附件

2025年度生物样本库建设

技术攻关项目榜单

1. 需求目标

为贯彻落实《北京市支持创新医药高质量发展若干措施（2025年）》，支撑生物样本库建设，需开发基因测序和蛋白质组学检测的低成本、自主可控关键设备及技术平台。同时，搭建人工智能随访系统，以辅助优化随访效率和质量，为生物样本库持续提供丰富的、高价值的患者数据资源，支持开展更加深入、全面的医学研究工作。

二、技术攻关榜单

（一）基因测序平台

1.榜单任务

建立高通量基因测序平台，应实现以下指标。

（1）自动化水平：拥有从样本到数据全流程自动化能力，包括自动样本前处理和文库建库系统。

（2）通量要求：测序读长不低于150bp，测序仪及样本自动化处理设备可兼容支持靶向测序、全基因组测序及全基因组甲基化测序。

（3）数据质量：①人全基因组测序：有效测序平均深度为30-35X，PE150平均测序质量Q30≥85%，经过质量控制处理后的测序数据（clean数据）GC含量在39%-43%，基因组比对率≥99%，当使用人基因组DNA实物标准及配套标准变异集评估时，单核苷酸多态性位点（SNP）检测准确率≥99%，灵敏度≥98%；插入缺失突变位点（InDel）检测准确度≥90%，灵敏度≥90%。②靶向测序（此方向技术要求以样本提供方的实际需求为准）：针对肿瘤组织样本，预期单核苷酸多态性（SNVs）/插入缺失突变（InDels）的检测限应不高于1%，基因融合突变检测限应不高于5%，目标区域的平均有效测序深度应不低于800X；针对液体活检样本，预期单核苷酸多态性/插入缺失突变的检测限应不高于0.1%，目标区域的平均有效测序深度应不低于10000X；PE150平均测序质量Q30≥85%；≥20%平均有效测序深度的覆盖区域不低于目标区域的95%；需经国家参考品或企业参考品对具体检测项目的检测限、准确度、重复性进行验证。③全基因组甲基化测序：有效测序平均深度为30X；采用重亚硫酸盐转化的转化率应不小于99%，采用酶法转化的转化率应不小于95%；胞嘧啶测序覆盖度不低于90%；使用人全基因组甲基化标准物质评估，应符合标准物质甲基化位点数及甲基化率的阈值范围，甲基化位点的准确度应不小于95%，灵敏度应不小于90%；同一管标准物质或同一管DNA样本的2个技术（从建库到测序）重复的位点甲基化率的皮尔森相关系数不小于0.95。

（4）生信分析能力：研发配套的生信分析流程，日处理能力不少于120人份全基因组测序数据（包括下机数据质量过滤、序列比对、检测变异和注释）。

（5）成本控制：本方向核心目标是在满足上述技术指标基础上，以5万例样本为测算依据，探索单样本基因测序极限成本。同等技术指标下，优先选择成本更低的揭榜团队或联合体。

2.考核指标

（1）建立包含全自动样本前处理系统、文库建库系统、具有自主知识产权的基因测序仪、标准化生物信息分析流程的系统1套；

（2）委托第三方检验机构出具符合上述指标的全基因组测序数据质量报告、全基因组甲基化测序数据质量报告、靶向测序数据质量报告；

（3）开展不低于5000例样本的基因检测服务，其中全基因组测序、靶向测序（测序深度及panel规模需符合样本提供方要求）、全基因组甲基化测序例数以具体样本提供方为准，提交可体现关键性能指标的测序报告和样本提供方的验收报告。

（4）提交每个检测项目单样本尽可能详尽的成本核算方案及必要的价格证明信息。

（二）蛋白质组学检测平台

1.榜单任务

建立规模化靶向蛋白质组学检测平台（针对体液样本），应实现以下指标。

（1）自动化水平：检测全流程操作自动化。

（2）样本量：可实现对低至5ul样本的检测。

（3）通量要求：单套检测通量不低于300样本/天，年检测量不低于7万个。

（4）浓度检测：针对单样本的全部蛋白，实现不低于5个数量级别动态跨度的蛋白检测。

（5）数据质量：在常见临床样本中可检测不少于2000种蛋白质，至少覆盖4种疾病队列相关的生物标志物，平均CV≤15%，需用不少于5个临床已成熟应用的蛋白靶标临床检测结果作为定量相关性质控内参。

