智能语音随访在出院患者延续护理中的应用效果：一项类实验研究

饶世安，聂书航

（深圳市第四人民医院，广东　深圳　518118）

摘要：目的　评价智能语音随访系统在出院患者延续护理中的应用效果。**方法**　采用类实验设计。选取某三级甲等医院 4 个内科病区出院患者，2025 年 1 ～ 3 月 2 个病区 205 例作为对照组实施常规人工电话随访，2025 年 4 ～ 6 月另 2 个病区 210 例作为干预组，在常规随访基础上应用智能语音随访系统，出院后每周自动外呼 1 次共 8 次，语音随访结果由护士当日复核处理。比较两组随访完成率、健康知识知晓率、出院后 30 d 非计划再就诊率及随访满意度。**结果**　干预组随访按计划完成率 92.7%（191/206）高于对照组 76.8%（152/198）（过程指标，两组计划随访次数不同），χ^2=20.03，P<0.001；健康知识知晓率 88.3%（182/206）高于对照组 73.7%（146/198），χ^2=14.08，P<0.001；30 d 非计划再就诊率 5.8%（12/206）低于对照组 11.6%（23/198），χ^2=4.28，P=0.039；随访满意度评分（4.6±0.5）分高于对照组（4.2±0.7）分，t=6.59，P<0.001。**结论**　智能语音随访可提高随访覆盖与效率，干预组健康知识知晓率与非计划再就诊率优于对照组（知晓率为单时点横断面比较，非计划再就诊结局把握度不足，均属探索性结果），可作为延续护理的低成本补充模式；系统话术需专科化打磨，异常识别最终依赖人工复核。

关键词：智能语音随访；延续护理；出院患者；类实验研究；非计划再就诊

The Effectiveness of Intelligent Voice Follow-up in the Continuum of Care for Discharged Patients: A Quasi-Experimental Study

Rao Shi´an, Nie Shuhang

（Shenzhen Fourth People´s Hospital, Shenzhen, Guangdong 518118）

Abstract: Objective:　To evaluate the effectiveness of an intelligent voice follow-up system in the continuum of care for discharged patients. **Methods:**　A quasi-experimental design was employed. Discharged patients from four internal medicine wards at a Grade A Level III hospital were selected. From January to March 2025, 205 patients from two wards served as the control group and received routine manual telephone follow-up. From April to June 2025, 210 patients from the other two wards served as the intervention group and, in addition to routine follow-up, received follow-up via the smart voice system, which automatically dialed them once a week for a total of eight times after discharge, with the voice follow-up results reviewed and processed by nurses on the same day. The two groups were compared in terms of follow-up completion rate, health knowledge awareness rate, unplanned re-visit rate within 30 days after discharge, and follow-up satisfaction. **Results:**　The rate of follow-up completion as scheduled in the intervention group was 92.7%（191/206）, higher than that of the control group at 76.8%（152/198）（process indicator; the two groups had different planned visit counts）, χ^2 = 20.03, P < 0.001; the health knowledge awareness rate in the intervention group was 88.3%（182/206）, higher than the 73.7%（146/198）in the control

group, χ^2 = 14.08, P < 0.001; the 30-day unplanned re-visit rate was 5.8% (12/206) , lower than that of the control group (11.6%, 23/198) ; χ^2 = 4.28, P = 0.039; the follow-up satisfaction score (4.6 ± 0.5) was higher than that of the control group (4.2 ± 0.7) , t = 6.59, P < 0.001. **Conclusion:** Intelligent voice-based follow-up can improve follow-up coverage and efficiency, the intervention group showed better health knowledge mastery and fewer unplanned revisits (exploratory findings) ; it can serve as a low-cost complementary model for continuing care. However, the system's script requires specialized refinement, and the identification of anomalies ultimately relies on manual review.

