口腔初诊患者人工智能智慧问答系统使用体验量表的研制与心理测量学验证

刘睿夕，匡碧芬，杨依凡，李宏，赵娅琼，冯云枝，袁临天，郭玥

(中南大学湘雅二医院口腔科，长沙 410011)

[摘 要] **目的：**口腔健康是全身健康的重要组成部分，然而口腔初诊患者使用人工智能智慧问答系统场景下尚缺乏专用测评工具。本研究旨在通过系统研制并验证适用于中国口腔科初诊患者的人工智能智慧问答系统使用体验量表(User Experience Scale for Artificial Intelligence Question-Answering Systems，UES-AIQA)，为口腔门诊人工智能(artificial intelligence，AI)服务质量评估提供标准化工具。**方法：**通过文献分析与专家小组会议构建初始条目池，经2轮德尔菲专家函询完成条目筛选。采用方便抽样法选取305例口腔科初诊患者，第1批233例用于探索性因子分析(exploratory factor analysis，EFA)，第2批72例用于验证性因子分析(confirmatory factor analysis，CFA)，综合采用项目分析、因子分析、信度分析及效度分析评价量表的心理测量学属性。**结果：**最终量表包含10个条目，呈单维度结构。各条目水平的内容效度指数(item-level content validity index，*I*-CVI)为0.900~1.000，量表水平的平均内容效度指数(scale-level content validity index/average，*S*-CVI/Ave)为0.980。平行分析与Kaiser准则均支持提取1个公因子，单因子累积方差贡献率为83.593%，各条目标准化因子载荷为0.879~0.937。CFA拟合结果显示卡方自由度比(ratio of chi-square to degrees of freedom，χ^2/df)为3.190，比较拟合指数(comparative fit index，CFI)为0.929，Tucker-Lewis指数(Tucker-Lewis index，TLI)为0.900，标准化均方根残差(standardized root mean square residual，SRMR)为0.035，主要拟合指标均达到可接受标准；近似误差均方根(root mean square error of approximation，RMSEA)为0.174；各条目标准化因子载荷为0.845~0.928。总量表克龙巴赫α系数为0.981，McDonald's ω系数为0.983，Spearman-Brown分半信度为0.970，重测组内相关系数(intra-class correlation coefficient，ICC)为0.830。**结论：**UES-AIQA内容效度与信度达到心理测量学标准，结构效度部分指标达到可接受水平，可作为评估口腔科初诊患者AI智慧问答系统使用体验的标准化工具。

[关键词] 人工智能；门诊患者；智慧问答系统；量表编制；信度；效度

Development and psychometric validation of the User Experience Scale for Artificial Intelligence Question-Answering Systems among first-visit dental patients

LIU Ruixi, KUANG Bifen, YANG Yifan, LI Hong, ZHAO Yaqiong, FENG Yunzhi, YUAN Lintian, GUO Yue

(Department of Stomatology, Second Xiangya Hospital, Central South University, Changsha 410011, China)

收稿日期(Date of reception)：2026-04-22

第一作者(First author)：刘睿夕，Email: Liurx211@163.com, ORCID: 0009-0007-0744-6784；匡碧芬，Email: zndxxyeyykbf@csu.edu.cn, ORCID: 0009-0006-1990-3789

通信作者(Corresponding author)：郭玥，副研究员，Email: guoyue@csu.edu.cn, ORCID: 0000-0003-2403-2034

基金项目(Foundation item)：湖南省自然科学基金(2025JJ50687，2024JJ9209)；中南大学中央高校基本科研业务(2024ZZTS0891)。This work was supported by the Natural Science Foundation of Hunan Province (2025JJ50687, 2024JJ9209) and the Fundamental Research Funds for the Central Universities of Central South University (2024ZZTS0891), China.

开放获取(Open access)：本文遵循知识共享许可协议，允许第三方用户按照署名-非商业性使用-禁止演绎4.0(CC BY-NC-ND 4.0)的方式，在任何媒介以任何形式复制、传播本作品(https://creativecommons.org/licenses/by-nc-nd/4.0/)。

©Journal of Central South University (Medical Science). All rights reserved.

ABSTRACT

Objective: Oral health is an essential component of overall health; however, specific assessment tools for first-visit patients in dental outpatient waiting scenarios are currently lacking. This study aims to systematically develop and validate the User Experience Scale for Artificial Intelligence Question-Answering Systems (UES-AIQA) for Chinese first-visit dental patients, providing a standardized instrument for evaluating the service quality of artificial intelligence (AI)-based services in dental outpatient settings.

Methods: An initial item pool was developed through literature analysis and expert group discussions. Item selection was conducted through 2 rounds of Delphi expert consultation. A total of 305 first-visit dental patients in the Department of Stomatology were recruited using convenience sampling. The first sample of 233 participants was used for exploratory factor analysis (EFA), and the second sample of 72 participants was used for confirmatory factor analysis (CFA). Psychometric properties of the scale were evaluated through item analysis, factor analysis, and reliability analysis.

Results: The final scale consisted of 10 items and demonstrated a unidimensional structure. The item-level content validity index (*I*-CVI) ranged from 0.900 to 1.000, and the average scale-level content validity index (*S*-CVI/Ave) was 0.980. Parallel analysis and the Kaiser criterion both supported the extraction of one common factor. The cumulative variance contribution rate of the single factor was 83.593%, and standardized factor loadings of all items ranged from 0.879 to 0.937. CFA showed that the ratio of chi-square to degrees of freedom (χ^2/df) was 3.190, the comparative fit index (CFI) was 0.929, the Tucker-Lewis index (TLI) was 0.900, and the standardized root mean square residual (SRMR) was 0.035, with the major fit indices reaching acceptable levels. The root mean square error of approximation (RMSEA) was 0.174. Standardized factor loadings in the CFA model ranged from 0.845 to 0.928. The Cronbach's α coefficient of the total scale was 0.981, McDonald's ω coefficient was 0.983, Spearman-Brown split-half reliability coefficient was 0.970, and the test-retest intra-class correlation coefficient (ICC) was 0.830.

Conclusion: The UES-AIQA demonstrated satisfactory content validity and reliability according to psychometric standards, while structural validity indicators reached acceptable levels. The scale may serve as a preliminary standardized tool for evaluating user experiences with AI question-answering systems among first-visit dental patients.

