·数智医学教育·

# 医学生医疗人工智能准备度量表汉化及信效度检验

何登鑫　陈芳　赵梅村
湖北中医药大学护理学院实验教学中心，武汉 430065
通信作者：何登鑫，Email：3261@hbucm.edu.cn

【摘要】 **目的**　本研究旨在汉化医学生医疗人工智能准备度量表，并检验其在医学生群体中的信效度，以此为评估中国医学生的医疗人工智能准备度提供可靠工具。**方法**　依据Brislin翻译模型进行量表汉化，过程包括翻译、回译和文化调适。采用便利抽样法选取267名医学生开展问卷调查，运用R 4.3.1软件进行统计分析以检验信效度。**结果**　中文版医学生医疗人工智能准备度量表包含基础知识、实践能力、态度和价值观3个维度，共14个条目，累计方差贡献率达54.03%。模型适配情况良好（$\chi^2/df$=1.63，RMSEA=0.05，TLI=0.94，GFI=0.93，CFI=0.95）。量表各维度Cronbach's $\alpha$系数在0.71~0.78之间，量表整体Cronbach's $\alpha$系数为0.85，重测信度为0.67。**结论**　汉化后的医学生医疗人工智能准备度量表信效度达标，可用于评估中国医学生的医疗人工智能准备度，能为医疗人工智能教育改革提供实证依据。

【关键词】 医学生；　医疗人工智能准备度量表；　量表汉化；　信效度检验
【中图分类号】 R-05
**基金项目**：2024年度湖北省教育厅哲学社会研究项目青年项目（24Q179）
DOI: 10.3760/cma.j.cn116021-20250923-02197

**Chinese version of Medical Artificial Intelligence Readiness Scale for Medical Students and its reliability and validity tests**

*He Dengxin, Chen Fang, Zhao Meicun*
*Experiment Teaching Center, School of Nursing, Hubei University of Chinese Medicine, Wuhan 430065, China*
*Corresponding author: He Dengxin, Email: 3261@hbucm.edu.cn*

【**Abstract**】 **Objective**　To translate the Medical Artificial Intelligence Readiness Scale for Medical Students into Chinese, test its reliability and validity among medical students, and provide a reliable tool for assessing medical artificial intelligence readiness of medical students in China. **Methods**　The scale was translated following Brislin's translation model, which included translation, back-translation, and cultural adaptation. A total of 267 medical students were selected using convenience sampling to complete the questionnaire survey. R 4.3.1 software was used for statistical analysis to evaluate reliability and validity. **Results**　The Chinese version of the Medical Artificial Intelligence Readiness Scale for Medical Students consisted of 14 items across three dimensions: basic knowledge, practical ability, and attitude and values, with a cumulative variance contribution rate of 54.03%. The confirmatory factor analysis indicated a good model fit ($\chi^2/df$=1.63, root mean square error of approximation=0.05, Tucker-Lewis index=0.94, goodness-of-fit index=0.93, comparative fit index=0.95). The Cronbach's $\alpha$ coefficients for the three dimensions ranged from 0.71 to 0.78, and the overall Cronbach's $\alpha$ coefficient of the scale was 0.85. The test-retest reliability was 0.67. **Conclusions**　The Chinese version of the Medical Artificial Intelligence Readiness Scale for Medical Students demonstrated acceptable reliability and validity, and can be used to assess medical artificial intelligence readiness among Chinese medical students. These results provide an empirical basis for reforms in medical artificial intelligence education.

【**Key words**】 Medical student;　Medical Artificial Intelligence Readiness Scale;　Translation of scale into Chinese;　Reliability and validity tests
**Fund program:** Philosophy and Social Research Project of Hubei Provincial Department of Education

in 2024, Youth Project (24Q179)
DOI: 10.3760/cma.j.cn116021-20250923-02197

随着人工智能技术飞速发展，其在医疗领域的应用已成为热点和趋势[1-3]。医疗人工智能能够实现海量医学数据的自动化处理、智能辅助诊断、助力健康管理与远程医疗、加速药物研发及临床试验[4]，不仅能提升医生的诊疗能力，还能显著降低医疗成本、提高医疗效率[5]。然而，人工智能在医疗领域的发展和应用，也对医学教育提出了新要求。医学教育者不仅要向医学生传授医疗知识，还要帮助他们掌握人工智能这一最新科技成果。医疗人工智能教育是培养医学生运用人工智能相关知识分析和解决问题能力、提升其人工智能准备度的重要途径[6]，已成为医学教育中不可或缺的一部分。目前，中国医疗人工智能教育尚处于发展阶段，存在培养目标不明确、教育内容不完善等问题[7]，这些问题严重阻碍了医学与人工智能复合型人才的培养，原因在于缺乏经科学验证的本土化测评工具来系统评估中国医学生的人工智能准备度。因此，迫切需要借助经科学验证的本土化测评工具完善医疗人工智能教育体系。