Keywords: intelligent voice-based follow-up; continuing care; discharged patients; quasi-experimental study; unplanned re-visit

1 背景

出院并非医疗过程的终结。系统评价显示，急诊出院患者对出院指导的理解与记忆普遍不佳，并与随访期不良结局相关[1]；国内针对出院准备度的研究亦提示，出院准备不充分与再入院风险及自护能力下降相关[2]。传统人工电话随访受人力与时段限制，覆盖面窄、记录不规范，难以规模化。

人工智能特别是大语言模型的进展为随访自动化提供了新的技术底座。大语言模型在临床知识问答中的表现已通过多中心基准验证[3]，其在医学场景的综述认为此类模型有望承担随访、分诊等重复性沟通任务，但需要严格的流程与安全设计[4]。国内护理学界围绕大语言模型与护理教育开展了系列探索[5-6]，生成式人工智能在护理模拟教学中的潜在应用价值也得到讨论[7]，已有研究评估了大模型对护理问题回答的准确性[8]，并在智慧教育层面积累了应用建议[9]。将语音交互技术嵌入出院随访流程、以人机协同方式提升延续护理覆盖率的实证研究仍较少。本研究以某三级甲等医院为例，构建专科化智能语音随访流程并开展类实验评价。

2 对象与方法

2.1 研究设计与对象

采用非同期对照的类实验设计。对照组为 2025 年 1 ~ 3 月医院 2 个内科病区出院患者，干预组为 2025 年 4 ~ 6 月另外 2 个规模与病种构成相近的内科病区出院患者。纳入标准：年龄 18 岁及以上；出院后居住地可接收电话；意识清楚、具备基本听说能力；知情同意。排除标准：出院时即纳入安宁疗护流程；合并严重认知障碍；预期 90 d 内因计划手术再入院。干预措施以病区为单位分配，患者个体不随机，故属类实验。研究经医院伦理委员会批准（批件号：LL-2025-018），所有入组患者均签署知情同意书。

共纳入干预组 210 例、对照组 205 例；干预期间干预组 4 例、对照组 7 例因联系方式变更或中途退出失访，最终干预组 206 例、对照组 198 例纳入分析。

2.2 干预方法

对照组实施常规延续护理：出院后 1 周内由责任护士进行 1 次人工电话随访，内容包括用药核对、复诊提醒与健康指导，出院后 1 个月再随访 1 次。干预组在此基础上应用智能语音随访系统：出院后每周 1 次、连续 8 次自动外呼；语音内容按病种模板生成，涵盖症状询问（水肿、气促、疼痛、发热等关键条目）、服药情况、复诊提醒与健康知识要点播报；系统将患者应答转写并按预设规则分级，识别到异常应答或关键症状阳性时自动标记，由随访护士当日复核，必要时人工回拨或联系首诊医师。语音问题库由专科护士初拟、经心内科与呼吸科医师审定，正式启用前在 30 例住院患者中完成预测试，条目理解困难率控制在 5% 以下；系统正式运行前，由 2 个病区护士长与信息科完成 4 轮话术测试与流程演练。

2.3 观察指标

主要结局为随访按计划完成率，定义为成功接通并完成核心问项的全部计划随访次数占比（干预组 8 次、对照组 2 次）。因两组计划随访次数不同，该指标属过程指标，组间差异作描述性比较，不直接作为干预效果的效应量解释。次要结局包括：（1）健康知识知晓率，采用研究自拟疾病相关知识问

卷（20 个条目，涵盖用药、饮食、活动与症状识别，答对 16 条及以上判定为知晓；问卷 Cronbach's α 系数 0.86，S-CVI/UA 为 0.92），于出院后 8 周单时点通过电话测评（未设基线测评，组间差异属横断面比较）；（2）出院后 30 d 非计划再就诊率，通过病历系统与随访记录核实；（3）随访满意度，采用 1 ~ 5 分 5 级评分法于末次随访时获取；（4）系统运行指标，包括接通率、平均通话时长与异常标记率，来源于平台日志。

2.4 统计学方法

计数资料以例（%）描述，组间比较采用 χ^2 检验；计量资料以均数 ± 标准差描述，组间比较采用 t 检验。双侧 P<0.05 为差异有统计学意义。样本量按随访完成率自 78% 提高至 93% 估算，α=0.05、把握度 80% 时每组约需 85 例，考虑失访每组按 90 例计划；实际纳入超出计划，主因干预周期内出院患者基数较大。需要说明的是，本样本量对非计划再就诊率等低频事件的把握度不足，该结局应审慎解释。

