

DOI:10.13288/j.11-2166/r.2024.22.017

中医教育

大规模中医师考试评价中认知诊断模型的构建及应用

江哲涵^{1,2}, 徐玲玲², 潘姝丞³, 金阿宁¹✉

1. 国家中医药管理局中医师资格认证中心, 北京市西城区北三环中路3号, 100029; 2. 北京大学全国医学教育发展中心;
3. 北京大学护理学院

[摘要] 目的 采用认知诊断模型对中医师岗位胜任力第二阶梯考核理论考试进行分析, 为大规模医学考试的认知诊断研究提供参考。方法 选择2023年参加中医师岗位胜任力第二阶梯考核理论考试的16 310位考生作答情况开展认知诊断分析。检验拓广决定型输入噪声与门模型(G-DINA模型)和考试数据之间的拟合程度; 在认知诊断框架下进行项目参数分析和信度分析; 从不同层面综合分析考生的认知属性掌握情况。结果 本研究中16 310名考生平均分数为(189.76±40.86)分, 共有9994名考生通过了此次考试, 通过率为61.28%。考试总题量300题, 五个模块中中医基础考核频率最高(132题, 占44%), 其次是中医临床(87题, 占29%)、西医基础(60题, 占20%), 中医经典(12题, 占4%)和医学人文(9题, 占3%)考核比例较少。模型拟合指标结果显示, G-DINA模型曲线下面积(AUC)为0.98, 说明该模型与本研究数据拟合优秀。项目参数分析结果显示, 300道测试题目认知诊断模型的所有区分度均大于0, 平均区分度为0.25, 整体题目区分度的表现良好。信度分析结果显示, 测验水平分类准确率为0.75, 属性水平分类准确率为0.83~0.99, 表明G-DINA模型的属性分类准确度高。认知诊断分析结果显示, 相较于未通过考试的考生, 已通过考试的考生对于各模块认知属性的掌握比例更高。结论 基于认知诊断的大规模中医考试分析, 可为提高中医学人才教育和培养质量提供支撑。

[关键词] 中医执业医师资格考试; 中医师岗位胜任力; 认知诊断; 中医教育; 标准化考试

根据当前《深化新时代教育评价改革总体方案》^[1]要求, 新一轮教育评价改革背景下, 传统的“以结果为导向的鉴定式评价”正逐渐转向“以改进为目的的诊断式评价”。以认知诊断为代表的第三代测评理论引起广泛关注, 成为新时代教育评价改革的重要着力点^[2]。中医学是医学的重要组成部分, 中医执业医师资格考试是中医师人才评价的核心制度, 对发挥中医药人才基础性、战略性、决定性作用起到关键推动作用^[3]。为全面加强新时代中医药人才工作, 深入贯彻《“十四五”中医药人才发展规划》, 国家中医药管理局中医师资格认证中心组织开展中医师岗位胜任力阶梯实证研究(post competency ladder assessment, PCLA)^[4], 覆盖医学

生向医师培养转化的全周期, 促进院校教育、毕业后教育的有效衔接, 旨在提高中医学人才培养和准入质量, 推动中医药教育改革与高质量发展。本研究以中医师岗位胜任力第二阶梯考核为例, 对我国中医药大规模标准化考试进行认知诊断分析与评价。通过科学建模与参数分析, 探讨该类考试测验质量、考试区分度与考生认知属性, 构建具有诊断功能的中医药考试分数报告案例, 以期为我国中医学考试命题及人才选拔提供理论依据和数据支持。

1 研究对象与方法

国家中医药管理局中医师资格认证中心于2023年将中医学类专业(本科)水平测试更名为中医师岗位胜任力第二阶梯考核。2023年度中医师岗位胜任力第二阶梯考核于2023年5月12—14日举行, 全国41所院校考生参加考试。该考试主要评估中医学专业本科生, 在完成全部理论学习并

基金项目: 国家自然科学基金(72104006); 国家中医药考试2023年度科研课题(TA2022002); 国家资助博士后研究人员计划(GZC20240067)

