

doi: 10.3969/j.issn.1672-4933.2026.04.010

社区老年听觉认知减退的智能简化评估方法

Intelligent Simplified Assessment Method for Elderly Hearing Loss in Communities

朱蓓¹ 王雪丽¹ 鲁秀玲² 陈洁² 保志军² 张敏²

ZHU Bei, WANG Xue-li, LU Xiu-ling, CHEN Jie, BAO Zhi-jun, ZHANG Min

【摘要】目的 构建并验证适用于社区人群的听觉认知功能智能简化评估方法。**方法** 选取35名健康大学生建立时间压缩言语测试、听觉数字广度及听觉Stroop测试常模,以95%置信区间下限作为功能减退判别标准,另选取第10百分位数界值进行敏感性分析。随后对466名社区居民采集完整测试数据,按2:1划分训练集与验证集,基于贝叶斯自适应测试、计算机自适应测试及支持向量机等算法对测试流程进行简化,在验证集中评估其一致性与诊断效能。对88名受试者进行3个月随访以验证稳定性。**结果** 简化版测试较完整版耗时减少50%~75%,且与完整版结果高度相关($r=0.88\sim0.94$)。受试者工作特征曲线(receiver operating characteristic curve, ROC)分析显示分析显示曲线下面积(area under the curve, AUC)为0.85~0.89,敏感度 $\geq 90\%$ 。验证集结果表明模型泛化能力良好($\Delta AUC < 0.05$)。随访分析显示,重复测量一致性较高组内相关系数(intraclass correlation coefficient, ICC=0.87)。**结论** 本研究构建并验证的基于智能算法的简化听觉认知评估工具,在严格保证测评信度与诊断效度的前提下,能够显著缩短评估时间,提升基层筛查效率。该工具克服了传统评估工具的冗长弊端,具有良好的模型泛化能力、重测稳定性和临床转化价值。将其集成于移动设备或智能手机平台中,适合在社区中老年人中开展大规模听觉与认知障碍早期筛查。

【关键词】 听觉认知;时间压缩言语;数字广度;听觉斯特鲁普;智能简化;社区筛查

【中图分类号】 R764.43

【文献标识码】 A

【文章编号】 1672-4933(2026)04-0377-07

【Abstract】 Objectives Construct and validate an intelligent simplified assessment method of auditory cognitive function for community people. **Methods** The norms of time compression speech test, auditory digit span test and auditory Stroop test were established by 35 healthy college students. The lower limit of 95% confidence interval was used as the criterion of functional impairment (the 10th percentile limit was selected for sensitivity analysis). The Data of 466 community residents were collected, and the training set and verification set were divided by 2 : 1. The test process was simplified based on Bayesian adaptive test, computer adaptive test and support vector machine, and its consistency and diagnostic efficiency were evaluated in the verification set. Another 88 subjects were followed up for 3 months to verify the stability. **Results** The test time of the simplified version was 50%-75% less than that of the full version, and was highly correlated with the results of the full version ($r=0.88\sim0.94$). ROC analysis showed that the AUC was 0.85-0.89, and the sensitivity was $\geq 90\%$. The results of validation set showed that the generalization ability of the model is good ($\Delta auc < 0.05$). The follow up analysis showed that the consistency of repeated measurements was high (ICC=0.87). **Conclusions** the simplified auditory cognitive assessment tool based on intelligent algorithms (including Bayesian adaptive and SVM feature selection) constructed and verified in this study can significantly shorten the assessment time and improve the efficiency of grass-roots screening under the premise of strictly ensuring the reliability and diagnostic validity of the assessment. This tool overcomes the tedious drawbacks of traditional evaluation tools, and has good model generalization ability, retest stability and clinical transformation value. Integrating it into the mobile device or smart phone platform is very suitable for large-scale early screening of hearing and cognitive impairment in the middle-aged and elderly people in the community.

【Key words】 Auditory cognition; Time-compressed speech; Digit span; Auditory Stroop; Intelligent simplification; Community screening

1 背景

全球范围内60岁以上老年人中有显著比例的个体遭

受不同程度的听觉功能减退^[1~5]。我国情况尤为严峻,超2000万的老年人群面临这一问题且未得到干预^[6,7]。作

基金项目:上海市松江区科技攻关项目“社区老年听觉认知减退的智能简化评估方法”(2024SJKJGG047)

作者单位:1 上海市松江区九亭镇社区卫生服务中心 上海 201615

2 复旦大学附属华东医院 上海 200040

作者简介:朱蓓 本科 主任医师;研究方向:眼耳鼻咽喉科

通讯作者:张敏, E-mail: miz18@pitt.edu

为老年人常见慢性病之一且位居老年痴呆症可控因素之首的老年听觉功能减退是健康老龄化的重大障碍^[8],其不仅限制了老年人的社交活动,还可能加速认知功能的衰退,影响整体健康水平^[9~11]。早期评估和干预老年听觉功能减退能够有效提高老年人群生活质量,降低社会负担,是健康老龄化和主动健康的重要切入点^[12, 13]。