国际学术界已开发出医疗人工智能准备度量表，用于评估医学生的医疗人工智能准备情况，其中，Karaca 等[8]开发的医学生医疗人工智能准备度量表（medical artificial intelligence readiness scale for medical students，MAIRS-MS）较为成熟。Karaca 等[8]依据医学生应具备的医疗人工智能能力与特点编制了 MAIRS-MS，该量表总体信效度良好，能为医学教育提供有价值的参考，帮助了解医学生医疗人工智能准备度的优势和不足，为改进教学方法、优化课程设置提供依据。不过，由于国内外医学教育体系存在差异，且医疗人工智能技术应用场景具有区域特性[7]，直接移植原量表可能会引发文化适应性问题，因此需要对 MAIRS-MS 进行系统性汉化与文化调适。

本研究结合专家咨询，遵循国际通行的 Brislin 双人互译模型[9]，对 MAIRS-MS 开展系统性汉化与文化调适工作。运用汉化后的量表收集数据，并检验其信效度，旨在为中国医学生医疗人工智能准备度评估提供科学工具。

## 1　资料与方法

### 1.1　原量表介绍

2021 年，Karaca 等[8]根据医学生所需具备的医疗人工智能相关准备能力及其特征，编制了 MAIRS-MS 量表，用于评估医学生的医疗人工智能准备度。该量表包含 4 个维度，共 22 个条目。其中，认知维度和能力维度各含 8 个条目，视野维度和伦理维度各含 3 个条目。所有条目均采用正向计分，遵循 Likert 5 级量表评分标准："强烈不同意"计 1 分，"不同意"计 2 分，"中立"计 3 分，"同意"计 4 分，"强烈同意"计 5 分。量表总分范围为 22~110 分，得分越高，表明医学生的医疗人工智能准备度越高。研究结果显示，该量表整体的 Cronbach's $\alpha$ 系数为 0.87，具有良好的信效度[8]。

### 1.2　量表汉化

#### 1.2.1　翻译

本研究在获得 MAIRS-MS 原作者授权后，依据 Brislin 模型开展量表汉化工作[9]。具体流程如下：首先，由 2 名精通英文、母语为中文的医学硕士研究生，独立将量表翻译成中文版量表 A1 和 A2。研究小组在保留原量表核心含义的前提下，整合 2 个翻译版本，形成中文版量表 A3。随后，邀请 2 名未接触过原量表且有量表翻译经验的英语专业专家，将中文版量表 A3 回译，得到回译版量表 B1 和 B2。研究小组将 2 份回译量表与原量表逐条对比，集体讨论、修改，必要时与原作者沟通，确保中文版量表与原量表语义、概念一致，最终形成中文版量表 A4。

#### 1.2.2　跨文化调试

邀请 6 名专家组成专家组，开展 Delphi 咨询，完成量表的跨文化调试。专家纳入标准如下：①具有学士及以上学历；②具有副高及以上职称；③熟练掌握评估工具开发及心理测量特征测量方法；④具备扎实的医学教育知识及丰富的临床实践经验。专家对汉化量表的内容相关性、语言表达清晰度、文化背景适用性等方面进行评价，并提出删除或修改建议。综合专家评价结果，对量表进行调整、修改，形成跨文化调试版量表 A5。

#### 1.2.3　预调查

2025 年 2 月至 3 月，研究团队采用便利抽样法，选取医学生，并运用跨文化调试版量表 A5 开展预调查。预调查中，研究对象的纳入标准为：①年龄≥18 岁；②为医学专业本科生或研究生；③调查期间处于在校就读或实习阶段。排除标准为：①存在规律作答行为，例如连续 10 个项目周期性作答或选择同一选项；②作答时间少于 2 min 或超过 20 min。

预调查共纳入20名医学生。调查启动前，研究人员向研究对象详细说明了本研究的目的和内容，在获得其明确同意后开展预调查工作。医学生完成问卷后，研究人员请他们指出不明白或不清楚的条目和词项。研究小组根据反馈信息，对量表条目进行修正，最终形成中文版MAIRS-MS A6。