✉ 通讯作者: jinaning@qq.com

准备进入临床实习前是否具备该阶段所要求的医学人文素养、医学基础理论知识与临床基本技能。该考核由两部分组成：医学综合考试（简称理论考试，采取计算机化形式）和实践技能考试。本研究仅针对医学综合考试进行认知诊断分析。

1.1 研究对象

本研究涉及对象为所有参加 2023 年度中医师岗位胜任力第二阶梯理论考试的中医学类专业（包括中医学、针灸推拿学、中医康复学、中医养生学、中医儿科学、中医骨伤科学等）五年及以上学制本科四年级在校学生，考生人数共 16 310 名。

1.2 研究方法

1.2.1 认知诊断基础 认知诊断的核心之一是 Q 矩阵，作为认知属性的量化表现形式，Q 矩阵体现了试题与认知属性间的关系，是整个认知诊断分析的基础和关键^[5]。然而对目前广泛存在的非诊断性大型标准化测验来说，确定 Q 矩阵是一项极具挑战的工作^[6]。一般而言，Q 矩阵的构建可参见双向细目表。本研究依托国家中医药管理局中医师资格认证中心，在开发试题过程中已组织全国教育教学、临床医学和心理测量等领域专家，对该理论考试的认知属性进行详尽探讨和标定，并在专家标定的双向细目表的基础上构建医学综合考试 Q 矩阵，如表 1 所示，Q 矩阵的标定为本研究建立模型提供基础和前提。

表 1 医学综合考试 Q 矩阵（示例）

Table 1 Q-matrix for the Comprehensive Medical Examination (example)

项目	属性				
	A1	A2	A3	A4	A5
题号 1	0	1	0	0	0
题号 2	1	0	0	0	0
题号 3	0	1	0	0	0
题号 4	0	0	0	1	0
题号 5	1	0	0	0	0
题号 6	0	0	0	0	1
题号 7	0	0	0	0	1
题号 8	1	0	0	0	0
题号 9	0	0	1	0	0
题号 10	0	0	1	0	0
.....					

注：A1，西医基础；A2，医学人文；A3，中医基础；A4，中医经典；A5，中医临床；1 表示该项目对应的认知属性，0 表示未对应属性。

1.2.2 模型筛选与拟合评估 认知诊断模型（cognitive diagnostic model, CDM）按其适用范围广度

可分为简化 CDM 和饱和 CDM。其中，拓广决定型输入噪声与门模型（G-DINA 模型）作为饱和模型的代表，是目前研究的主流选择^[7]。该模型假设相对宽松，认为试题各认知属性对试题答对概率有不同的贡献比例，掌握部分认知属性的被试也有一定答对概率。医学综合考试能力的各组成部分之间非简单的线性关系，而是一种复杂、多元关系，即各种知识、技能等相互作用、相互影响且不可分割。中医师岗位胜任力第二阶梯考核的目标属性具有抽象、综合的特征，对认知诊断模型要求高，补偿型模型可有效契合医学综合考试能力的综合性和多元性，饱和模型则可以比较理想地对应医学综合考试能力的抽象性和难区分性。因此，本研究选用兼具补偿性和饱和性特征的 G-DINA 模型进行认知诊断分析。

模型评估采用绝对拟合指标，即曲线下面积（area under curve, AUC）检验拟合结果。AUC > 0.50 表明模型的表现性能优于随机猜测，AUC 为 0.50~0.69 意味着模型拟合效果可接受，AUC 为 0.70~0.90 提示拟合效果良好，AUC > 0.90 表明该模型拟合优秀^[8]。

1.2.3 信度分析 信度主要从测验可靠性、一致性和稳定性评价测验质量，系衡量测验或测量工具质量的关键指标，是分数报告合理性的重要保证^[9]。因此，信度分析是认知诊断的必要步骤。本研究采用自然指标分类准确率评估测验水平和属性水平诊断的信度^[10-11]，准确率取值范围 [0, 1]，数值越接近 1，表明模型的分类精准度越高。