当前基层听觉评估主要聚焦外周感知功能,中枢听觉认知功能缺乏可推广的评估工具。笔者通过智能简化纯音测听流程^[13, 14]和老年听力障碍筛查量表(hearing handicap inventory for the elderly - screening version, HHIE-S)^[15]实现了社区人群听觉感知功能减退的高效早期筛查,但针对听觉认知功能的简化方案仍存在缺口,使得后续干预措施^[16]难以实现个体化与精准化。因此,开发一套兼具高效性与可靠性的听觉认知功能简化评估工具,成为推动听觉健康从感知筛查迈向认知评估的关键环节,以减缓认知功能下降和老年心理问题的发生。

本研究采集完整听觉认知功能数据,基于贝叶斯自适应、支持向量机等人工智能算法对评估流程进行科学简化,在保障90%及以上敏感度的条件下,大幅缩短测试耗时并加以验证,旨在提供一套适用于基层社区的听觉认知功能综合评估工具。

2 材料与方法

2.1 研究对象

本研究共纳入两阶段受试者。第一阶段选取35名健康大学生(18~25岁)建立听觉认知测试常模。第二阶段从2025年1~6月九亭镇社区体检人群中招募466名居民,年龄18~55岁,平均年龄 35.7 ± 11.2 岁,男性201人,女性265人,用于模型训练与验证,按2:1划分为训练集与验证集。随机抽取88名受试者,平均年龄 38.3 ± 9.6 岁,男性36人,女性52人进行3个月后的随访,验证重测信度。纳入标准:①年龄18~50岁(社区样本)或18~25岁(常模样本);②母语为中文;③自报无听力障碍或神经系统疾病史;④能够独立完成移动设备操作;⑤同意参与研究并签署知情同意。排除标准:①明确的听力损失史,如中重度感音神经性聋;②神经或精神疾病,如脑卒中、癫痫、抑郁症;③服用影响认知药物;④视听障碍;⑤依从性差(反应遗漏率>20%)。本研究经复旦大学附属华东医院伦理委员会批准(2024K107)。

2.2 方法

本研究采用两阶段设计。首先基于常模样本建立3项测试时间压缩言语、听觉数字广度、听觉Stroop正常参考范围,以95%置信区间下限为功能减退判定阈值。随后在社区样本中采集完整测试数据,训练集采用机器

学习算法简化流程,验证集评估诊断效能。最终构建简化版测试工具,并通过3个月随访验证其临床稳定性。流程如图1。

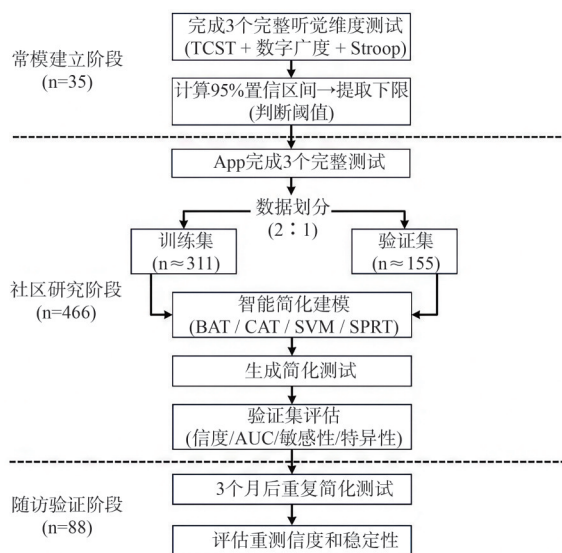


图1 实验流程图

2.2.1 测试环境和设备 所有测试均在相对安静的室内环境中完成(环境噪声<40 dB SPL,受试者主观报告无干扰)。受试者佩戴立体声耳机,通过个人智能终端设备(手机或平板电脑)运行听觉认知评估应用程序完成测试。程序严格遵循经典实验范式设计,复制华东医院听力实验室标准范式。为保证数据一致性,所有音频刺激均采用统一采样率(44.1 kHz)与标准化音量输出(约65 dB SPL)录制。

2.2.2 时间压缩言语测试(time-compressed speech test, TCST)简化 TCST是评估中枢听觉时程处理能力的重要客观范式^[17~19]。完整版测试采用标准化行为学范式^[17],以中文短句为刺激材料^[20],部分句子经词汇替换构建语义异常条件。设置5个压缩率(0%、12.5%、25%、37.5%、50%)。受试者聆听句子并完成语义逻辑判断任务,正式实验共40个随机试次,记录反应准确率与反应时间。

针对冗余试次,本研究采用贝叶斯自适应测试(bayesian analysis toolkit, BAT)^[21]结合最大似然估计(maximum likelihood estimation, MLE)^[22]优化流程。基于受试者实时作答表现动态选择最优压缩率,采用最大信息熵策略优先测试不确定性最高的难度区间,拟合心理物理学函数估计50%识别阈值,阈值估计达预设精度(90%)即自动终止。