### 1.3 量表的信效度检验

#### 1.3.1 研究对象

2025年3月至10月，研究团队采用便利抽样法，选取符合预调查所设定纳入与排除标准的医学生开展调查。根据探索性因子分析要求，样本量应为量表条目数的5~10倍[10]。原量表共有22个条目，据此估计需收集220份问卷。

#### 1.3.2 研究工具

一般人口学资料调查表：采用自行设计的问卷，涵盖性别、年龄、培养层次、所在年级、民族和生源地等信息。人工智能素养量表：以人工智能素养量表作为效标[11]，检验中文版MAIRS-MS的效标关联效度。该量表包含知识、技能、态度/价值观、伦理4个维度，共计25个条目，各维度Cronbach's α系数为0.755~0.911。

#### 1.3.3 资料收集与质量控制

本次研究采用多样化的海报宣传方式招募参与者。具体流程如下：对研究内容感兴趣的潜在参与者添加调查人员联系方式后，调查人员会向其介绍项目详情并初步核验信息；在核验环节，调查人员需确认研究对象学生证信息的真实性，且所有参与者均须在自愿的前提下签署知情同意书，完成上述步骤后才能填写问卷。

为保障数据质量，问卷设置了2项用于质量筛查的问题："您是否认为自己是世界前十的人工智能研究人员？"以及"您是否认为自己是中国最好的人工智能研究人员？"只有这2项回答均为"否"的问卷，才可纳入数据分析。

本次研究共回收366份问卷，其中99份因不符合质量控制标准被排除，最终有267份有效问卷被纳入分析，以检验量表信效度。

### 1.4 统计学方法

使用R 4.3.1软件进行统计分析，检验水准$\alpha$=0.05。定性数据以频数及百分比表示，定量数据以均数±标准差呈现。量表信效度检验内容包括内容效度、结构效度、收敛效度、区分效度、效标关联效度、内部一致性信度、折半信度、重测信度。

内容效度：内容效度包括量表条目水平的内容效度指数（item-level CVI，I-CVI）和量表水平的内容效度指数（scale-level CVI，S-CVI），S-CVI含全体一致S-CVI和平均S-CVI。若3项指标分别大于或等于0.78、0.80、0.90提示量表内容效度良好[12]。

结构效度：首先根据Kaiser-Meyer-Olkin（KMO）值和Bartlett's球形检验结果评估是否适合进行因子分析，若KMO值接近1且Bartlett's球形检验结果具有统计学意义（$P$<0.05），提示可以进行因子分析。在探索性因子分析过程中，应用方差最大化正交旋转法提取特征值大于或等于1的公因子、删除因子主载荷小于0.50的条目[13]，最终模型累计方差贡献率大于或等于50%表明量表结构效度良好[14]。在验证性因子分析过程中，采用模型参数卡方/自由度比（$\chi^2/df$）、近似误差均方根（root mean square error，RMSEA）、Tucker-Lewis指数（Tucker-Lewis index，TLI）、拟合优度指数（goodness of fit index，GFI）、比较拟合指数（comparative fit index，CFI）评估模型优劣，若$1<\chi^2/df<3$、RMSEA<0.08、TLI>0.90、GFI>0.90、CFI>0.90表明模型拟合良好[15]，量表结构效度较好。

收敛效度：基于验证性因子分析结果计算量表各维度的组合信度、平均方差抽取值（average variance extracted，AVE），组合信度大于0.70、平均方差抽取值大于0.50表示量表收敛效度良好[16]。

区分效度：计算量表各维度间的相关系数与各维度平均方差抽取值的平方根，相关系数小于对应维度的平均方差抽取值的平方根表明量表区分效度达到标准[16]。

效标关联效度：选择人工智能素养量表作为效标[11]，检验中文版MAIRS-MS的效标关联效度，使用Pearson相关系数法计算二者总分的相关系数$r$，$r$绝对值越接近1，其效标关联效度越佳。

内部一致性信度：采用Cronbach's α系数评价量表的内部一致性信度，量表各维度Cronbach's α系数大于0.60、量表总体Cronbach's α系数大于0.80表明该量表的内部一致性较高，信度良好。