1.2.4 区分度分析 区分度亦称鉴别力，反映了测验题目对考生实际水平的区分程度或鉴别能力^[12]。一般来说，区分度在 0.40 以上表明试题区分程度很好，区分度 0.30~0.39 表明较好，区分度 0.20~0.29 表明尚可，区分度 0.20 以下表明较差。本研究基于标定的 Q 矩阵框架，对 300 道测试题采用 G-DINA 模型^[13]进行参数估计。参数估计的过程使用 R 语言 G-DINA 包^[14]，模型参数估计采用极大似然估计（maximum likelihood estimation, MLE），属性掌握状态估计采用极大后验估计（maximum a posteriori estimation, MAP）。认知诊断模型分析过程中，根据 De La Torre^[15]提出的指标计算题目区分度，该指标公式被定义为掌握了题目 j 所有属性的被试答对概率 $p_j(1)$ 减去未掌握题目 j 任何属性的被试答对概率 $p_j(0)$ ，表达式为： $d_j = p_j(1) - p_j(0)$ 。本研究对所有测验题目均进行区分度分析。

1.2.5 认知诊断分析 认知属性描述了考生完成

特定任务时所需的知识、技能和策略等，反映了考生解决问题时心理加工过程的内在机制。通过分析考生的认知属性掌握模式，可精确识别其知识掌握类型、当前知识状态及未来需改进的方向，为教师提供有针对性的补救教学依据。基于中医执业医师资格考试大纲五大部分，该理论考试考生的认知属性包括西医基础（A1）、医学人文（A2）、中医基础（A3）、中医经典（A4）、中医临床（A5）五个模块。本研究基于 G-DINA 模型的估计结果，分析不同考生群体对于五个模块认知属性的掌握情况（即计算考生答对的题目数量与该认知属性下的所有题目数量的比例），并通过雷达图和知识点关系图可视化呈现考生的认知属性掌握模式及属性间的关系。

1.3 统计学及绘图方法

本研究中的所有统计分析过程均依托 R 语言软件进行。针对考生基本信息的描述性统计分析与模型拟合度检验，本研究采用 R 语言自编程序对数据进行分析与验证。在评估题目区分度、测验信度及考生属性掌握模式方面，本研究利用 G-DINA 包，并在此基础上通过自编程序进一步拓展了分析维度。此外，为了直观展示考生的属性掌握模式以及属性掌握模式的概率，本研究采用 fmsb 包与 psych 包联合绘制了雷达图；为了揭示知识点之间的复杂关系与结构，本研究引入了 igraph 包进行知识点关系图的绘制。

2 结果

2.1 考试基本情况

中医医师岗位胜任力第二阶梯考核的理论考试满分为 300 分，本研究中 16 310 名考生考试得分 3~289 分，平均分数为（189.76±40.86）分，中位数为 193 分。按照 180 分合格标准，共有 9994 考生通过了此次考生，本轮考试通过率为 61.28%。

2.2 认知属性分布情况

考试总题量 300 题，五个模块中 A3 考核频率最高（132 题，占 44%），其次是 A3（87 题，占 29%）、A1（60 题，占 20%），A4（12 题，占 4%）和 A2（9 题，占 3%）考核比例较少。Q 矩阵中行为试题、列为属性，并以 0/1 计分，如某题双向细目考察 A2，则 Q 矩阵中该对应行为 [0, 1, 0, 0, 0]。本研究所有题目均系单属性。

2.3 CDM 分析结果

2.3.1 模型拟合检验

本研究对中医医师岗位胜任力第二阶梯考核的理论考试数据进行拟合检验，G-DINA 模型 AUC 为 0.98，该模型的拟合效果优秀，表明所选模型与中医医师岗位胜任力第二阶梯考核理论考试的拟合程度高，即数据对模型具有强大支撑作用。

2.3.2 分类准确度 本研究信度分析结果显示，2023 年中医医师岗位胜任力第二阶梯考核理论考试的测验水平分类准确率为 0.75；A1 分类准确率为 0.99，A2 为 0.94，A3 为 0.98，A4 为 0.96，A5 为 0.83。除 A5 模块外，其他模块属性水平分类准确率均在 0.90 以上，表明 G-DINA 模型的属性分类准确度高。