2.2.3 听觉数字广度测试(auditory digit span test)简化 听觉数字广度测试主要用于评估短时听觉记忆和工

作记忆容量^[23~26]。完整版测试采用预录标准化语音材料,速率1个/秒,顺背从3位开始,倒背从2~3位开始,每个长度2次机会,连续2次失败终止。受试者通过键盘输入记忆序列,系统自动记录正确率与反应时间。为提高测试效率,本研究引入计算机自适应测试(computerized adaptive testing, CAT)策略^[27],采用自适应阶梯法结合二分逼近逻辑^[28],快速定位个体记忆广度边界。顺背测试以中等难度(如5位数)为起点,根据作答表现动态调整步长(初始 ± 2 ,转折后 ± 1),并在确定“通过 n 位、失败 $n+1$ 位”的临界点后自动终止。该方法通过优先采样能力边缘区,显著减少无效试次。倒背测试采用跨任务锚定自适应算法,以顺背得分(N)作为起点($N-2$)^[23],根据作答表现进行跨步调整(± 2 或下调),以连续2次失败作为终止标准。

2.2.4 听觉 Stroop 测试简化 该测试是神经心理学中评估大脑听觉执行功能和抑制控制的核心范式^[29]。完整版测试包含音量、音高和时长3个声学维度,每个维度设置一致、不一致及中性条件。受试者需忽略语义,仅根据声音的物理特征进行判断并按键反应。系统记录反应准确率与反应时间,完整测试需遍历全部维度与条件,试次数近百。本研究采用支持向量机(support vector machine, SVM)^[30]进行特征筛选,并结合序贯概率比检验(sequential probability ratio test, SPRT)^[31]实现动态终止。基于特征权重分析,剔除区分度较低的中性条件及音高、时长维度,仅保留音量维度下的一致与不一致条件作为核心任务。测试过程中交替呈现两类刺激,实时更新 Stroop 干扰效应均值($RT_{不一致}-RT_{一致}$)及方差;当干扰效应的置信区间达到预设精度或达到最大试次数(20题)时,自动终止测试并输出结果。

2.3 简化版测试的验证和统计方法

在测量一致性方面,采用皮尔逊相关分析及 Bland-Altman 一致性分析评估简化版与完整版测试结果之间的关系。在诊断效能方面,以常模样本 95% 置信区间下限作为分类标准,将受试者分为听觉认知功能正常与功能减退两类,基于此构建受试者工作特征曲线(receiver operating characteristic curve, ROC),对简化版测试的判别能力进行评估。在测试效率方面,通过配对样本 t 检验对耗时进行比较。为进一步验证简化模型的稳定性与临床可重复性,在3个月后将部分受试者进行随访测试。鉴于随访阶段仅实施简化版测试,为保证重复测量比较的严格等价性与科学性,本研究预先通过盲法设定了数据映射提取规则。严格从基线完整版测试数据中提取与最终算法输出的简化版完全相同的试次集合作为等价简化子集。随后,采用双向随机效应模型的组内相关系数

(intraclass correlation coefficient, ICC)评估基线等价简化子集与3个月随访简化版测试结果之间的一致性,从而在消除试次内容差异干扰的前提下,真实客观地反映测试工具在时间维度上的稳定性与重测信度。3项测试的智能简化均使用 Python 编程环境进行算法构建与蒙特卡洛模拟验证。数据统计使用 SPSS 25.0。

3 结果

3.1 描述性统计结果

常值人群($n=35$)与社区人群($n=466$)在听觉认知功能测试结果时间压缩言语测试中,常值人群的压缩率阈值为 $72.2\pm 8.14\%$,社区人群为 $61.3\pm 12.03\%$;听觉数字广度测试中,常值人群的正背广度为 7.5 ± 0.4 个数字串长度,倒背广度为 6.0 ± 0.3 个数字串长度,社区人群的正背广度为 6.4 ± 0.2 个数字串长度,倒背广度为 4.6 ± 0.2 个数字串长度;听觉斯特鲁普测试中,常值人群的干扰效应均值为 99.1 ± 114.3 ms,社区人群为 222.6 ± 187.3 ms。

3.2 时间压缩言语测试简化结果

完整版测试经过简化算法剔除了在极低压缩率(0%)下天花板效应显著的题目,以及在极高压缩率(50%)下呈地板效应的题目。最终提取出15道具有最佳区分效度的题目构成智能简化版时间压缩言语测试,重点保留了25%与37.5%压缩率的题目。简化版内部一致性 Cronbach's α 系数为0.84,表明精简后的题目间仍具有高度的同质性。采用奇偶折半法经 Spearman-Brown 公式校正后,折半信度系数为0.86, $P<0.01$,进一步证实简化版测试在随机误差控制上的稳定性。在155例验证集数据中,简化版时间压缩言语测试的总正确率与完整版呈高度相关($r=0.89$, $P<0.001$),压缩率阈值一致性良好($r=0.86$, $P<0.001$);工作特征曲线 ROC 分析显示其 AUC 为0.85(95%CI: 0.79~0.90),敏感度为90.6%,特异度为80.3%。