折半信度：采用Spearman-Brown公式进行折半信度分析，计算相关系数$r$，$r$大于0.60表明量表折半信度能够接受。

重测信度：本研究选取30名医学生，2周后检验重测信度。用相关系数$r$表示重测信度。$r$越接近1，量表稳定性越佳。

## 2 结果

### 2.1 跨文化调试结果

根据专家意见，研究小组研讨后对量表的6个条目进行修订，具体情况见表1。

**表 1**　跨文化调试修改内容

| 条目 | 修改后条目 |
|---|---|
| 5. 我能够正确分析人工智能在医疗保健中获得的数据 | 5. 我能够准确分析人工智能在医疗服务中获得的数据 |
| 7. 我能够根据人工智能的逻辑组织工作流程 | 7. 我能够基于人工智能的逻辑和原则组织工作流程 |
| 11. 我能够根据人工智能应用的目的正确使用它们 | 11. 我能够依据人工智能应用的预期目的合理使用它们 |
| 13. 我能够解释医疗人工智能应用如何解决哪个问题 | 13. 我能够解释医疗人工智能应用旨在解决哪些问题 |
| 16. 我能够根据医疗保健中遇到的问题选择合适的人工智能应用 | 16. 我能够依据医疗场景下遭遇的具体问题选择合适的人工智能应用 |
| 20. 我能够根据法律和伦理规范使用健康数据 | 20. 我能够遵循法律法规及伦理准则使用健康数据 |

**2.2**　预调查结果

使用经过文化调适的量表对 20 名医学生进行预调查，医学生反馈能够理解问卷，未发现条目表达不清晰或填写困难的问题，因此未对量表的条目进行修订。

**2.3**　信效度检验结果

**2.3.1**　医学生的一般资料

本次调查研究共纳入 267 名医学生，年龄为（22.14 ± 2.25）岁，涉及临床医学、护理学、口腔医学、药学等多个医学专业。83.52%（223/267）的研究对象的培养层次为本科。具体情况见表 2。

**2.3.2**　效度检验结果

内容效度：中文版 MAIRS-MS 的 I-CVI 值处于 0.83~1.00 之间，全体一致的 S-CVI 为 0.92，平均 S-CVI 为 0.86，这表明中文版 MAIRS-MS 的内容效度较好。

结构效度：KMO 值为 0.89 且 Bartlett's 球形检验结果有统计学差异（$\chi^2$=925.73，$P$<0.001），提示可以进行因子分析。通过主成分分析法，研究小组提取了 3 个公因子，结合其内容特性及专家意见，分别将其命名为基础知识、实践能力、态度和价值观。在探索性因子分析过程中，以下条目因主载荷低于 0.50 被删除：条目 3“我能够解释人工智能系统是如何被训练的”、条目 5“我能够准确分析人工智能在医疗中获得的数据”、条目 7“我能够基于人工智能的逻辑和原则组织工作流程”、条目 8“我能够阐述数据收集、分析、评估和安全性对医疗人工智能发展的重要性”、条目 9“我能够结合我的专业知识使用基于人工智能的信息”、条目 11“我能够依据人工智能应用的预期目的合理使用它们”、条目 16“我能够依据医疗场景下遭遇的具体问题选择合适的人工智能应用”。删除主载荷低于 0.50 的条目后，剩余 14 个条目的中文版 MAIRS-MS 累计方差贡献率为 54.03%。各条目的主载荷为 0.56~0.86（表 3）。验证性因子分析结果显示模型参数 $\chi^2/df$=1.63，RMSEA=0.05，TLI=0.94，GFI=0.93，CFI=0.95。模型的各参数均在参考范围内[12]，提示量表结构效度良好，模型拟合良好。

**表 2**　医学生的一般资料[$n$=267，例（%）]

| 项目 | 分类 | 人数 |
|---|---|---|
| 性别 | 男 | 108（40.45） |
| | 女 | 159（59.55） |
| 专业 | 临床医学 | 68（25.47） |
| | 护理学 | 39（14.61） |
| | 口腔医学 | 37（13.86） |
| | 药学 | 28（10.49） |
| | 医学技术学 | 20（7.49） |
| | 康复医学 | 16（5.99） |
| | 法医学 | 15（5.62） |
| | 基础医学 | 14（5.24） |
| | 中医学类 | 14（5.24） |
| | 公共卫生与预防医学 | 10（3.74） |
| | 其他 | 6（2.25） |
| 培养层次 | 本科 | 223（83.52） |
| | 研究生 | 44（16.48） |
| 生源地 | 城镇 | 139（52.06） |
| | 农村 | 128（47.94） |
| 民族 | 汉族 | 264（98.88） |
| | 其他 | 3（1.12） |
| 是否独生子女 | 是 | 124（46.44） |
| | 否 | 143（53.56） |
| 父母受教育程度 | 均在小学及以下 | 20（7.49） |
| | 至少 1 人初/高中（包括中专） | 140（52.43） |
| | 至少 1 人大专/本科 | 82（30.71） |
| | 均为本科及以上 | 25（9.36） |
| 家庭平均月收入 | <5 000 元 | 27（10.12） |
| | 5 000~元 | 120（44.94） |
| | 10 000~元 | 120（44.94） |