KEY WORDS　artificial intelligence; outpatients; intelligent question-answering system; scale development; reliability; validity

口腔健康是全身健康的重要组成部分。国家卫生健康委员会办公厅于2019年印发《健康口腔行动方案(2019—2025年)》[1]，明确提出提升全人群、全生命周期口腔健康服务能力，强调创新口腔健康教育载体与信息服务形式。初诊患者对病情、诊疗流程及注意事项普遍缺乏充分了解，候诊阶段易产生焦虑情绪，既往研究[2]显示，牙科焦虑水平较高的患者满意度较低且等待时间是影响牙科服务满意度的重要因素之一。而医师在有限接诊时间内需优先完成诊断，难以充分提供个体化信息支持[3-5]。传统服务模式在信息可及性与个体化方面的结构性不足[6]。因此，亟须新型信息支持工具[如人工智能(artificial intelligence，AI)智慧问答系统]弥补这一缺口。

近年来，AI在医疗卫生领域的应用获得国家政策的持续推动。2024年11月，国家卫生健康委员会、国家中医药管理局、国家疾病预防控制局联合印发《卫生健康行业人工智能应用场景参考指引》[7]，明确了包括患者服务、健康咨询在内的84类AI应用方

©Journal of Central South University (Medical Science). All rights reserved.

向；2025年10月，国家卫生健康委员会等5部门进一步联合印发《关于促进和规范“人工智能+医疗卫生”应用发展的实施意见》[8]，将患者就诊智能服务列为重点推进方向，明确要求到2027年实现患者就诊智能服务在医疗卫生机构的广泛应用。在技术层面，基于自然语言处理(natural language processing，NLP)与大语言模型(large language model，LLM)构建的AI智慧问答系统已在多个医疗场景中展现出应用价值。在患者教育、诊断决策辅助及临床流程优化等方面均有广泛探索[9-13]，在口腔医学领域，AI技术已被应用于影像辅助诊断及牙髓病学教育[14]，然而将AI智慧问答系统应用于口腔初诊候诊场景以改善患者信息获取与就诊体验的研究目前较少，相关标准化测评工具亦有待开发。

用户体验是决定医疗信息系统采纳率与持续使用意愿的核心因素，直接影响患者满意度、依从性乃至健康结局[15-16]。现有医疗AI用户体验研究主要依赖3类工具：1)通用可用性量表，如系统可用性量表(System Usability Scale，SUS)，其条目设计未涵盖病情解释准确性、情绪安抚效果等医疗特有元素[17]；2)通用技术接受度问卷，其对口腔初诊场景的专科需求覆盖不足；3)少数面向医疗场景的专科量表[如电子健康素养量表(The eHealth Literacy Scale，eHEALS)]，虽具备一定场景适用性，但未涉及AI智慧问答系统的专业信息质量与个性化服务维度[18]。上述工具均难以全面评估口腔初诊患者在AI辅助候诊场景下的核心体验，包括病情解释准确性、就医焦虑缓解效果、个性化建议适配性及后续推荐意愿等关键维度。此外，经对PubMed、万方数据知识服务平台及中国知网数据库的系统检索(共检索到261篇，筛选后纳入5篇相关研究)，现有研究虽已涉及口腔AI工具的患者可用性评估[19-20]及AI智慧问答系统的患者满意度报告[21-23]，但均未开发针对患者使用体验的专用量表，亦未进行系统性心理测量学验证，口腔初诊候诊场景下AI智慧问答系统使用体验的量化测评工具仍有待开发。

因此，本研究以DeLone & McLean信息系统成功模型(D&M模型)及技术接受模型(technology acceptance model，TAM)为理论框架，旨在研制具有口腔专科特异性的AI智慧问答系统使用体验量表(User Experience Scale for Artificial Intelligence Question-Answering Systems，UES-AIQA)。具体目标如下：在文献分析与专家小组会议基础上构建初始条目池；通过预调查对条目可读性进行验证，经2轮德尔菲专家函询完成条目筛选，并基于2批独立正式调查样本，综合运用项目分析、探索性因子分析(exploratory factor analysis，EFA)及验证性因子分析(confirmatory factor analysis，CFA)，对量表的信度(如克龙巴赫α系数)与效度进行系统检验，确保量表具备稳健的测量学性能，为口腔门诊AI服务的质量监测与持续优化提供标准化、量化的评价工具。

1 对象与方法

1.1 伦理声明

本研究已获得中南大学湘雅二医院临床研究伦理委员会批准(审批号：LYEC2025-0087)。

1.2 对象

1.2.1 预调查对象

2025年6月，采用方便抽样选取中南大学湘雅二医院口腔科(以下简称“本科”)初诊患者20例进行预调查，收集其对条目表述的意见。纳入标准：1)有口腔诊治需要；2)具有阅读理解能力；3)同意参与本研究。排除标准：1)存在身心障碍；2)本科复诊。根据患者对条目表述的反馈意见，对量表进行修订，形成调查版量表。

1.2.2 调查对象

问卷采取患者自填方式，由10名医护人员在本科分诊台向候诊患者发放问卷星二维码链接，患者使用个人手机独立扫码作答，医护人员不介入填写过程，以减少诱导性偏倚。问卷星平台设置同一IP地址仅可提交1次，以控制重复作答。纳入和排除标准同预调查。EFA的样本量应为量表条目数的5~10倍[24]，本研究调查版问卷包含10个条目，鉴于电子问卷不存在纸质问卷的物理丢失，据预调查经验估计此类情况发生率约为10%，据此在按条目数估算的最低样本量基础上上浮10%，作为目标样本量估算依据(计算样本量为111例)。正式调查阶段采用“强制完成”模式，确保提交问卷的完整性，故未对正式调查的丢失率进行单独统计。为避免数据过拟合，本研究严格采用2批独立的样本进行分析。第1次调查收集有效问卷233份(样本1)用于EFA；随后在同质人群中再次独立收集有效问卷72份(样本2)用于CFA。依据Nunnally等[25]在1994年提出的每个测量指标对应5~10个样本的标准，本研究10个条目所需最低样本量为50~100例，72例满足该要求。此外，Wolf等[26]的模拟研究亦表明，当模型结构简单且因子载荷较高(≥0.700)时，小样本(n≥60)可获得稳定拟合结果，本研究所有条目λ≥0.845，说明样本量合理。

©Journal of Central South University (Medical Science). All rights reserved.