对88名被试由2个版本实际测得的压缩率阈值进行相关分析结果显示,两者之间呈现高度正相关, $r=0.88$, $P<0.001$ (图2左)。该结果提示,提取出的15道核心题目能高度代表原40道题目的测量结果,信息损失量在可接受范围($<10\%$)内。配对样本 t 检验结果显示,受试者完成简化版测试的平均总耗时为 3.2 ± 0.5 分钟,显著低于完整版测试的 9.8 ± 1.2 分钟, $t=22.45$, $P<0.001$, Cohen's $d=6.42$ 。基于常值95%置信区间社区人群分为听觉认知功能正常组与功能减退组,ROC 分析简化版测试在识别听觉感知功能减退上的临床诊断效能,计算曲线下面积(AUC)结果显示(图2右),简化版的 AUC 达到0.87(95% CI: 0.81~0.93),表明其具有良好的诊断辨别力。

通过约登指数(Youden index)最大化原则确定最佳截断值为压缩率阈值=55%,在该截断值下,简化版测试的灵敏度为90.7%,特异度为82.4%,而完整版的AUC=0.88,灵敏度93.2%,特异度85.0%。该结果提示简化版的筛查精度仅发生小幅度的衰减,但在社区大规模早筛场景中,其时间成本收益远大于精度衰减。对被试错报率的方差分析结果显示,在50%高压压缩率条件下,完整版测试后半段的错报率显著增加,提示老年被试出现了明显的测试疲劳;而简化版结果中的错报率从完整版的4.5%显著下降至1.1%, $\chi^2=8.76$, $P=0.003$ 。随访一致性分析结果显示,简化版时间压缩言语测试与基线等价子集之间具有良好的重复测量稳定性($ICC=0.87$,95%CI: 0.81~0.92),提示该工具在短期内具有较高的时间一致性与可靠性。

3.3 听觉数字广度测试简化结果

完整版数字广度测试(含顺背与倒背)平均需要执行 14.5 ± 3.2 个测试项,而基于跨任务锚定自适应算法的简化版测试平均仅需执行 6.8 ± 1.5 个测试项, $t=32.45$, $P<0.001$ 。在时间成本上,数字广度评估环节的平均耗时从8.2分钟缩减至3.4分钟(缩减约58.5%)。在155例验证集中,简化模型的整体判别性能保持稳定,其AUC较训练集相差 $\Delta AUC=0.02\sim 0.04$,敏感性与特异性变化 $<5\%$,未见明显过拟合现象,表明模型具有良好的泛化能力与稳健性。

对88名被试由两个版本实际测得的数字广度进行相关分析结果显示,顺背广度简化版与完整版得分显著相关, $r=0.94$, $P<0.001$ (图3左)。进一步Bland-Altman一致性分析显示,96.4%的模拟预测结果均落在95%一致性界限内,最大误差不超过 ± 1 个数字广度。倒背广度简化版与完整版得分亦呈现显著正相关, $r=0.92$, $P<0.001$ 。基于常值95%置信区间社区人群分为听觉认知功能正常组与功能减退组,工作特征曲线(ROC)评估2个版本数字广

度的筛查效能结果显示(图3右),简化版顺背广度得分的AUC为0.86(95%CI: 0.81~0.90),显示出良好的诊断鉴别力。根据约登指数确定的最佳临界值为5,简化版测试的敏感度达到了90.5%符合预期保障90%及以上敏感度的前置条件,特异度为81.2%。与完整版测试的诊断效能(AUC=0.88,敏感度92.1%)相比,两者无统计学显著差异(DeLong's test, $Z=1.05$, $P=0.29$)。倒背测试结果显示,简化版AUC=0.87,敏感度91.2%,完整版AUC=0.89,敏感度93.4%。对全量测试数据的错报率分析发现,在完整版中,经过10个试次后,受试者的无意义按键或超时未响应率上升至8.5%;而在自适应简化版中,由于试次总量控制在8次以内,疲劳相关的超时未响应率被有效控制在1.2%, $\chi^2=15.6$, $P<0.001$ 。随访一致性分析结果显示,简化版听觉数字广度测试与基线等价子集之间具有良好的重复测量稳定性($ICC=0.88$,95%CI: 0.81~0.93),提示该工具在短期内具有较高的时间一致性与可靠性。