注：由于四舍五入，统计结果可能不等于 100.00%。

收敛效度：基础知识、实践能力、态度和价值观 3 个维度的组合信度依次为 0.78、0.78 和 0.74，平均方差抽取值依次为 0.55、0.37 和 0.37，平均方差抽取值大于 0.36[17]，这表明中文版 MAIRS-MS 收敛效度可以接受。

**表3**　中文版MAIRS-MS探索性因子分析结果

| 条目 | 基础知识 | 实践能力 | 态度和价值观 |
|---|---|---|---|
| 1.我能够定义数据科学的基本概念 | 0.86 | 0.17 | 0.08 |
| 2.我能够定义统计学的基本概念 | 0.80 | 0.08 | 0.18 |
| 4.我能够定义人工智能的基本概念和术语 | 0.72 | 0.27 | 0.13 |
| 6.我能够区分人工智能相关工具和应用的功能与特点 | 0.36 | 0.60 | 0.25 |
| 12.我能够使用信息和通信技术获取、评估、使用、分享和创造新知识 | 0.20 | 0.56 | 0.33 |
| 13.我能够解释医疗人工智能应用旨在解决哪些问题 | 0.20 | 0.70 | –0.02 |
| 15.我能够向患者解释在医疗服务中使用的人工智能应用 | 0.15 | 0.60 | 0.18 |
| 17.我能够解释人工智能技术的局限性 | 0.06 | 0.61 | 0.22 |
| 18.我能够解释人工智能技术的优势和劣势 | 0.02 | 0.75 | 0.13 |
| 10.我能够在医疗服务中有效且高效地使用人工智能技术 | 0.20 | 0.33 | 0.57 |
| 14.我认为将人工智能用于教育、服务和研究是有价值的 | 0.17 | 0 | 0.71 |
| 20.我能够遵循法律法规及伦理准则使用健康数据 | 0.12 | 0.16 | 0.70 |
| 21.我能够在使用人工智能技术时遵循伦理原则 | 0.05 | 0.37 | 0.59 |
| 22.我能够遵守在医疗领域使用人工智能技术的相关法律法规 | 0.02 | 0.14 | 0.73 |
| 特征值 | 2.23 | 2.86 | 2.51 |
| 累计方差贡献率(%) | 16.74 | 36.11 | 54.03 |

区分效度：基础知识、实践能力、态度和价值观3个维度平均方差抽取值的平方根依次为0.74、0.61、0.61，各维度间的相关系数均小于对应维度的平均方差抽取值的平方根（表4），提示量表区分效度达标。

**表4**　中文版MAIRS-MS各维度相关系数与平均方差抽取值的平方根

| 维度 | 基础知识 | 实践能力 | 态度和价值观 |
|---|---|---|---|
| 基础知识 | 0.74[a] | | |
| 实践能力 | 0.57 | 0.61[a] | |
| 态度和价值观 | 0.48 | 0.57 | 0.61[a] |

注：a，基础知识、实践能力、态度和价值观3个维度平均方差抽取值的平方根。

效标关联效度：本研究选择人工智能素养量表作为中文版MAIRS-MS的效标。Pearson相关系数法分析结果显示，二者总分呈正相关关系（$r$=0.44，$P$<0.001），相关系数大于0.40[18]，提示中文版MAIRS-MS效标关联效度达标。

**2.3.3**　信度检验结果

内部一致性信度：量表的基础知识、实践能力、态度和价值观3个维度的Cronbach's $\alpha$系数分别为0.78、0.76、0.71，量表总体Cronbach's $\alpha$系数为0.85，说明量表整体及各维度的内部一致性符合标准。

折半信度：量表的折半信度为0.89，表明量表折半信度达标。

重测信度：间隔2周后，选取30名医学生进行再次测量，结果显示各维度的重测相关系数介于0.61~0.93之间，总量表的重测相关系数为0.67，重测相关系数大于0.60[19]，提示量表重测信度可以接受。