1.3 方法

1.3.1 理论基础

本研究基于D&M模型和TAM双重理论框架，结合口腔科初诊患者分诊咨询的实际需求与AI智慧问答系统使用体验的核心要素(包括信息质量、系统质量、感知有用性、感知易用性及使用意向等维度)构建初始条目池，形成“专业性与实用性”“交互体验”“系统体验与行为意向”3个维度(包含10个条目)的初版量表。

1.3.2 文献分析

以“人工智能/智慧医疗/智能问答系统”“口腔/初诊/门诊分诊”“满意度/用户体验/技术接受度”为检索词，检索中国知网、中国生物医学文献服务系统、PubMed和Web of Science等数据库，收集相关文献并阅读，提取与患者人机交互体验、分诊准确性感知及使用意愿相关的内容，同时借鉴D&M模型与TAM相关理论维度，形成量表条目池。

1.3.3 专家小组会议

邀请5名专家针对问卷条目结构合理性、条目确立及筛选、条目表述明确性等作出评定。讨论过程做好记录。讨论结束后，结合专家意见对问卷进行讨论与修改，初步形成包含3个维度、10个条目的函询量表，其中，专业性与实用性4个条目(P1~P4)、交互体验3个条目(I1~I3)、系统体验与行为意向3个条目(R1~R3)。

1.3.4 德尔菲专家函询

于2025年5月至6月通过面对面发放邀请函，共纳入本科15名与本研究密切相关的领域(以口腔临床为主)专家。当函询专家意见趋于一致时，终止函询。专家纳入标准：1)本科及以上学历；2)从事过医学等相关专业工作；3)对本研究有一定的积极性。专家权威程度(C_r)是衡量咨询结果可靠性的关键指标，由学术造诣系数、判断依据系数(C_a)及熟悉程度系数(C_s)共同决定，一般认为C_r≥0.700为可靠。函询问卷包括前言、量表评价、专家一般情况及判断依据4个部分。条目重要性评分采用Likert 5级评分法，从“很不重要”到“很重要”依次计1~5分，也可说明修改或删除意见。专家意见的协调程度采用肯德尔和谐系数(Kendall’s *W*)检验。结合专家提出的定性修改意见，研究组对问卷条目进行了实质性修订与调整，并开展第2轮函询。以条目重要性赋值均数≥4.0分、变异系数(coefficient of variation，*CV*)≤0.2作为筛选标准，肯德尔系数显著提高且*P*<0.05，结合专家意见趋同后终止函询[27]。综合专家意见及研究小组讨论结果，对条目进行补充、删除及修改。

1.3.5 量表信效度检验

1.3.5.1 调查工具

包括基本信息、使用情况调查、用户体验度评估与深度反馈与建议。基本信息包括性别、年龄、学历、职业等。使用情况调查包括就诊目的、AI使用情境偏好、AI建议与医务人员反馈一致性等。用户体验度评估共10个条目，采用Likert 5级计分法(1=非常不同意，5=非常同意)，涵盖理解准确性、回答易懂性、指导有效性、专业性、个性化、缓解焦虑、流畅性、界面清晰度、再使用及推荐意愿。

1.3.5.2 资料收集方法及质量控制

本研究组建了调查小组，并对调查员进行标准化培训，确保其熟练掌握问卷填写指南及注意事项，调查员于门诊候诊区选取符合纳入和排除标准的初诊患者。调查采用问卷星电子平台，通过平板电脑或手机进行现场实测，系统后台预设为“强制完成”模式以避免缺项。数据收集完毕后，导出原始数据库，由2名研究人员进行独立的数据清洗。清洗标准包括剔除作答时间过短的问卷。最终数据经双人核对确认无误后锁定。

1.3.5.3 条目质量分析

为确保量表的鉴别力与敏感度，本研究依据经典测量理论(classical test theory，CTT)，采用以下4种统计学指标对初版问卷条目进行筛选[28]。当某一条目同时满足以下2种及以上方法的剔除标准时，予以删除：1)临界比值法(决断值法)。将所有被试的量表总分按降序排列，界定总分最高的前27%为高分组，最低的后27%为低分组。采用独立样本*t*检验比较2组在各条目得分上的差异。若某条目在2组间的差异无统计学意义(*P*>0.05)，提示该条目无法有效区分不同体验水平的受访者，应予删除。2)相关系数法。计算各条目的校正项总计相关系数(corrected item-total correlation，CITC)。依据心理测量学标准，若条目与总分的相关系数*r*<0.40或差异无统计学意义(*P*>0.05)，表明该条目与量表整体构念的一致性较差，应予剔除。3)*CV*法。计算各条目的*CV*，公式为*CV*=($\bar{x}$/*s*)×100%。该指标反映条目的敏感程度，若*CV*过低，需结合其他指标综合判定。4)克龙巴赫α系数法。计算“删除该条目后的克龙巴赫α系数”。若删除某条目后，量表整体的系数数值不仅未下降反而出现升高，说明该条目降低了量表的内部一致性信度，应予删除[25]。

1.3.5.4 信度检验

为验证量表测量结果的精准度与稳定性，本研究采用以下3种统计学指标进行综合评价[18]：1)内部

©Journal of Central South University (Medical Science). All rights reserved.

一致性信度。采用克龙巴赫α系数及McDonald's ω系数作为核心评价指标，测算总量表的同质性水平。依据心理测量学通用标准，当克龙巴赫α系数≥0.700时，判定量表内部一致性达到可接受水平。二者差值<0.020时，表明τ等价假设成立，α系数报告可信。2)分半信度。采用折半法(split-half method)检验量表的内在稳定性。将量表条目按题目顺序均分为2部分，计算2部分得分的相关性，并用Spearman-Brown公式进行校正，得出最终分半信度系数。3)重测信度。在首轮调查结束后的第3周，随机抽取20例曾参与调查的初诊患者进行再次施测，采用组内相关系数(intra-class correlation coefficient，ICC)评估量表的跨时间稳定性。4)亚组信度分析。为初步评估量表在不同人群中的测量稳健性，按年龄(≤45岁 vs ≥46岁)、学历(高中及以下 vs 大专及以上)及性别(男性 vs 女性)对EFA样本(n=233)进行分组，分别计算各亚组的克龙巴赫α系数及条目因子载荷范围，以验证量表的跨人群适用性。

1.3.5.5 *效度检验*

内容效度：本研究组建由15名专家构成的评审小组，对初版量表中各条目与其所属内容领域的契合程度进行专业评定。基于专家评分结果，采用条目水平的内容效度指数(item-level content validity index，*I*-CVI)及量表水平的平均内容效度指数(scale-level content validity index/average，*S*-CVI/Ave)作为核心评价指标，并计算调整后的Kappa值评价专家意见的一致性[27]。参照内容效度评价的相关测量学标准，量表的*I*-CVI≥0.780、*S*-CVI/Ave≥0.800，且调整后的Kappa值≥0.740，提示该量表内容效度处于可接受水平。