2.3.3 测试题目区分度 本研究对理论考试的 300 道测试题目进行认知诊断区分度分析。结果显示，所有题目的区分度均值为 0.25，标准差为 0.11，最小值为 0.01，最大值为 0.51。该结果表明医学基本知识考试测试题目一定程度上可以区分完全掌握和未掌握考核认知属性的被试群体，少部分试题的区分度仍有待提高。

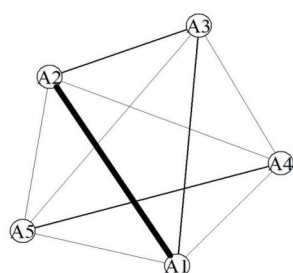
2.3.4 诊断分析结果 按照考生是否通过考试，比较不同考生群体在五个模块认知属性的掌握情况。认知诊断结果显示，相较于未通过考试的考生，已通过考试的考生对于各模块认知属性的掌握比例更高，这表明属性掌握情况估计结果的合理性。已通过考试的考生掌握程度最高的模块是 A3 为 88%，而未通过考试的考生掌握程度最高的模块是 A5 为 11%，详见表 2。

表 2 中医医师岗位胜任力第二阶梯考核理论考试不同考生群体对各属性掌握情况（%）

Table 2 Mastery of each attribute by different groups of candidates in the theory examination of the second-tier assessment of competence for traditional Chinese medicine physicians（%）

模块	全体		通过		未通过	
	未掌握	掌握	未掌握	掌握	未掌握	掌握
西医基础（A1）	47	53	16	84	97	3
医学人文（A2）	49	51	19	81	96	4
中医基础（A3）	44	56	12	88	94	6
中医经典（A4）	57	43	32	68	97	3
中医临床（A5）	54	46	33	67	89	11

基于考生对各认知属性的掌握情况，本研究进一步构建了不同模块之间关系的无向图，如图 1 所示。连线的粗细代表模块之间关系的强弱，图中仅保留相关系数值为 0.5 以上的关系连线，可以看出，认知属性之间大多具有中等以上的相关关系，其中模块 A1 与 A2 之间关联最紧密。



注：A1，西医基础；A2，医学人文；A3，中医基础；A4，中医经典；A5，中医临床。

图1 中医师岗位胜任力第二阶梯考核理论考试的知识点关系无向图

Figure 1 Undirected graph of the relationship of knowledge points in the theory examination of the second-tier assessment of competence for traditional Chinese medicine physicians

2.3.5 编制考生分数报告 进一步分析每位考生在不同属性上的掌握概率，即知识结构，这是一类以属性为基本分析单位的精细化测量模型。如图2所示，两位考生的得分相同，都为180分，但是两位考生表现出了不同的知识结构。通过认知诊断G-DINA模型进行参数估计，得到其中1位考生(图中粗实线条表示)知识状态为(10011)，这说明该考生掌握了A1、A4、A5，而未掌握A3和A2。另1名考生(图中虚线表示)知识状态为(11010)，即掌握了A1、A2和A4，而未掌握A3和A5。借助雷达图，可以了解考生不同的属性掌握模式以及属性掌握模式的概率，同时为制定个性化评价和进一步制定补救学习方案提供重要依据。

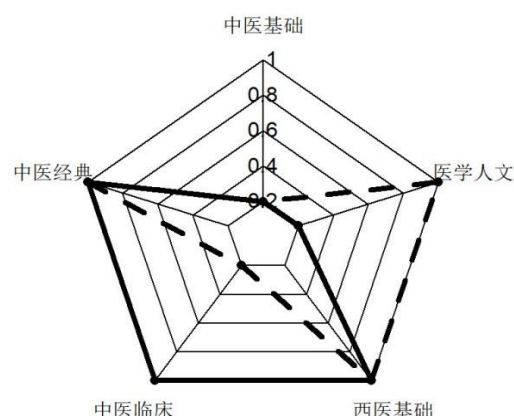


图2 中医师岗位胜任力第二阶梯考核理论考试得分相同的两位考生的知识结构比较雷达图

Figure 2 Comparative radar chart of the knowledge structure of the two candidates who scored the same on the theory examination of the second-tier assessment of competence for traditional Chinese medicine physicians