**2.4**　医学生医疗人工智能准备度现状

研究结果显示，医学生医疗人工智能准备度所有条目均分为（4.12 ± 0.43），提示医学生医疗人工智能准备度整体处于较高水平。基础知识、实践能力、态度和价值观3个维度的条目均分分别为（3.77 ± 0.78）、（4.09 ± 0.50）、（4.36 ± 0.44），提示医学生医疗人工智能相关的基础知识处于中等水平，实践能力与态度和价值观水平处于较高水平。

## 3　讨论

**3.1**　中文版MAIRS-MS汉化过程科学严谨

中文版MAIRS-MS汉化过程严格遵循Brislin模型[9]。在翻译阶段，由研究小组成员共同研讨并达成共识后确定最终翻译版本；在跨文化调试阶段，针对与国内医学教育情况不符的条目，咨询专家后修改条目，确保量表条目准确且符合国内实际情况。专家团队具有一定权威性，6名咨询专家均具备副高及以上职称，具备扎实的医学教育知识与丰富的临床实践经验。在正式调查前，本研究通过预调查收集20名医学生的真实感受与反馈，确保量表内容适合中国医学生实际情况。最后，通过正式调查收集数据，对量表进行信效度检验、并根据检验结果调整量表的维度及条目。通过以上翻译、跨文化调试、信效度检验及量表调整过程，最终形成中文版MAIRS-MS。

需注意的是，原始量表包含4个维度、22个条目，而汉化后的中文版MAIRS-MS最终删除8个条目、保留14个条目，并重新整合为3个维度。维度划分的差异可能源于医学生所处的文化背景与医疗体系环境的不同。从维度对应关系来看，中文版新增的“基础知识”“实践能力”2个维度，与原始量表的“认知”“能力”维度在核心概念与内容指向性上差异较小，可视为对原始构念的本土化延续。二者的关键差异集中在维度整合层面：原始量表中独立的“视野”与“伦理”维度，在中文版中被归并融合为“态度和价值观”单一维度。这一整合的文化逻辑具体表现为：在国外文化语境下，“视野”维度更强调个体对医疗人工智能发展趋势的独立判断能力，“伦理”维度则侧重对伦理观点的包容与自主权衡；而中国医疗体系具有鲜明的“权威导向”特征（如政策法规的刚性约束、行业专家共识的指导作用），因此中文版“态度和价值观”维度的测量重点发生了本土化调整，更聚焦于两点核心：一是对国家层面医疗人工智能相关政策的认同与遵从度，二是对医疗行业权威伦理规范的理解与执行度，而非单纯考查个体的主观判断倾向。维度缩减并非简单的条目删减，而是将原始量表中的4个构念，基于中国医疗教育实践逻辑，整合为3个更具内在关联性与逻辑递进性的核心维度。调整后的“基础知识—实践能力—态度和价值观”三维结构，形成了完整的逻辑闭环：其中，“基础知识”维度解决“知道什么”的问题，是医学生应用医疗人工智能的前提；“实践能力”维度解决“会用什么”的问题，是连接理论知识与临床实践的核心纽带；“态度和价值观”维度解决“愿意怎么用”的问题，是确保医疗人工智能应用符合伦理要求与法律法规的底线保障，整体更契合中国医学生医疗人工智能准备度的培养目标与评价逻辑。

**3.2**　中文版MAIRS-MS效度合格

在本研究中，量表效度检验包括内容效度、结构效度、收敛效度、区分效度、效标关联效度5个方面。量表的I-CVI为0.83~1.00、全体一致S-CVI为0.92、平均S-CVI为0.86，提示量表内容效度良好[12]，能够有效反映医学生医疗人工智能准备的各个方面。本研究提取出3个公因子，量表的累计方差贡献率为54.03%，提示中文版MAIRS-MS结构效度达标[14]。收敛效度方面，中文版MAIRS-MS的3个维度组合信度大于0.70、平均方差抽取值大于0.36[17]，说明中文版MAIRS-MS收敛效度可以接受，内部一致性合格。此外，量表各维度间的相关系数均小于各维度对应的平均方差抽取值的平方根，表明中文版MAIRS-MS区分效度符合标准。本研究选择人工智能素养量表作为中文版MAIRS-MS的效标，Pearson相关系数法分析结果显示二者总分呈正相关关系，且相关系数大于0.40[18]，提示中文版MAIRS-MS效标关联效度达标。综合来说，中文版MAIRS-MS的效度合格，能够反映测评结果与实际测量内容的一致性程度。