结构效度：本研究采用EFA[29]与CFA[30]相结合的方法，全面检验量表的结构效度。1)EFA：通过Kaiser-Meyer-Olkin(KMO)样本适切性检验与Bartlett球形检验来评估数据是否适合进行因子分析。依据统计学标准，若KMO值>0.700且Bartlett检验的P<0.05，则视为数据结构适切。在因子数确定阶段，采用平行分析(1 000次蒙特卡洛模拟)作为主要判断依据，同时参考Kaiser准则(特征根>1)及碎石图。当实际数据特征根超过随机模拟数据对应值时，保留该因子。采用最大似然法(maximum likelihood，ML)进行因子提取。条目删减标准设定为：条目因子载荷<0.400，或存在明显的双重载荷现象(即在2个及以上因子上的载荷值均>0.400且差值<0.200)。2)CFA：基于EFA结果，采用ML对单维度模型进行拟合度检验，并根据修正指数(modification indices，MI)释放具有理论依据的条目残差相关。模型拟合参考标准：卡方自由度比(ratio of chi-square to degrees of freedom，χ^2/df)<5.000；近似误差均方根(root mean square error of approximation，RMSEA)<0.080；比较拟合指数(comparative fit index，CFI)>0.900；标准化均方根残差(standardized root mean square residual，SRMR)<0.080。聚合效度评价采用标准化因子载荷(λ)作为判断指标，λ>0.700提示聚合效度良好。

1.4 统计学处理

采用R软件(版本4.6.0)对数据进行分析，主要使用base R内置函数进行项目分析、信度分析及EFA，CFA使用lavaan包(版本0.6.21)实现。数据描述遵循以下原则：对于性别、学历、职业等计数资料以频数(构成比)表示；对于量表得分等计量资料，在满足正态分布的前提下，以均数±标准差表示。所有统计检验均为双侧检验，P<0.05为差异具有统计学意义。

2 结 果

2.1 德尔菲专家函询结果

2.1.1 *专家的一般资料*

共纳入本科的15名与本研究密切相关的领域专家，其中口腔临床专家8名(53.33%)，医疗信息系统/数字健康专家3名(20.00%)，卫生统计学/心理测量学专家2名(13.33%)，护理学专家1名(6.67%)，人机交互设计专家1名(6.67%)，男女比4∶11；专业年限集中在5~20年；其中本科4名，硕士3名，博士8名；高级职称3名，中级职称5名，初级职称7名。

2.1.2 *专家的积极程度和权威程度*

共开展了2轮德尔菲专家函询：第1轮函询共发放问卷15份，回收有效问卷15份，有效回收率为100%；基于首轮的反馈修订后，第2轮函询向上述15名专家再次发放问卷，予以全部回收(15份)，有效回收率达100%。

统计分析显示，专家对问卷内容的判断依据系数为0.920，熟悉程度系数为0.840。经计算，专家的C_r为0.880。该数值远高于可接受标准。

第1轮函询结果显示专家意见尚存分歧(Kendall's W=0.099，χ^2=13.191，P=0.154)，专家组对首轮指标体系的评分尚未达到统计学意义上的一致，专家意见离散度较高。经过条目重构与意见整合后，第2轮函询的Kendall's W显著提升至0.607(P<0.001)。

©Journal of Central South University (Medical Science). All rights reserved.

2.1.3 专家函询结果

依据专家的定性反馈与定量评分，对问卷初稿进行了以下实质性修订。内容增补：8名专家指出，原问卷中涉及“就医决策”及“反馈”等强书面化表达的描述，超出了普通初诊患者的理解能力；另有3名专家提出部分条目表述冗余，建议精简语言表达。据此，我们调整了相关表述。例如，将“AI提供的建议对我做出就医决策有帮助”修改为更通俗的“AI能有效指导我进行初步就医判断”；同时，为确保语义清晰，删除了“我觉得”等表述。最后，对各条目的语言进行了精简化处理。经过2轮函询与修订，专家意见已趋于一致。最终形成的理论预设为3个维度，10个条目的调查版量表(表1)。各保留条目的重要性评分均值为4.11~5.00，*CV*为0~0.20，显示出良好的专家共识度。

至此，量表条目池经2轮专家函询后得以确立，具备了进入预调查与正式施测阶段的条件。

表1　调查版量表条目

Table 1　Items of the survey version of the scale

条目	条目内容
P1	AI能有效指导我进行初步就医判断
P2	AI回答的专业性和可信度很高
P3	AI可给予个性化建议(如根据年龄、病情推荐内容)
P4	AI能帮我很好地缓解就医焦虑
I1	AI能准确理解我的提问
I2	AI的回答通俗易懂
I3	使用过程流畅、无卡顿
R1	AI问答回复界面清晰
R2	我愿再次使用该AI智慧问答系统
R3	我会推荐他人使用这个AI智慧问答系统

AI：人工智能。

2.2 预调查结果

20名口腔科初诊患者填写修订版量表后表示条目合理，表述清晰，无理解偏差，无修改内容，能在5 min左右完成。

独立样本*t*检验结果显示，各条目高低分组差异均具有统计学意义(*t*=17.054~22.669，*P*<0.001)，条目区分度良好。各条目与量表总分的Pearson相关系数为0.896~0.940(*P*<0.001)，均大于0.400标准，全部符合保留标准。各条目*CV*为18.57%~21.98%，均低于25%，提示参与者回答存在一定集中趋势，结合其余指标综合判断，不予删除。克龙巴赫α系数法显示，删除任一条目后量表整体α系数均有所下降(Δα=−0.001~−0.003)，全部条目均符合保留标准。经项目分析，10个条目全部予以保留。

2.3 口腔科初诊患者AI智慧问答系统使用体验测评量表的信效度检验结果

2.3.1 调查对象的一般资料

调查对象年龄大多集中在18~30岁；其中，男83名(35.62%)，女150名(64.38%)。在受教育程度方面：初中及以下13名(5.58%)，高中/中专24名(10.30%)，大专34名(14.59%)，本科134名(57.51%)，硕士及以上28名(12.02%)。在职业性质方面：学生45名(19.31%)，在职人员162名(69.53%)，待业及其他12名(5.15%)，退休人员8名(3.43%)，其他6名(2.58%，表2)。

表2　受访者人口统计学特征(*n*=233)

Table 2　Demographic characteristics of respondents (*n*=233)

指标		数值/[例(%)]
性别	男	83(35.62)
	女	150(64.38)
年龄/岁	<18	7(3.00)
	18~30	106(45.49)
	31~45	78(33.48)
	46~60	38(16.31)
	>60	4(1.72)
最高学历	初中及以下	13(5.58)
	高中/中专	24(10.30)
	大专	34(14.59)
	本科	134(57.51)
	硕士及以上	28(12.02)
职业性质	学生(从基础教育到研究生教育的各个阶段)	45(19.31)
	在职	162(69.53)
	待业(包括求职中、家庭主妇/夫等未在职状态)	12(5.15)
	退休人员	8(3.43)
	其他	6(2.58)

此外，用于CFA的第2批调查对象(*n*=72)中，男性22名(30.56%)，女性50名(69.44%)；年龄结构以18~30岁为主(48.61%)，经检验，2批样本在性别、年龄、学历等人口统计学特征上的分布差异均无统计学意义(均*P*>0.05)。

©Journal of Central South University (Medical Science). All rights reserved.