3.4 听觉斯特鲁普测试简化结果

利用支持向量机算法对全量测试数据进行特征重要性分析结果表明,在音量、音高与时长3个维度中,音量维度的干扰效应对认知状态分型的基尼不纯度递减贡献最大,特征重要性因子FI=0.62,显著高于音高(FI=0.21)与时长(FI=0.17)。同时,中性测试项的诊断贡献比近乎为0。因此,简化版方案成功将测试维度从 3×3 降至 1×2 (仅保留音量维度的一致与不一致条件),使基础题库量锐减78.5%。原始版测试平均需要执行 85.4 ± 12.1 个测试项,而引入序贯概率比检验动态终止算法的智能简化版,平均仅需 18.6 ± 4.3 个测试项即可达到收敛标准。配对样本 t 检验显示,简化版测试总耗时(3.5 ± 0.4 分钟)显著低于完整版(16.8 ± 1.1 分钟), $t=45.21$, $P<0.001$ 。在155例验证集数据中,简化版Stroop干扰效应与完整版测量结果呈高度相关($r=0.91$, $P<0.001$),一致性分析未见系统性偏

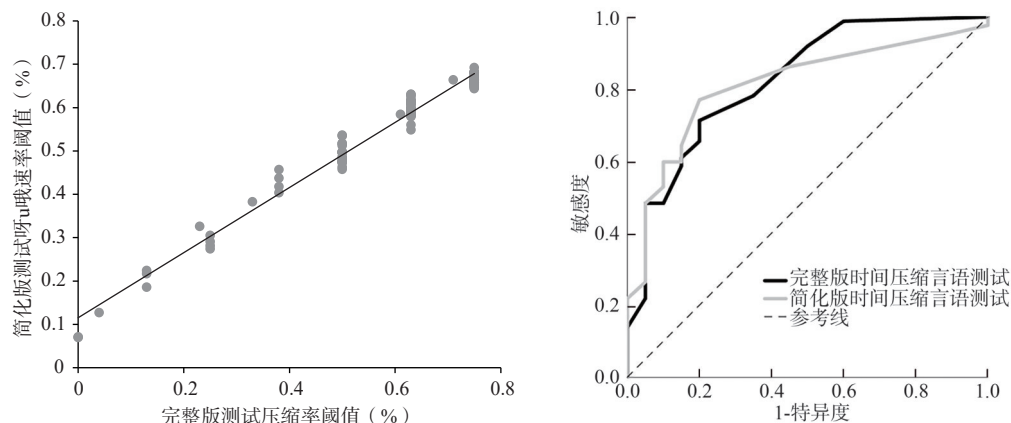


图2 88例随访者简化版与完整版测试压缩率阈值的相关性(左)及完整版和简化版时间压缩言语测试的工作特征曲线(ROC)比较(右)

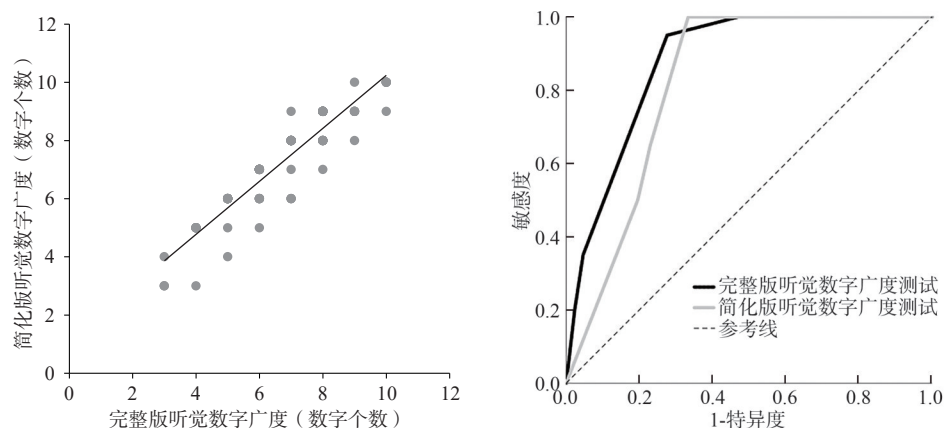


图3 88例随访者简化版与完整版听觉数字广度的相关性(左)及完整版和简化版听觉数字广度的工作特征曲线(ROC)比较(右)

倚;其ROC曲线下面积为0.88(95%CI: 0.83~0.92),敏感度为91.0%,特异度为83.5%。

对88名随访受试者在两个版本实际测试中获得的干扰效应均值进行皮尔逊相关分析结果显示(图4左),两者呈高度正相关, $r=0.93$, $P<0.001$ 。Bland-Altman一致性检验表明,两者的均值差为8.5 ms,且95.8%的样本点位于95%一致性界限内。基于常值95%置信区间社区人群分为听觉认知功能正常组与功能减退组,采用工作特征(ROC)曲线比较两个版本的诊断效果,结果显示,基于简化版 Stroop 干扰量构建的二元分类模型,其曲线下面积(AUC)为0.89(95%CI: 0.84~0.93)(图4右)。在最佳截断值下,简化版测试的敏感度为92.4%(满足预设的>90%敏感度要求),特异度为83.6%。完整版测试结果分析显示,AUC=0.86(95%CI: 0.81~0.90),在最佳截断值下敏感度为90.3%,特异度为84.5%。对错报率的方差分析显示,在完整版测试的后半段,受试者的平均漏按率7.2%;而在简化版的序列中,受试者的漏按率为1.5%,卡方检验

