管理		展示样本采集日期、类型、保存条件、可用量、使用记录； 支持样本多组学数据（WES、RNA-seq 等）存储与展示。	
17	影像数据管理		支持 DICOM 影像上传、在线调阅（二维、MPR）、关键帧标注； 支持影像与患者、病灶、病理报告关联； 提供影像脱敏处理。	
18		系统集成		支持与医院统一身份认证集成。

19	数据深度治理	治理范围	<p>△基于卵巢癌的字建设需求,实现从入院记录现病史提取相关诊疗信息,包括但不限于分子病理检查、药物治疗、手术治疗等内容;</p> <p>基于卵巢癌的字建设需求,实现从入院记录婚育史提取相关信息,包括但不限于婚育状态、生育次数、生育方式等内容;</p> <p>基于卵巢癌的字建设需求,实现从入院记录辅助检查提取相关检查信息,包括但不限于分子病理检查、影像学检查等内容;</p> <p>基于卵巢癌的字建设需求,实现从病程记录提取相关治疗信息,包括但不限于术后并发症等内容;</p> <p>基于卵巢癌的字建设需求,实现从检查报告提取相关检查信息,包括但不限于CT检查、PET_CT检查、MRI检查等内容;</p> <p>基于卵巢癌的字建设需求,实现从病理报告提取相关信息,包括但不限于细胞学、组织病理学、免疫组化、分子病理等内容;</p> <p>对于提取的信息完成数据归一,包括但不限于诊断归一、检查归一、药物归一、病理归一、手术归一等内容;</p> <p>对于提取的信息完成关系挂载、包括但不限于手术与手术时间、检查与时间、药物与时间、诊断与时间、检查与部位等内容;</p> <p>关键字段准确率≥99%、召回率≥99%;</p> <p>支持新增字段规则配置及模型增量更新。</p>	
20			数据质控	<p>提供完整性校验规则配置;</p> <p>定期自动生成质控报告;</p> <p>支持问题数据追溯与修复流程(标注异常、人工审核)。</p>
21			能力建设	<p>△实现可视化功能,支持从现病史文本中提取药物治疗相关实体(药物类型、药物名称、给药途径、持续时间)信息和实体关联(药物名称与疑似词/数值/频率词/时间/给药途径、药物类型与持续时间)关系信息,且形成包含药物名称、持续时间标准化数值、持续时间、开始时间、状态词、药物类型、用药频次、活性成分标准化、给药途径属性的图谱化内容;</p> <p>系统支持点击溯源及人工修正。</p>
22				<p>实现可视化功能,支持从现病史文本中提取检查项目、检查部位、次级解剖部位、否定词、时间等实体信息和检查项目与时间、检查项目与部位、诊断名称与状态词否定等关系信息,且形成涵盖部位标准化、检查项目概念治理、扫描类型、检查项目、医院(本院/外院)、检查类型、检查时间、报告时间、扫描方式的图谱化内容;</p> <p>支持点击溯源及人工修正。</p>

23				<p>实现可视化功能，支持从现病史文本中提取淋巴结、病灶描述、转移、部位等实体信息和转移与状态词否定、病灶描述与时间、诊断名称与状态词否定、病灶描述与数值、部位与诊断名称、淋巴结与转移等关系信息，且形成涵盖标本部位、解剖结构部位、病灶病理大小、次级解剖部位、诊断名称、医院名称（本院/外院）、病灶描述数值、病灶描述单位、病理报告时间、状态词否定、病理诊断、常规病理分化程度标准化、病理转移的图谱化内容；支持点击溯源及人工修正。</p>
24				<p>△实现可视化功能，支持从现病史文本中提取蛋白质名称（BRCA、HRD、p53、Ki67等）、时间、蛋白质说明、标本类型、外显子、基因表达评分指标等实体信息和蛋白质名称与数值、蛋白质名称与定性结果、蛋白质名称与时间、蛋白质名称与蛋白质说明、蛋白质名称与检测方法、蛋白质名称与医院、蛋白质名称与外显子、蛋白质名称与基因表达评分指标、外显子与时间等关系信息，且形成涵盖诊断、部位、标本类型、分子检测方法、病理诊断、基因名、病理标本类型、诊断名称、数值、医院（本院/外院）、基因表达评分指标、检测方法、基因表达评分数值（HRD评分等）、报告时间、免疫组化检测方法、定性结果、定性描述、程度、变量类型的图谱化内容；支持点击溯源及人工修正。</p>
25				<p>实现可视化功能，支持从现病史文本中提取手术名称、状态词-疑似、部位、病灶描述等实体信息和手术名称与医院、手术名称与诊断名称、手术名称与部位等关系信息，且形成涵盖手术名称、医院名称、部位、手术名称标准化的图谱化内容；支持点击溯源及人工修正。</p>
26	卵巢癌 AI 分子分型与辅助决策系统	模型开发	多模态数据整合	<p>支持多模态数据（临床、影像、病理、基因组学）的特征筛选、降维、归一化及相关性分析；集成常用机器学习/深度学习框架（如 XGBoost、LightGBM、Scikit-learn、PyTorch、TensorFlow），并提供超参数自动调优工具（如 Grid Search、Bayesian Optimization）。</p>
27			联邦学习	<p>支持多中心数据不出院墙的联合建模，提供横向或纵向联邦学习框架，保护各中心数据隐私，并支持联邦模型性能评估</p>
28		模型管理与评估	模型版本管理	<p>支持 AI 模型上传、版本控制、回滚及部署状态管理；自动计算并可视化展示模型评估指标，包括但不限于：AUC、ROC 曲线、F1-score、敏感度、特异度、阳性预测值、阴性预测值、校准曲线、决策曲线分析（DCA）；</p>

				支持交叉验证及外部验证集评估。
29			模型可解释性	提供 SHAP、LIME、Integrated Gradients 等多种可解释性工具，输出特征重要性排序、个体预测解释及可视化图表（如瀑布图、力图），辅助临床医生理解模型决策逻辑。
30		临床辅助决策应用	分子分型与疗效预测	△基于多模态数据输出卵巢癌分子分型（如 P53 突变型、BRCA 突变型、HRD 阳性型等）概率及置信区间； 预测铂类药物化疗耐药性、PARP 抑制剂响应疗效、免疫治疗响应概率，并提供预测依据摘要。
31			预后评估与临床决策	△输出复发风险等级、无进展生存期（PFS）及总生存期（OS）的预测曲线和生存概率； 在患者全景视图中嵌入 AI 预测卡片，如动态展示预测结果随数据更新的变化，支持医生对预测结果进行标注、接受或拒绝反馈，反馈数据用于模型持续优化。
32	工期	180 个自然日		
33	质保期	最终验收达到合格后 3 年		