2.3.2 信度检验

2.3.2.1 克龙巴赫α系数

信度检验结果显示，最终形成的10条目量表总体克龙巴赫α系数为0.981，McDonald's ω系数为0.983，二者差值为0.002(<0.020)，提示本质τ等价假设基本成立，α系数作为信度估计较为稳健。专业性与实用性、交互体验、系统体验与行为意向维度的克龙巴赫α系数分别为0.964、0.939、0.959，各条目的校正项总计相关性为0.872~0.925(表3)，均高于0.40的保留标准；删除任一条目后量表克龙巴赫α系数均有所下降(Δα=−0.001~−0.003)，表明各条目均对量表内部一致性有所贡献，全部予以保留，上述数据表明该量表具有较好的内在一致性。

2.3.2.2 分半信度

正式版(10个条目)的分半信度为0.970，量表具有稳定的内在一致性。

2.3.2.3 重测信度

正式版(10个条目)的重测信度结果显示其ICC为0.830(P<0.001)。

2.3.2.4 亚组信度分析

为初步评估量表在不同人群中的稳健性，本研究对EFA样本(n=233)按年龄、学历及性别进行亚组信度分析。结果显示，≤45岁组(n=191)与≥46岁组(n=42)的克龙巴赫α系数分别为0.981和0.981，条目载荷范围分别为0.893~0.938和0.825~0.969；高中及以下学历组(n=71)与大专及以上学历组(n=162)的克龙巴赫α系数分别为0.983和0.979，条目载荷范围分别为0.889~0.970和0.894~0.940；男性组(n=83)与女性组(n=150)的克龙巴赫α系数分别为0.975和0.983，条目载荷范围分别为0.853~0.931和0.894~0.954。各亚组信度指标均维持在较高水平，进一步采用Feldt检验比较各亚组间克龙巴赫α系数差异，结果显示年龄亚组及学历亚组差异均无统计学意义(均P>0.050)，性别亚组间差异具有统计学意义(F=0.663，P=0.041)，但绝对差值仅为0.008(表4)，初步提示量表在不同年龄、学历及性别群体中具有一定的稳健性。

综合上述信度检验结果，总量表克龙巴赫α系数为0.981，McDonald's ω系数为0.983，斯皮尔曼-布朗分半系数为0.970，重测ICC为0.830，表明UES-AIQA具有良好的内部一致性与跨时间稳定性。

表3 量表的信度分析(n=233)

Table 3 Reliability analysis of the scale (n=233)

条目	CITC	项删除后的克龙巴赫α系数	各维度克龙巴赫α系数
I1	0.898	0.979	0.939
I2	0.885	0.979	
I3	0.872	0.979	
P1	0.925	0.978	0.964
P2	0.923	0.978	
P3	0.909	0.978	
P4	0.884	0.979	
R1	0.907	0.978	0.959
R2	0.924	0.978	
R3	0.916	0.978	

CITC：校正项总计相关系数。

表4 UES-AIQA亚组信度分析结果(n=233)

Table 4 Subgroup reliability analysis results (n=233)

分组		n	克龙巴赫α系数	条目载荷范围	组间比较
年龄	≤45岁	191	0.981	0.893~0.938	F=0.993, P=0.936
	≥46岁	42	0.981	0.825~0.969	
学历	高中及以下	71	0.983	0.889~0.970	F=1.206, P=0.337
	大专及以上	162	0.979	0.894~0.940	
性别	男	83	0.975	0.853~0.931	F=0.663, P=0.041
	女	150	0.983	0.894~0.954	

组间α系数差异采用Feldt检验。性别亚组间α系数差异有统计学意义(P=0.041)，但绝对差值仅0.008，实际意义有限；其余亚组间差异均无统计学意义(均P>0.05)。

©Journal of Central South University (Medical Science). All rights reserved.

2.3.3 效度检验

2.3.3.1 内容效度

所有条目*I*-CVI为0.900~1.000，均大于0.780；全体一致量表内容效度指数(scale-level content validity index based on universal agreement，*S*-CVI/UA)为0.800，*S*-CVI/Ave为0.980，均≥0.800。对专家意见评分的随机一致性进行校正后，量表各条目调整后的Kappa值为0.899~1.000，均大于0.740，量表具有良好的内容效度。

2.3.3.2 结构效度

本研究量表初始设有3个理论维度，10个条目，基于233份有效问卷开展EFA。KMO值为0.950(>0.900)，Bartlett检验*P*<0.001，数据适合做因子分析。平行分析结果显示，仅第1个成分的实际特征根(8.521)超过随机模拟数据对应值(1.435)，第2成分实际特征根(0.327)低于随机均值(1.232)，支持提取1个因子。Kaiser准则结果显示仅第1个成分特征根>1.000，与碎石图结论一致(图1)。据此提取1个公因子，并采用ML进行因子提取。10个条目的标准化因子载荷为0.879~0.937，均高于0.400的保留标准，单因子累积方差贡献率为83.593%，量表呈单维度结构(表5)。

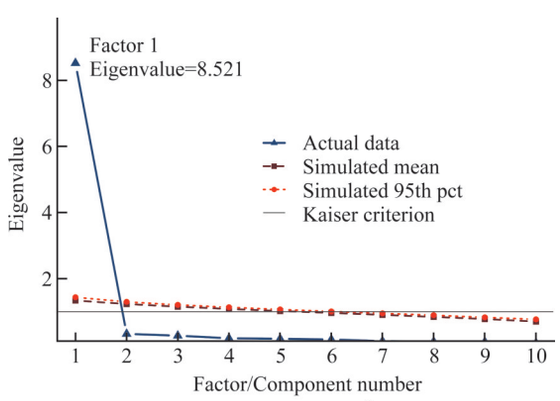


图1　平行分析碎石图(*n*=233，1 000次迭代)

Figure 1　Parallel analysis scree plot (*n*=233, 1 000 iterations)

The solid line (▲) represents the actual data eigenvalue, the red dotted line (■) represents the mean value of random simulation, the orange dotted line (●) represents the 95% percentile of random simulation, and the horizontal solid line represents the Kaiser criterion (eigenvalue=1), with only the actual eigenvalue of the first factor (8.521) exceeding the corresponding value of the 95% percentile of random simulation (1.437), which supports the extraction of one factor.