（6）成本控制：本方向核心目标是在满足上述技术指标基础上，以5万例样本为测算依据，探索单样本蛋白组学检测极限成本。同等技术指标下，优先选择成本更低的揭榜团队或联合体。

2.考核指标

（1）按产品技术指标形成试剂及检测样机。

（2）提供满足产品技术指标的两家应用单位测试报告。

（3）开展不低于1000例样本蛋白质组学检测服务，提交可体现关键性能指标的检测报告和样本提供方的验收报告。

（4）提交每个检测项目单样本尽可能详尽的成本核算方案及必要的价格证明信息。

（三）人工智能随访系统

1.榜单任务

搭建人工智能辅助随访系统，应实现以下指标。

（1）多途径随访：支持智能外呼、短信/微信、视频、可穿戴设备等多途径随访方式，满足不同随访需求。结合不同病种特点、研究需求以及数据质量要求，能够设置不同随访路径。智能外呼并发支持能力≥200路。

（2）高仿真多并发自然对话及语音数据自动整理：具备高精准度的语音合成与识别技术，支持不同口音，能够实现医疗场景下与患者自然流畅的语音对话，且具备多轮连续对话能力。将随访过程中患者的语音回复自动转换成文本，解决部分方言、表述不清等识别难点，提高识别率和准确率，经结构化处理后更新到随访数据中。语音识别准确率≥97%，语义理解正确率≥97%，语音合成自然度≥4.5分。

（3）支持智能穿戴设备：支持接入至少2款具备二类医疗器械证资质的可穿戴智能硬件设备。系统应同时具备实时采集患者体征数据和通过患者端OCR识别检查化验记录的功能，采集信息经结构化分析后上传至研究队列。根据随访计划，对于需要定期监测指标的患者，提醒患者按时进行指标上传，应支持手动录入和智能上传。

（4）支持患者端OCR识别：随访患者端支持门诊病历、住院病历、检查检验报告单上传，系统完成OCR识别后，结构化数据按规则要求更新至患者随访数据中的准确率需达到90%以上。

（5）系统管理能力：系统需支持管理至少100个研究队列，总患者容量不少于10万人，且支持未来数据量增长的平滑扩展。

（6）患者服务：通过语音电话、患者小程序等手段，精准推送科室、疾病、手术、用药相关知识。

（7）随访异常与结果质控：支持异常随访的闭环管理。若出现患者回答病情未控制、患者死亡、电话未接通/无法联系等异常情况时，系统应自动识别，并生成重点关注的随访对象清单。

（8）人群疾病筛查与复诊提醒：结合各病种的专业评估量表，利用人工智能电话进行大面积人群初筛和定期筛查，发现高风险人群后由人工随访引导就医或入组。

2.考核指标

（1）按照上述任务列出的技术指标和成果形式作为验收的主要依据，完成人工智能随访系统一套，并提供操作手册和系统测试报告。软件需通过第三方专业机构的安全性、稳定性、技术指标等方面的测试，测试报告完整、真实。系统需支持管理至少100个研究队列，总患者容量不少于10万人，且支持未来数据量增长的平滑扩展。

（2）选择至少2种不同病种接入并进行实际环境使用验证，验证时间不少于30天，并提供验收报告。

（3）项目交付后，开发团队需根据实际应用情况，在6个月内提供至少1次系统功能优化升级服务，优化升级内容需切实解决应用过程中发现的问题，提升系统性能和应用体验。

三、项目周期

方向一：执行期原则上不超过6个月

方向二：执行期原则上不超过18个月

方向三：执行期原则上不超过6个月

四、榜单金额

方向一：不超过300万元

方向二：不超过900万元

方向三：不超过200万元

五、知识产权归属

本次验证所产生的组学及随访等相关数据知识产权归生物样本和数据等资源提供方所有。