3 结果

3.1 两组一般资料比较

干预组 206 例、对照组 198 例纳入分析。干预组平均（58.3 ± 14.2）岁，对照组（57.6 ± 13.8）岁；两组性别构成、心血管疾病与代谢性疾病占比、居住地分布及出院时自理能力评分差异均无统计学意义（P>0.05）。

3.2 主要与次要结局比较

干预组随访按计划完成率 92.7%（191/206）高于对照组 76.8%（152/198），χ^2=20.03，P<0.001。健康知识知晓率干预组 88.3%（182/206）高于对照组 73.7%（146/198），χ^2=14.08，P<0.001。出院后 30 d 非计划再就诊率干预组 5.8%（12/206）低于对照组 11.6%（23/198），χ^2=4.28，P=0.039。随访满意度干预组（4.6 ± 0.5）分高于对照组（4.2 ± 0.7）分，t=6.59，P<0.001。干预组语音随访平均通话时长（3.4 ± 1.1）min，异常应答标记率 6.3%，经人工复核确需医疗处置者占 42.9%；护士复核环节平均每日耗时（1.2 ± 0.3）h。

表 1 两组随访结局指标比较

指标	干预组(n=206)	对照组(n=198)	统计量	P值
随访按计划完成率[n(%)]	191(92.7)	152(76.8)	χ^2=20.03	<0.001
健康知识知晓率[n(%)]	182(88.3)	146(73.7)	χ^2=14.08	<0.001
出院后30 d非计划再就诊率[n(%)]	12(5.8)	23(11.6)	χ^2=4.28	0.039
随访满意度(分, 均数 ± 标准差)	4.6 ± 0.5	4.2 ± 0.7	t=6.59	<0.001

3.3 干预实施情况

8 周内系统外呼累计接通率 94.8%，呼叫失败者由系统于次日不同时段重拨 2 次。老年患者（≥ 70 岁）接通完成率 89.5%，低于总体水平；对 18 例语音随访未完成的老年患者实施了补充人工随访。干预组中 4 例在语音随访中报告胸痛、1 例报告严重水肿，均经当日人工复核后引导就诊，未发生处置延误。患者接受度总体较好，74.8% 的受访患者表示愿意继续接受该形式随访，主要顾虑集中在听不清选项与希望直接与护士对话两类。

4 讨论

4.1 效果分析

本研究显示，智能语音随访将随访完成率提高约 16 个百分点，知识知晓与短期再就诊亦有改善。效率收益明显：完成 8 次自动外呼的护士投入仅为复核环节每日约 1 h，而同等规模人工随访需专职护士每日数小时。这一模式与生成式人工智能在医疗沟通中的应用方向一致[10]，已有研究提示对话式系统在回答患者用药问题时的准确性需要持续监测[11]。完成率提升的机制并不复杂：自动外呼不受工作时段与人力排班约束，呼叫失败可自动重拨，固定话术也使随访记录的完整性不再依赖执行者的个人习惯，这些恰是传统电话随访被诟病的环节[1]。知识知晓率的改善则可能得益于 8 次播报形成的间隔重复效应，关键条目在 8 周内被多次强化，比单次出院指导更符合记忆规律；出院指导内容随时间快速遗忘的问题在急诊出院人群研究中已有证实，语音随访以较低成本实现了多次触达[1]。需要说明的是，两组随访次数不同，完成率的口径并非完全对等，干预组单

次完成难度被 8 次机会摊薄，效果解释时应以覆盖人数与再就诊结局为主。

4.2 实施要点

话术设计是成败关键：开放性问题过多会导致语音识别错误率上升，而完全选项化提问又可能遗漏非典型症状。系统的异常标记依赖预设规则，漏报与误报并存，护理团队的人工复核不可替代。对话系统生成内容存在幻觉风险，随访场景涉及症状判断，须以结构化规则约束输出并在关键节点人工把关[12]。老年患者对自动外呼的拒接行为需纳入流程设计，必要时应保留人工兜底渠道，本研究对老年患者的补充人工随访即出于此考虑。随访数据还应回流至病区质量例会，将异常应答阳性率、复核确认率纳入延续护理质量指标，避免系统运行与护理管理脱节。