表5　效度分析结果

Table 5　Validity analysis results

条目	因子载荷	共同度
I1	0.906	0.821
I2	0.893	0.797
I3	0.879	0.773
P1	0.937	0.878
P2	0.932	0.869
P3	0.922	0.850
P4	0.896	0.803
R1	0.915	0.837
R2	0.935	0.874
R3	0.926	0.857

采用最大似然法提取，平行分析与Kaiser准则均支持单因子结构；单因子累积方差贡献率为83.593%。h^2为共同度。

基于第2批独立样本(*n*=72)数据，采用ML对单维度模型进行拟合度检验。结果显示，χ^2/df=3.190，CFI=0.929(>0.900)，Tucker-Lewis指数(Tucker-Lewis index，TLI)=0.900，规范拟合指数(normed fit index，NFI)=0.901(>0.900)，SRMR=0.035(<0.080)，主要拟合指标均达到可接受标准。RMSEA=0.174，高于0.08的常规参考值。各条目标准化因子载荷为0.845~0.928，均大于0.700，且均具有统计学意义(*P*<0.001，表6)，聚合效度良好。

表6　验证性因子分析标准化载荷

Table 6　Standardized factor loadings in confirmatory factor analysis

条目	标准化载荷	*SE*	*z*	*P*
I1	0.928	0		
I2	0.898	0.073	13.156	<0.001
I3	0.922	0.071	14.318	<0.001
P1	0.885	0.082	12.574	<0.001
P2	0.889	0.089	12.754	<0.001
P3	0.845	0.088	11.105	<0.001
P4	0.869	0.083	11.944	<0.001
R1	0.891	0.079	12.818	<0.001
R2	0.893	0.073	12.932	<0.001
R3	0.881	0.079	12.408	<0.001

I1为参照指标(marker variable)，非标准化载荷固定为1，故无*SE*、*z*值及*P*值；其余条目标准化载荷均具有统计学意义(均*P*<0.001)。*SE*：标准差。

©Journal of Central South University (Medical Science). All rights reserved.

3 讨 论

本研究以D&M模型与TAM为双重理论框架[31-32]，研制了专为口腔科初诊患者设计的UES-AIQA。UES-AIQA在条目设计上覆盖了病情解释准确性、就医焦虑缓解效果及个性化建议适切性等口腔专科核心维度，与Wang等[33]开发老年数字健康素养问卷时强调场景特异性条目设计的理念及Zhang等[34]针对专科患者需求进行条目细化的思路相呼应，但其专科适用性优势有待与现有通用工具开展直接对比研究后加以验证。

本研究的专家咨询结果具有较高的权威性与可信度，采用2轮德尔菲专家函询对条目进行系统筛选与修订[35]，并参照量表编制中内容效度指数的报告要求，对条目水平和量表水平内容效度进行评价[36]。专家建议将条目语言贴近患者通俗语境。这一调整符合口腔初诊场景的特殊性——患者常处于候诊焦虑状态，信息处理能力暂时下降[37]，测评工具须以患者为中心。专家组成员对量表各条目的评价较为一致，可终止函询，最终确立量表条目池，同时患者均可在5 min内完成填写。

信度分析结果表明，UES-AIQA各项信度指标均超过心理测量学通用标准，量表具有良好的内部一致性与跨时间稳定性。总量表克龙巴赫α系数偏高(>0.95)在理论上提示需警惕条目冗余的可能性[38]，但删除任一条目后量表克龙巴赫α系数均有所下降，提示各条目均对量表整体信度均有贡献，无条目需删除。既往研究[36]显示，AI聊天机器人在回应患者健康问题时，其回答质量和共情表现均受到关注，提示患者对AI智慧问答系统的评价并非仅限于信息准确性，还可能同时涉及回应充分性、可理解性与情感支持等体验维度。口腔初诊患者在候诊过程中面临信息获取、情绪调节与就诊决策等多重需求，对AI智慧问答系统的评价往往难以在心理层面加以细分，条目间较高的相关性更多反映了这一测量情境的特殊性，而非条目内容的简单重叠。结合本量表仅含10个条目的精简设计，此克龙巴赫α系数在本研究中具有合理性。效度结果方面，内容效度各项指标均达到推荐标准[39]，表明量表条目与口腔初诊患者AI智慧问答系统使用体验的核心概念高度契合。亚组信度分析初步提示量表在不同年龄、学历及性别群体中具有一定稳健性，但老年及低学历亚组样本量相对有限，跨人群效度有待后续研究进一步验证。同时，性别亚组间克龙巴赫α系数差异虽达统计学显著水平，但绝对差值极小(0.008)，且2组克龙巴赫α系数均远超0.900的可接受标准，提示该统计学差异更可能源于女性样本量较大带来的检验效力提升，而非实质性的测量学差异。结构效度方面，本研究最初基于D&M模型与TAM框架预设三维度结构，但EFA结果并不支持该预设。宜作为单维度量表使用。本研究综合数据证据最终采用单维度结构，后续研究可通过双因子模型进一步探索潜在子维度结构，或增设差异性条目以增强区分度。CFA主要拟合指标均达到可接受标准，RMSEA偏高可能与CFA样本量较小对该指标敏感有关[40]，本研究CFA样本量为72例，这可能是RMSEA偏高的重要原因之一。结合多项指标综合评判，模型拟合具有一定合理性，未来可通过更大样本的独立数据集，对该单维度模型的跨样本稳定性与普适性进行进一步检验。

本研究尚存在以下局限：1)样本来源于单一医疗机构，且以中青年、高学历患者为主，外部效度有待多中心、多人群研究验证；尽管亚组信度分析初步提示量表在不同年龄、学历及性别群体中具有一定稳健性，但老年及低学历亚组样本量相对有限，跨人群推广性结论须审慎表述。2)CFA验证样本量偏小(n=72)，结构稳定性需通过扩大样本进一步确认。3)横断面设计仅反映患者即时体验，未能追踪AI系统对后续诊疗效率的客观影响，未来可结合电子病历数据加以深化。4)未能评估量表对不同AI系统版本的跨系统适用性，随着AI技术的快速迭代，量表条目的时效性亦有待持续评估。

综上，本研究通过专家函询确立条目、预调查验证语言适用性、项目分析筛选条目质量，并依次完成信度与效度的系统检验。UES-AIQA呈单维度结构，各条目载荷均良好，量表结构效度达到可接受水平。结果表明，UES-AIQA在内部一致性、跨时间稳定性、内容效度等方面均达到心理测量学通用标准，结构效度部分指标达到可接受水平，可作为评估口腔科初诊患者AI智慧问答系统使用体验的标准化工具，但量表的结构稳定性与跨人群适用性尚需后续扩大样本进一步验证。

作者贡献声明：刘睿夕、匡碧芬 资料收集，数据分析与解释，论文初稿撰写；杨依凡、李宏、赵娅琼 资料收集，数据分析与解释；冯云枝 宏观指导，学术思路评审及论文质量最终审核；袁临天 数据分析与解释，论文撰写与修改；郭玥 选题构思，研究方案设计，关键性理论内容指导与修订。所有作者阅读并同意最终的文本。

利益冲突声明：作者声称无任何利益冲突。

©Journal of Central South University (Medical Science). All rights reserved.