差异显著, $\chi^2=18.3$, $P<0.001$ 。简化版听觉斯特鲁普测试与基线等价子集之间具有良好的重复测量稳定性(ICC=0.86,95%CI: 0.80~0.92),提示该工具在短期内具有较高的时间一致性与可靠性。

为验证采用青年常模界值判别社区人群听觉认知功能减退的稳健性,本研究补充进行了敏感性分析。将听觉认知功能减退的分类界值从常模样本的95%置信区间下限调整为第10百分位数界值。调整后的界值分别为:听觉数字广度顺背界值=5、倒背界值为4;听觉 Stroop 测试干扰效应均值能力界限对应为245.4 ms;时间压缩言语测试压缩率阈值为42.0%。在新的分类标准下重新评估简化版模型在验证集($n=155$)中的诊断效能。结果显示,3项简化测试的AUC波动范围为0.01~0.03,时间压缩言语AUC波动于0.84~0.86;数字广度AUC波动于0.85~0.88;听觉 Stroop AUC波动于0.87~0.89,且各测试的敏感度均稳定保持在88.5%以上。

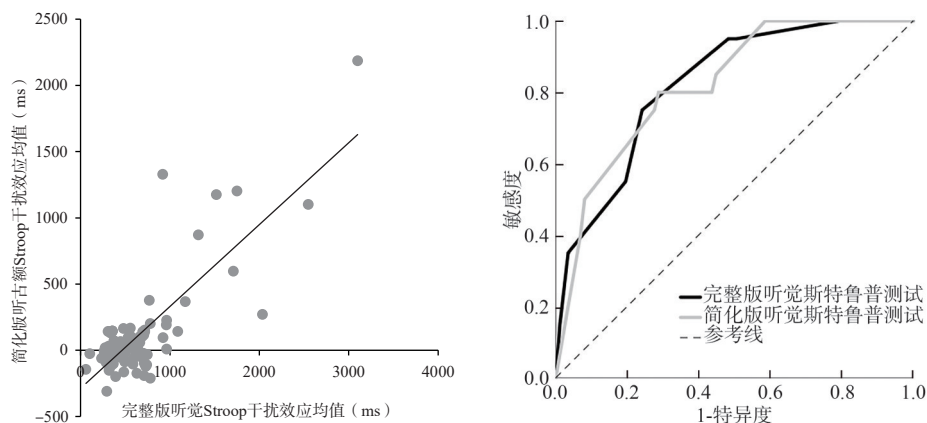


图4 88例随访者简化版与完整版听觉 Stroop 干扰效应均值的相关性(左)及完整版和简化版 Stroop 干扰效应均值的工作特征(ROC)曲线比较(右)

3.5.3 项测评综合耗时

受试者基线完整版测评平均总耗时达 34.8 ± 8.7 分钟,分别为 9.8 ± 3.4 分钟、 8.2 ± 6.1 分钟和 16.8 ± 7.7 分钟。智能简化流程后,单例受试者完成三项核心测试的平均总测评时间显著下降至 10.1 ± 4.3 分钟,分别为 3.2 ± 1.1 分钟、 3.4 ± 1.5 分钟和 3.5 ± 0.8 分钟。综合耗时整体缩减了71.0%。

4 讨论

本研究构建验证了基于智能算法的多维度简化测试体系。结果显示,经过贝叶斯自适应、跨任务锚定及支持向量机特征降维等算法优化后,听觉核心认知测试在总耗时削减71%的同时,各项诊断指标AUC均维持在0.85以上,敏感度稳定在约90%。表明复杂的认知评估可通过信息驱动的自适应策略在最小化信息损失的前提下实现高效测量^[13,27]。

时间压缩言语测试简化后保留的15个核心试次仍然能够代表原始40题测试的信息结构。算法成功捕捉了个体在时间压缩条件下的核心加工能力,而未引入系统性偏差^[17,18]。Cronbach's α 系数及折半信度进一步证明,精简后的测试在内部一致性与随机误差控制方面依然稳健。简化策略并非简单删减题目,而是基于区分度与信息量进行结构性优化。低压缩率条件下的“天花板效应”题目及高压缩率下的“随机作答噪音”题目被系统性剔除,使得测试更聚焦于个体能力的临界区间。这与心理物理学中阈值测量的原理一致,解释了为何在题量显著减少的情况下仍可维持高相关性。与此同时,测试时间从近10分钟缩短至约3分钟,这一效应量在行为科学研究中属于极大效应,表明简化策略不仅具有统计学意义,也具有明确的实际应用价值。