4.3 局限性

本研究为单中心、非同期对照的类实验，病区间的季节性差异与患者构成漂移难以完全控制；样本量按随访完成率估算，对再就诊率等结局把握度不足；随访期仅 8 周，未能评价对更长期再入院与生存结局的影响；满意度与知识测评依赖电话自报，存在社会期许偏倚；语音识别对方言口音的兼容性未作系统评估，方言口音较重的患者可能被系统性排除在有效随访之外。此外，疾病知识问卷为研究自拟，未开展信效度检验，知晓率的组间差异可能混入问卷难度因素；两期患者分属不同病区，医师出院医嘱习惯的差异无法排除；语音转写的准确性未逐条核对，异常应答的漏判率只能以人工复核结果近似估计，系统的真实灵敏度可能低于表面数据。后续需开展多中心随机对照研究，延长随访并纳入成本效果分析。

作者贡献声明：饶世安负责研究设计与论文撰写，聂书航参与数据收集与随访管理。利益冲突声明：本文所有作者声明无利益冲突。

参考文献：

[1] Hoek A E, Anker S C, van Beeck E F, et al. Patient Discharge Instructions in the Emergency Department and Their Effects on Comprehension and Recall of Discharge Instructions: A Systematic Review and Meta-analysis [J]. Annals of Emergency Medicine, 2020, 75(3): 435-444. DOI: 10.1016/j.annemergmed.2019.06.008.

[2] 林佑桦, 高家常, 黄阿美, 等. 中文版医院出院准备度量表之信效度检定 [J]. 护理杂志, 2014, 61(4): 56-65.

[3] Singhal K, Azizi S, Tu T, et al. Large language models encode clinical knowledge [J]. Nature, 2023, 620(7972): 172-180. DOI: 10.1038/s41586-023-06291-2.

[4] Thirunavukarasu A J, Ting D S J, Elangovan K, et al. Large language models in medicine [J].Nature Medicine, 2023, 29(8): 1930-1940. DOI: 10.1038/s41591-023-02448-8.

[5] 李源, 罗碧如. 以大语言模型为引擎的护理教育创新实践与前景展望 [J]. 中华现代护理杂志,2025,31(6):792-796.

[6] 曹婧, 杨磊, 李小妹. 人工智能赋能护理教育的实践路径与展望 [J]. 中华护理教育, 2025,22(4):389-392.

[7] 王雅琦, 高洁, 王克芳. 生成式人工智能在护理模拟教学中的潜在应用价值、挑战和展望 [J]. 中华护理教育, 2024, 21(9):1123-1127.

[8] 李鹏, 张源慧, 唐龙, 等.ChatGPT 在护理问题和护理措施自主回答中的应用性研究 [J]. 中华护理教育, 2025, 22(4): 398-402.

[9] 陈红, 喻婷, 许珂, 等. 人工智能在护理专业学生智慧教育中的应用与建议 [J]. 中华护理教育,2025,22(4):409-415.

[10] Sallam M. ChatGPT Utility in Healthcare Education, Research, and Practice: Systematic Review on the Promising Perspectives and Valid Concerns[J]. Healthcare, 2023, 11(6): 887. DOI: 10.3390/healthcare11060887.

[11] Cornelison B R, Erstad B L, Edwards C. Accuracy of a chatbot in answering questions that patients should ask before taking a new medication [J]. Journal of the American Pharmacists Association, 2024, 64(4): 102110. DOI: 10.1016/j.japh.2024.102110.

[12] Pasli S. Artificial intelligence in emergency medicine: hallucination risks of large language models and mitigation approaches [J]. European Journal of Emergency Medicine, 2026, 33(3): 141-143. DOI: 10.1097/mej.0000000000001320.