参考文献

[1] 国家卫生健康委办公厅. 健康口腔行动方案(2019—2025年)[EB/OL]. (2019-01-31)[2026-02-05]. https://www.gov.cn/zhengce/zhengceku/2019-12/28/content_5464758.htm.
National Health Commission of China. Healthy oral action plan (2019—2025) [EB/OL]. (2019-01-31) [2026-02-05]. https://www.gov.cn/zhengce/zhengceku/2019-12/28/content_5464758.htm.

[2] Tirpude TN, Warhekar SA, Kokane NB, et al. The role of dental anxiety in shaping patient satisfaction with dental services among patients visiting a tertiary care hospital: a cross-sectional study[J]. Int J Community Med Public Health, 2025, 12(12): 5650-5656. https://doi.org/10.18203/2394-6040.ijcmph20254044.

[3] 何英剑, 欧阳涛, 范照青, 等. 医院临床科室线上初诊咨询系统优势和随访结果分析[J]. 中国医院管理, 2021, 41(11): 73-75.
HE Yingjian, OUYANG Tao, FAN Zhaoqing, et al. A study on the advantages and follow-up results of online consultation system of clinical departments in hospital[J]. Chinese Hospital Management, 2021, 41(11): 73-75.

[4] 王雯雯. 基于患者选择行为的互联网诊疗系统资源优化配置研究[D]. 沈阳: 沈阳建筑大学, 2024.
WANG Wenwen. Research on optimal allocation of Internet diagnosis and treatment system resources based on patients' choice behavior[D]. Shenyang: Shenyang Jianzhu University, 2024.

[5] 李琴, 孙爱娟, 谭明英, 等. 大型综合医院门诊设立初诊患者服务区的探讨[J]. 华西医学, 2016, 31(11): 1907-1909. https://doi.org/10.7507/1002-0179.201600524.
LI Qin, SUN Aijuan, TAN Mingying, et al. Discussion on setting up service area for first-time patients in outpatient department of large general hospital[J]. West China Medical Journal, 2016, 31(11): 1907-1909. https://doi.org/10.7507/1002-0179.201600524.

[6] 魏来. 整合型医疗服务系统构建研究[J]. 中国卫生政策研究, 2018, 11(10): 1-11.https://doi.org/10.3969/j.issn.1674-2982.2018.10.001.
WEI Lai. Research on the construction of integrated medical service system[J]. Chinese Journal of Health Policy, 2018, 11(10): 1-11. https://doi.org/10.3969/j.issn.1674-2982.2018.10.001.

[7] 国家卫生健康委员会办公厅, 国家中医药管理局, 国家疾病预防控制局. 卫生健康行业人工智能应用场景参考指引 [EB/OL]. (2024-11-06) [2026-02-05]. https://www.nhc.gov.cn.
National Health Commission of China, National Administration of Traditional Chinese Medicine, National Disease Control and Prevention Administration. Reference Guide for AI Application Scenarios in the Health Sector [EB/OL]. (2024-11-06)[2026-02-05]. https://www.nhc.gov.cn.

[8] 国家卫生健康委员会办公厅. 关于促进和规范“人工智能+医疗卫生”应用发展的实施意见[EB/OL]. (2025-11-09)[2026-02-05]. https://www.nhc.gov.cn.
National Health Commission of China. Implementation Opinions on Promoting and Regulating the Development of "AI+Healthcare" Applications[EB/OL]. (2025-11-09)[2026-02-05]. https://www.nhc.gov.cn.

[9] Goktas P, Karakaya G, Kalyoncu AF, et al. Artificial intelligence chatbots in allergy and immunology practice: where have we been and where are we going? [J]. J Allergy Clin Immunol Pract, 2023, 11(9): 2697-2700. https://doi.org/10.1016/j.jaip.2023.05.042.

[10] Yang HS, Wang F, Greenblatt MB, et al. AI chatbots in clinical laboratory medicine: foundations and trends[J]. Clin Chem, 2023, 69(11): 1238-1246. https://doi.org/10.1093/clinchem/hvad106.

[11] Li J. Security implications of AI chatbots in health care[J/OL]. J Med Internet Res, 2023, 25: e47551[2026-01-23]. https://doi.org/10.2196/47551.

[12] Yu H, Zhou J, Li L, et al. Simulated patient systems powered by large language model-based AI agents offer potential for transforming medical education[J]. Commun Med, 2026, 6: 27. https://doi.org/10.1038/s43856-025-01283-x.

[13] Rokhshad R, Bagherianlemraski M, Ehsani SS, et al. Large language models for the screening step in systematic reviews in dentistry[J]. J Dent, 2025, 160: 105877. https://doi.org/10.1016/j.jdent.2025.105877.

[14] Mohammad-Rahimi H, Setzer FC, Aminoshariae A, et al. Artificial intelligence chatbots in endodontic education: Concepts and potential applications[J]. Int Endodontic J, 2026, 59(6): 999-1012. https://doi.org/10.1111/iej.14231.

[15] Huang ES. The user experience of AI[J]. Med, 2022, 3(4): 228-232. https://doi.org/10.1016/j.medj.2022.03.005.

[16] Browne S, Umlauf A, Moore DJ, et al. User experience of persons using ingestible sensor-enabled pre-exposure prophylaxis to prevent HIV infection: cross-sectional survey study[J/OL]. JMIR Mhealth Uhealth, 2024, 12: e53596[2026-01-23]. https://doi.org/10.2196/53596.

[17] Simola S, Hörhammer I, Xu YH, et al. Patients' experiences of a national patient portal and its usability: cross-sectional survey study[J/OL]. J Med Internet Res, 2023, 25: e45974[2026-01-23]. https://doi.org/10.2196/45974.

[18] Newton AS, March S, Gehring ND, et al. Establishing a working definition of user experience for eHealth interventions of self-reported user experience measures with eHealth researchers and adolescents: scoping review[J/OL]. J Med Internet Res, 2021, 23(12): e25012[2026-01-23]. https://doi.org/10.2196/25012.

[19] Madfa AA, Alshammari AF, Alanazi BA, et al. Evaluating the effectiveness of an artificial intelligence-based chatbot for reducing dental anxiety among adult dental patients[J]. Front Oral Health, 2026, 7: 1778200. https://doi.org/10.3389/froh.2026.1778200.

[20] Al-Zubaidy D, Innes N, Galloway J, et al. Evaluating user perceptions and usability of an AI-powered smartphone application for at-home dental plaque screening[J]. Br Dent J, 2025, 239(1): 46-52. https://doi.org/10.1038/s41415-025-8502-0.