**3.3**　中文版MAIRS-MS信度可接受

本研究结果发现，量表各维度的Cronbach's $\alpha$系数为0.71~0.78，量表的总体Cronbach's $\alpha$系数为0.85，提示量表的内部一致性可以接受[12]；量表的折半信度为0.89，说明量表的内部一致性达标；总量表的重测相关系数大于0.60[19]，表明量表的稳定性可以接受。这些结果表明中文版MAIRS-MS信度可以接受，具有合格的稳定性和内部一致性。

**3.4**　中文版MAIRS-MS实用性较好

在人工智能与医疗深度融合的背景下，医学人才需兼备传统医学知识与现代人工智能技术[20]。当前国内缺乏经过科学验证的医疗人工智能准备度评估工具，无法准确定位医学生医疗人工智能相关能力的优势与短板，导致中国医疗人工智能教育目标不够明确[21]，进而制约复合型医学人才培养，需要经过科学验证的医疗人工智能准备度测评工具完善国内医疗人工智能教育。本研究对MAIRS-MS进行汉化，并通过严谨的跨文化调适，最终形成中文版MAIRS-MS。该量表涵盖基础知识、实践能力、态度和价值观3个维度，从多层面评估医学生的医疗人工智能准备状况，具有较强的针对性与实用性。中文版MAIRS-MS能够帮助医学教育者更全面、深入地掌握医学生在医疗人工智能准备度上的优势和不足，本研究的结果提示医学生医疗人工智能相关的价值观水平处于较高水平、实践能力处于相对较高水平但仍有提升空间、基础知识处于中等水平。可能的原因包括：一是近年来人工智能领域的快速发展与热潮，推动医学生对医疗人工智能形成了积极的态度与价值观；二是医学生了解人工智能的渠道较为单一，绝大多数依赖媒体或社交媒体，通过学校专业讲座等正规渠道获取相关信息的学生占比极低。同时，多数医学院校尚未针对医学生开设系统、专业的医疗人工智能知识与技能培训课程[22]，这直接限制了医学生医疗人工智能相关基础知识的积累与提升。鉴于医学生在医疗人工智能基础知识、实践能力上仍有提升潜力，且院校层面的系统专业培训存在明显欠缺，各医学院校应在现

有医学专业课程体系中，增设专门的医疗人工智能应用课程，帮助医学生系统夯实基础知识；同时积极与大型综合医院、专业医疗人工智能企业建立长期稳定的合作关系，搭建校外实习实践基地，为医学生提供更多医疗人工智能实践机会，提升其医疗人工智能实践能力。

### 3.5 研究局限性

本研究存在一定局限性。首先，本研究采用便利抽样，样本代表性受限，进而限制了研究结果的推广性。其次，本研究纳入的医学生数量有限。未来可选用更具代表性的抽样方法、扩大样本量，进一步验证中文版 MAIRS-MS 的信效度，并构建医学生医疗人工智能准备度常模。最后，在评估重测信度时，本研究仅选取 30 名医学生作为样本，样本量较小，可能致使重测信度结果的稳定性和代表性欠佳，未来需扩大样本量，验证中文版 MAIRS-MS 在不同专业医学生中的重测信度。

## 4 结语

本研究完成了对 MAIRS-MS 的翻译、回译、跨文化调适及信效度检验，最终构建出中文版 MAIRS-MS。该量表信效度达标，可用于评估中国医学生的医疗人工智能准备度，有助于了解中国医学生医疗人工智能准备度的现状，帮助医学教育者掌握医学生在这一方面的优势与不足，进而开展针对性的医疗人工智能教育改革。

**利益冲突**　所有作者声明不存在利益冲突

**作者贡献声明**　何登鑫：提出论文构思、撰写论文；陈芳、赵梅村：总体把关、审订论文

**参考文献**

[1] 吴敏敏，王鑫钰，王伟炳. 医学人工智能的技术发展和场景应用[J]. 复旦学报(医学版)，2025，52(3)：470-474.

[2] 清华大学万科公共卫生与健康学院，北京大学公共卫生学院，中国医师协会全科医师分会，等. 智能全科医生中国专家共识[J]. 中国全科医学，2025，28(2)：135-142. DOI: 10.12114/j.issn.1007-9572.2024.0453.

