［本报告数据及结论尚未公开发表，仅限内部传阅，请勿公开发布或转载］

AI 赋能工具包提升基层医疗服务质量有效性项目报告要点总结

陕西师范大学教育实验经济研究所

0. 核心结论

1) 基于标准化病人测试结果表明，村医小助手短期使用（13 天试用期）没有显著提升医生的医疗服务水平；
2) 基层医生的医疗 AI 技术素养和使用意愿度高，干预后有所改善，患者接受度高；
3) 基于医生的反馈，村医小助手日后能解决医疗工作中 75%的诊疗需求，具有推广性；
4) 在迭代改进的方向上，短期可以方案实施升级和提示词优化为主，长期可考虑大模型的底层升级；
5) 对基层医生的医疗服务水平改善情况需更长干预周期、大样本的评估，同时在模型上考虑国产模型的替代可行性的有效性。

1. 基本情况

1) 机构：10 个村，每村平均 1 个卫生室，1 名问诊医生，合计 10 名样本村医；平均接诊 106 人次/月，约 3.7%会转诊到上级医院。44.5%的村民常去卫生室看病。
2) 医生：平均年龄 56.7 岁，男性占比 70%，累计从医时间 33.8 年；半数最高教育水平为大专，50%持有乡村医生执业证书，90%无职称。

2. 医疗工作需求

60%村医自报告现有知识水平不能满足诊疗需求。约 36%的概率会遇到自己拿不准、看不了的病，主要应对办法是转诊，其次通过网络、书籍查询资料。约 30%提及使用网络平台查询医学资料，常用搜索引擎、医学平台或微信公众号。

3. 互联网技术素养

1) 熟练掌握微信、抖音等软件使用，但 40%存在打字困难和普通话不标准的问题。30%的村医听说过医疗 AI 产品，但没有使用过。70%村医的使用意愿度较高。
2) 在使用态度、焦虑感、有用性感知、信息的信任感和功能的信任感上，村医对医疗 AI 产品处于较高水平。
3) 使用村医小助手后，在使用态度、信息的信任感和功能的信任感方面有 6%-13%的提升，但焦虑感上升，有用性感知下降。

4. 村医小助手使用情况

1) 使用频次

干预的 13 个工作日内，村医为 32%的接诊患者使用了村医小助手进行辅助问诊。干预组内，5 名村医和 5 名乡镇医生接受医疗 AI 工具的辅助诊疗支持。平均每次问诊问询 2-3 轮对

话。10 名医生累计使用 141 次，平均 14.1 次/人，其中乡医平均 13 次/人，村医平均 15.2 次/人。

2) 问诊需求和场景

a) 医生自报告认为，村医小助手至少解决了其诊疗工作中 70%的难题。

b) 使用场景：85%为患者病情讨论，15%为医学知识查询。咨询需求主要围绕诊断和治疗建议（66.7%），以慢性病、常见病和老年人疾病为主（68.4%）。

c) 在患者病情讨论场景下，90%医生会在首轮对话中提供详尽症状，通过少轮对话快速获得诊断和治疗建议，且在乡医群体中这一表现更为突出。

d) 有 5 次因病情严重，超出基层诊疗能力范围，小助手建议立即转诊。

5. 基于标准化病人法测试的有效性评估

基于高血压和肺结核病例评估后发现，在开展为期 13 天的试用后，村医小助手没有显著提升基层医生的医疗服务水平。

通过后台数据分析发现，干预期内这两类疾病村医通过村医小助手的问诊较少，没有咨询过肺结核病例，仅咨询 4 次高血压病例。在过往问诊咨询中，主要寻求具体用药建议和治疗方案调整建议，较少关注建议问诊和检查条目及诊断方面的建议，对两个病例的学习情况较少。这可能是导致短暂试用后未能显著提升的主要原因。

6. 使用爽点：

1) 医生认为，村医小助手的专业性、便捷性、医疗知识面广等方面的优势突出。

2) 绝大多数医生的整体满意度较高，认为小助手具备推广性，能满足日后医疗工作中 75%的诊疗需求。

3) 83%的患者对医生使用村医小助手持支持态度，患者满意度高。

7. 使用难点

1) 工具本身

a) 场景适配度：仅关注医生的诊疗需求，对公共卫生服务需求和医技人员需求适配性较低；存在答非所问和提供建议冗长情况，提示词设定需细化，模型准确性需提升；受检查设备缺乏的限制，无法针对检查建议开展相关检查。

b) 产品体验：语音转文字精准性低、无法保存历史记录、缺乏诊疗建议反馈。

c) 模型能力：缺乏中医语料库，导致回复不精准、药品建议无法给出具体药品细类或商品名称、无法识别检查单或图片、缺乏对上下文的综合理解能力。

2) **医生视角**：没有形成科学问诊习惯；本职工作繁忙；患者人数少。

3) **项目实施**：培训时间较短、培训内容不全面、督导方案较单一。

8. 解决方案

1) 首先，**优化场景适配和产品体验**。设计针对性提示词，满足不同人群和场景下基本需求；完善语音转文字功能，并开发使用反馈和历史记录查询功能。

2) 其次，完善项目方案和落地实施。完善培训方案，强化科学问诊和使用技巧的培训和练习，健全监督指导机制。
3) 最后，升级模型或实现国产模型替代，提升模型理解力和交互体验。