[21] 陶毛毛, 刘隽, 廖骞, 等. 基于AI智能语音随访的口腔科互联网门诊患者使用现状分析[J]. 上海口腔医学, 2023, 32(2): 183-187. https://doi.org/10.19439/j.sjos.2023.02.013.
TAO Maomao, LIU Jun, LIAO Qian, et al. Analysis of the use

©Journal of Central South University (Medical Science). All rights reserved.

status of Internet outpatient in stomatological departments based on AI follow-up[J]. Shanghai Journal of Stomatology, 2023, 32(2): 183-187. https://doi.org/10.19439/j.sjos.2023.02.013.

[22] 张洪滔, 孙婉昕, 李晓英, 等. 中文人工智能大语言模型在口腔黏膜病疾病咨询中的应用评价[J]. 中华口腔医学杂志, 2026, 61(3): 332-338. https://doi.org/10.3760/cma.j.cn112144-20251201-00482.
ZHANG Hongtao, SUN Wanxin, LI Xiaoying, et al. Evaluation of Chinese artificial intelligence large language models in oral mucosal disease consultation[J]. Chinese Journal of Stomatology, 2026, 61(3): 332-338. https://doi.org/10.3760/cma.j.cn112144-20251201-00482.

[23] 武思齐, 庞潇, 李丹妮, 等. 人工智能在口腔互联网医院健康管理中的应用研究[J]. 中国卫生质量管理, 2025, 32(12): 8-12.
WU Siqi, PANG Xiao, LI Danni, et al. Research on the application of artificial intelligence in health management of oral Internet hospitals[J]. Chinese Health Quality Management, 2025, 32(12): 8-12.

[24] Sharpe L, Menzies RE, Richmond B, et al. The development and validation of the worries about recurrence or progression scale (WARPS)[J]. Br J Health Psychol, 2024, 29(2): 454-467. https://doi.org/10.1111/bjhp.12707.

[25] Nunnally JC, Bernstein IH. Psychometric Theory[M]. 3rd ed. New York: McGraw-Hill, 1994.

[26] Wolf EJ, Harrington KM, Clark SL, et al. Sample size requirements for structural equation models: an evaluation of power, bias, and solution propriety[J]. Educ Psychol Meas, 2013, 73(6): 913-934. https://doi.org/10.1177/0013164413495237.

[27] 陆晶, 杨璇, 石双姣, 等. 压力性损伤患者非正式照顾者知信行量表的编制及信效度检验[J]. 中南大学学报(医学版), 2024, 49(6): 972-980. https://doi.org/10.11817/j.issn.1672-7347.2024.240039
LU Jing, YANG Xuan, SHI Shuangjiao, et al. Development and its reliability and validity of the Knowledge, Attitude and Practice Scale for Informal Caregivers of Pressure Injury Patients[J]. Journal of Central South University. Medical Science, 2024, 49(6): 972-980. https://doi.org/10.11817/j.issn.1672-7347.2024.240039

[28] Hur Y, Seo DG. Development of a character qualities test for medical students in Korea using polytomous item response theory and factor analysis: a preliminary scale development study[J]. J Educ Eval Health Prof, 2023, 20: 20. https://doi.org/10.3352/jeehp.2023.20.20.

[29] Kim H, Schnall R, Yoon N, et al. Development and validation of a mobile-centered digital health readiness scale (mDiHERS): health literacy and equity scale[J/OL]. J Med Internet Res, 2024, 26: e58497[2026-01-23]. https://doi.org/10.2196/58497.

[30] Savalei V, Huang M. Fit indices are insensitive to multiple minor violations of perfect simple structure in confirmatory factor analysis[J]. Psychol Meth, 2025. https://doi.org/10.1037/met0000718.

[31] Zheng F, Wang K, Wang Q, et al. Factors influencing clinicians' use of hospital information systems for infection prevention and control: cross-sectional study based on the extended DeLone and McLean model[J/OL]. J Med Internet Res, 2023, 25: e44900[2026-01-23]. https://doi.org/10.2196/44900.

[32] Song T, Deng N, Cui T, et al. Measuring success of patients' continuous use of mobile health services for self-management of chronic conditions: model development and validation [J/OL]. J Med Internet Res, 2021, 23(7): e26670[2026-01-23]. https://doi.org/10. 2196/26670.

[33] Wang X, Zhang C, Qi Y, et al. Digital health literacy questionnaire for older adults: instrument development and validation study[J/OL]. J Med Internet Res, 2025, 27: e64193 [2026-01-23]. https://doi.org/10.2196/64193.

[34] Zhang J, Kong F, Wang W, et al. Reliability and validity of the quality of life with implant prostheses in China: a tool for assessing quality of life among implant-prostheses wearers[J]. J Dent, 2025, 162: 106106. https://doi.org/10.1016/j.jdent.2025.106106.

[35] 彭慕云, 喻风雷. 基于德尔菲法构建胸外科手术难度评价量表[J]. 中南大学学报(医学版), 2022, 47(5): 655-664. https://doi.org/10.11817/j.issn.1672-7347.2022.200782.
PENG Muyun, YU Fenglei. Establishment of thoracic surgical difficulty assessment scale based on Delphi method[J]. Journal of Central South University. Medical Science, 2022, 47(5): 655-664. https://doi.org/10.11817/j.issn.1672-7347.2022.200782.

[36] Almanasreh E, Moles R, Chen TF. Evaluation of methods used for estimating content validity[J]. Res Social Adm Pharm, 2019, 15(2): 214-221. https://doi.org/10.1016/j.sapharm.2018.03.066.

[37] 胡格, 马宇锋. 牙科焦虑症治疗方法的研究进展[J]. 临床口腔医学杂志, 2023, 39(7): 436-439. https://doi.org/10.3969/j.issn.1003-1634.2023.07.014.
HU Ge, MA Yufeng. Research progress on treatment methods of dental anxiety disorder[J]. Journal of Clinical Stomatology, 2023, 39(7): 436-439. https://doi.org/10.3969/j.issn.1003-1634.2023.07.014.

[38] Streiner DL. Starting at the beginning: an introduction to coefficient alpha and internal consistency[J]. J Pers Assess, 2003, 80(1): 99-103. https://doi.org/10.1207/s15327752jpa8001_18.

[39] 史静琤, 莫显昆, 孙振球. 量表编制中内容效度指数的应用[J]. 中南大学学报(医学版), 2012, 37(2): 152-155. https://doi.org/10.3969/j.issn.1672-7347.2012.02.007.
SHI Jingcheng, MO Xiankun, SUN Zhenqiu. Content validity index in scale development[J]. Journal of Central South University (Medical Science), 2012, 37(2): 152-155. https://doi.org/10.3969/j.issn.1672-7347.2012.02.007.

[40] Kenny DA, Kaniskan B, McCoach DB. The performance of RMSEA in models with small degrees of freedom[J]. Sociol Meth Res, 2015, 44(3): 486-507. https://doi.org/10.1177/0049124114543236.

（责任编辑 宋柳）

©Journal of Central South University (Medical Science). All rights reserved.