[3] 杨子雨，朱锡雨，虞隽扬，等. 人工智能在介入医学中的应用及前景[J]. 介入放射学杂志，2025，34(4)：441-444. DOI: 10.3969/j.issn.1008-794X.2025.04.019.

[4] 卢玉，于书帆，姜秉寅，等. 人工智能：新时代化学、药学与医学研究的强大助力：2024 年诺贝尔化学奖[J]. 首都医科大学学报，2024，45(6)：1131-1137. DOI: 10.3969/j.issn.1006-7795.2024.06.026.

[5] 郭华源，刘盼，卢若谷，等. 人工智能大模型医学应用研究[J]. 中国科学：生命科学，2024，54(3)：482-506.

[6] Paranjape K, Schinkel M, Nannan Panday R, et al. Introducing artificial intelligence training in medical education [J]. JMIR Med Educ, 2019, 5(2): e16048. DOI: 10.2196/16048.

[7] 贾萌，张瑛琪，李云峰，等. 医学人工智能教育的研究进展[J]. 医学教育研究与实践，2023，31(1)：1-6. DOI: 10.13555/j.cnki.c.m.e.2023.01.001.

[8] Karaca O, Çalışkan SA, Demir K. Medical artificial intelligence readiness scale for medical students (MAIRS-MS): development, validity and reliability study [J]. BMC Med Educ, 2021, 21(1): 112. DOI: 10.1186/s12909-021-02546-6.

[9] 郭金玉，李峥. 量表引进的过程及评价标准[J]. 中华护理杂志，2012，47(3)：283-285. DOI: 10.3761/j.issn.0254-1769.2012.03.039.

[10] 王家良. 临床流行病学临床科研设计、测量与评价[M]. 3 版. 上海：上海科学技术出版社，2009：231-232.

[11] 周琼，徐亚苹，蔡迎春. 高校学生人工智能素养能力现状及影响因素多维分析[J]. 图书情报知识，2024，41(3)：38-48. DOI: 10.13366/j.dik.2024.03.038.

[12] 张晨，周云仙. 我国护理测量工具文献中内容效度指数应用误区分析[J]. 护理学杂志，2020，35(4)：86-88，92. DOI: 10.3870/j.issn.1001-4152.2020.04.086.

[13] 李雯，魏丽丽，张艳，等. 产妇出院指导质量量表的汉化及信效度检验[J]. 中华护理杂志，2021，56(2)：315-320. DOI: 10.3761/j.issn.0254-1769.2021.02.028.

[14] Schreiber JB. Issues and recommendations for exploratory factor analysis and principal component analysis [J]. Res Social Adm Pharm, 2021, 17(5): 1004-1011. DOI: 10.1016/j.sapharm.2020.07.027.

[15] 温忠麟，侯杰泰，马什赫伯特. 结构方程模型检验：拟合指数与卡方准则[J]. 心理学报，2004，36(2)：186-194.

[16] 肖惠敏，林晓丹，吴红，等. 晚期癌症患者生存痛苦量表的汉化及信效度检验[J]. 护理学杂志，2022，37(9)：74-77. DOI: 10.3870/j.issn.1001-4152.2022.09.074.

[17] 余构雄，刘博，尹书华. 入境商务游客视角下的城市目的地形象维度构成及量表开发[J]. 资源开发与市场，2023，39(11)：1409-1417. DOI: 10.3969/j.issn.1005-8141.2023.11.001.

[18] 周钰，李刚卫，李弯月，等. 语言认知评估系统用于语言认知障碍评估的可行性分析[J]. 中国全科医学，2022，25(31)：3872-3876，3890. DOI: 10.12114/j.issn.1007-9572.2022.0474.

[19] 吴明隆. 问卷统计分析实务 SPSS 操作与应用[M]. 重庆：重庆大学出版社，2010：191-192.

[20] 唐雷，钟世镇. 人工智能及其在医学领域中的应用[J]. 中国实用妇科与产科杂志，2024，40(9)：876-878. DOI: 10.19538/j.fk2024090104.

[21] 王艳，初红，梁静静，等. 人工智能背景下高等医学教育的改革现状与探究[J]. 中国医药导报，2022，19(13)：64-67. DOI: 10.20047/j.issn1673-7210.2022.13.011.

[22] 宋元明. “人工智能+医学”新医科人才培养探索：以部分高校实践为例[J]. 中国高校科技，2020(8)：65-68. DOI: 10.16209/j.cnki.cust.2020.08.015.

（收稿日期：2025-09-23）

（本文编辑：兰明娟）