听觉数字广度测试简化版平均试次数量减少超过50%,但与完整版得分保留了高度的相关性,说明通过动态逼近个体能力边界,可以有效避免传统线性测试中大量无信息试次的累积。尤其是倒背测试中引入顺背成绩作为锚点的设计,本质上利用了工作记忆结构中的已知认知规律,从而减少了重复测量带来的冗余负担^[26]。诊断效能方面,简化版AUC达到0.86,且与完整版无显著差异,表明该算法不仅在测量层面保持一致,在分类判别层面同样具备可靠性。疲劳相关错误率显著下降,提示传统测试后期的数据可能已受到非认知因素,如注意力下降、动机降低的污染,而简化版在一定程度上去除了疲劳噪音,从而提高了测量的纯度。

听觉Stroop测试的简化结果也体现了特征筛选+动态终止策略的优势。通过支持向量机对特征重要性的评

估,发现音量维度在区分认知状态方面具有最高贡献,而音高与时长维度贡献有限^[30]。这一发现不仅具有方法学意义,也为理解听觉执行功能的敏感指标提供了新的实证依据^[29]。在此基础上进行降维,使测试结构从 3×3 条件压缩至 1×2 核心对抗范式,大幅降低了任务复杂度。结合序贯概率比检验实现动态终止后,试次数量减少约78%,而干扰效应指标与完整版仍保持高度一致。简化版AUC略高于完整版,提示在排除冗余维度与疲劳影响后,测试的判别能力反而得到优化。说明在复杂认知任务中,增加测试维度并不必然提升诊断效能,反而可能引入测试噪音。

综合3项测试结果可以发现,简化模型在多个维度上达到了预期目标。在测量一致性方面,所有核心指标均表现出高度相关性($r > 0.88$),说明信息压缩并未显著改变个体能力排序;在诊断效能方面,AUC均维持在0.85以上,敏感度基本达到或超过预设的90%阈值,满足筛查工具对高敏感性的要求;再次,在时间效率方面,总测试时长由原约50分钟显著缩短至约8分钟内,使其具备在社区大规模推广的可行性;在依从性与疲劳控制方面,错误率显著下降,提示用户体验得到实质性改善。这种多维度一致性支持了本研究提出的智能简化评估框架的可行性,即通过算法驱动的动态测量,在保持诊断能力的同时显著降低实施成本。

从理论层面看,本研究进一步强化了听觉认知功能的多维结构模型。时间压缩言语测试主要反映时间处理与语音整合能力,数字广度测试反映短时存储与工作记忆容量,而听觉Stroop任务则反映执行控制与抑制能力。三者共同构成了相对独立但相互关联的认知系统^[4]。本研究通过移动端平台将三者整合,并验证其在真实社区环境中的可操作性,为听觉认知功能的生态化评估提供了新的路径。此外,基于35名健康青年建立的95%置信区间下限作为判别标准,虽然在方法上提供了清晰的分类依据,但也隐含了一个重要假设,即青年样本可作为最佳功能状态参考。这一假设在一定程度上合理,但在跨年龄推广时仍需谨慎解读。

本研究亦存在若干局限。首先,建立常模参考的样本量较小($n=35$)且局限于青年群体,将青年最佳状态作为广年龄跨度人群的单一判定阈值可能引入老化与教育偏倚,未来需构建分层常模。其次,移动设备声学呈现差异未引入硬件校准或补偿算法。第三,未纳入纯音听阈等客观听力学指标作为协变量,无法完全量化剥离外周听力对中枢加工的相对贡献。在老龄化背景下,该智能简化工具兼具高效性与可及性,有望作为数字健康平台的关键模块,未来可结合多模态数据深入验证其临床应

用价值。

在当前老龄化背景下,听觉功能减退的早期识别与干预已成为认知健康管理的重要环节^[8]。本研究提出的智能简化评估工具,兼具高效性、可及性与较高的诊断性能,有望作为基层社区筛查与数字健康平台中的关键模块,为后续个体化干预^[16],如听觉训练、助听器验配或认知干预提供客观依据。未来研究可进一步结合纵向随访与多模态数据(如语噪测试或脑电指标),探索其在预测听觉认知衰退及干预效果评估中的应用价值。

参考文献

- [1] 张敏,王姣锋,傅蕾,等.听配能的行为学和自陈量表测试方法及其在老年人群中的应用[J].中国听力语言康复科学杂志,2020,18(4):257-261.
- [2] 张敏,Catherine VP.听配能(listening effort)调控研究与临床听觉康复[J].中国听力语言康复科学杂志,2019,17(1):1-8.
- [3] 张敏,Catherine V.Palmer.老年人听配能调控行为初探与启发[J].中国听力语言康复科学杂志,2019,17(5):327-331.
- [4] 张娟,商嘉琪,聂帅.老年性聋听觉系统及认知功能改变[J].中华耳科学杂志,2023,21(2):194-199.
- [5] Zhou Y, Han H, Zhang X, et al. Dual improvement of cognitive function and auditory ability in elderly patients with hearing impairment by transcranial direct current stimulation-assisted auditory rehabilitation training[J]. Frontiers in Aging Neuroscience, 2025, 17: 159-196.
- [6] 梁巍.提升全民健康素养促进主动听力健康——写在第24次全国“爱耳日”之际[J].中国听力语言康复科学杂志,2023,21(2):113-116.
- [7] 傅鑫萍,段吉茸,张敏.老年人听力筛查现状分析及其研究进展[J].中国听力语言康复科学杂志,2018,16(3):205-208.
- [8] Livingston G, Huntley J, Liu KY, et al. Dementia prevention, intervention, and care: 2024 report of the Lancet standing Commission [J]. The lancet, 2024, 404(10452): 572-628.
- [9] Merten N, Boenniger MM, Herholz SC, et al. The associations of hearing sensitivity and different cognitive functions with perception of speech-in-noise[J]. Ear and Hearing, 2022, 43(3): 984-992.
- [10] Dawes P, Munro KJ. Hearing loss and dementia: Where to from here? [J]. Ear and hearing, 2024, 45(3): 529-536.
- [11] 张多米,王凯.老年性聋患者的认知功能评估[J].中国听力语言康复科学杂志,2025,23(2):185-189.
- [12] Choi JS, Adams ME, Crimmins EM, et al. Association between hearing aid use and mortality in adults with hearing loss in the USA: a mortality follow-up study of a cross-sectional cohort[J]. The Lancet Healthy Longevity, 2024, 5(1): e66-e75.
- [13] Zhang M, Bi Z, Fu X, et al. A parsimonious approach for screening moderate-to-profound hearing loss in a community-dwelling geriatric population based on a decision tree analysis[J]. BMC geriatrics, 2019, 19(1): 214.
- [14] 朱蓓,段吉茸,袁春兰,等.基于决策树智能算法的2步纯音听力筛查在社区老年人中的应用[J].听力学及言语疾病杂志,2021,29(2):130-134.
- [15] 孙凯,朱蓓,杨建宏,等.中文版老年听力障碍筛查量表简化潜力研究[J].中国听力语言康复科学杂志,2023,21(2):149-153.
- [16] Lin FR, Pike JR, Albert MS, et al. Hearing intervention versus health education control to reduce cognitive decline in older adults with hearing loss in the USA (ACHIEVE): a multicentre, randomised controlled trial[J]. Lancet, 2023,402(10404):786-797.
- [17] Dias JW, Mccliskey CM, Harris KC. Time-Compressed Speech Identification Is Predicted by Auditory Neural Processing, Perceptuomotor Speed, and Executive Functioning in Younger and Older Listeners[J]. Journal of the Association for Research in Otolaryngology, 2019,22(11):73-88.
- [18] Nourski KV, Reale RA, Oya H, et al. Temporal envelope of time-compressed speech represented in the human auditory cortex[J]. The Journal of Neuroscience, 2009, 29(49): 15564-15574.
- [19] Massoud C, El Khoury-Malhame M. Normal speed versus time-compressed speech in long-term memory retention: a randomized, within subjects experimental design[J]. BMC Psychology, 2026, 14(1): 223-241.
- [20] 郑海昌,王枫,郝昕,等.安静条件下语句识别表的等价性研究[J].中国听力语言康复科学杂志,2011,9(5):16-18.
- [21] Van der Linden WJ. Bayesian item selection criteria for adaptive testing [J]. Psychometrika, 1998, 63(2): 201-216.
- [22] Myung IJ. Tutorial on maximum likelihood estimation[J]. Journal of mathematical Psychology, 2003, 47(1): 90-100.
- [24] Wechsler D. Wechsler adult intelligence scale-Fourth Edition (WAIS-IV)[J]. San Antonio, TX: NCS Pearson, 2008, 22(498): 1-132.
- [25] 张敏, Catherine V.Palmer. 语言理解易度模型与听配能[J]. 中国听力语言康复科学杂志, 2018, 16(3): 226-231.
- [26] Egeland J, Lund O, Raudeberg R. Measuring working memory span with WAIS-IV: Digit sequence is the superior span test[J]. Applied Neuropsychology: Adult, 2026, 33(1): 65-72.
- [27] Bjorner JB. Computerized adaptive testing: A primer (2nd Edition) [J]. Quality of Life Research, 2001. 10(8): 733-734.
- [28] Kaernbach C. Adaptive threshold estimation with unforced-choice tasks [J]. Perception & psychophysics, 2001, 63(8): 1377-1388.
- [29] Kestens K, Degeest S, Miatton M, et al. An auditory Stroop test to implement in cognitive hearing sciences: development and normative data[J]. International Journal of Psychological Research, 2021, 14(2): 37-45.
- [30] Zhu J, Rosset S, Tibshirani R, et al. l-norm support vector machines [J]. Advances in neural information processing systems, 2003, 20(16): 112-129.
- [31] Silva IR, Montalban J, de Oliveira FLP. Maximized sequential probability ratio test regression[J]. Biometrics, 2025, 81(4): 170-181.

收稿日期 2026-04-14

修回日期 2026-05-20

录用日期 2026-06-30

责任编辑 蒋 春