3 讨论

如何改革人才培养模式、优化人才成长途径对于促进中医药传承创新发展至关重要。中医师岗位胜任力第二阶梯考核是评估中医学专业本科生是否具备该阶段所要求的医学人文素养、医学基础理论知识与临床基本技能的重要标志，可有效衔接院校教育与毕业后教育，对于我国中医师人才培养具有重要作用。认知诊断作为新兴的测量理论，融合了认知心理学和现代测量学的理论和方法，在教育领域受到广泛关注^[16]。借助认知诊断模型，不仅可以对考生的微观认知结构进行诊断分析，还能对测试群体的宏观能力进行评估^[17]。本研究纳入分析2023年度16 310名考生中医师岗位胜任力第二阶梯考核理论考试情况，依据认知诊断模型的全面流程，采用G-DINA模型逐步进行实证Q矩阵、模型设定、模型与数据拟合评估、信度分析、区分度分析、项目诊断以及结果解读。通过认知诊断分析，对考生的中医基础、中医临床、中医经典、西医基础、医学人文模块的掌握情况加以预测和评估，实现了对考生的认知优势与劣势的诊断，有利于准确、高效地找出当前我国中医师岗位胜任力第二阶梯考核理论考试中可能存在的问题，从而有针对性地推进中医学类教育考试改革，促进我国中医药类人才选拔水平的提高。

3.1 确定认知属性是认知诊断建模的基础

对于中医知识的掌握情况、辨证论治的临床能力、中医思维等都是中医师执业考试的重点考核内容，且执业医师考试与中医师的岗位胜任力密切相关^[18]。在基于认知诊断理论的建模框架中，用“认知属性”来描述测量任务中被测的认知过程或认知技能，其可以是认知过程中所涉及的知识或技能，也可以是对被测心理内部加工过程的诊断基础。本研究将中医师岗位胜任力第二阶梯考核理论考试五个模块选定为认知诊断分析所需的认知属性，按照不同考试模块划分，认知属性包括中医基础、中医临床、中医经典、西医基础和医学人文。需要强调的是，中医药考试涵盖内容广泛、多样，除按模块分类外，还可以根据科目、题型或认知层次等方式标定题目认知属性。认知属性粒度越细，其测验结构越复杂，因此，随着属性数量增多，未来研究可能需要考虑计算瓶颈和测量问题，以确保计算效率和测量准确性。

3.2 合理标定Q矩阵是认知诊断的重要环节

目前，在外语等领域^[7]已经逐步实现认知诊

断建模自动化与诊断测评智能化,在中医学领域认知诊断研究相对较少。本研究纳入分析 2023 年度中医师岗位胜任力第二阶梯考核理论考试作答 300 道题情况,其考试规模较大,考题量较多。在进行中医考试认知诊断分析过程中,Q 矩阵是对测验项目与认知属性关系的描述。Q 矩阵的合理性构建直接影响项目参数估计及被试分类的准确性。本研究采用前期专家标定的 Q 矩阵,已有研究提出采用数据驱动的 Q 矩阵标定与验证方法^[19-20]和基于探索性因素分析的 Q 矩阵标定方法^[21],可以提高 Q 矩阵标定的精度。中医师岗位胜任力第二阶梯考核理论考试“以岗位胜任力为导向”,科学考评考生的综合医学知识水平,通过合理标定 Q 矩阵有助于未来构建科学的中医药专业考试框架。

3.3 编制考生分数报告以指导后续改进

传统测验结果提供宏观分数,这种反馈仅能体现考生的总体水平及其在群体中的排名,无法精准定位考试的问题,也无法提供具有针对性个性化的细致反馈^[22];同时对“具有相同分数或相同能力的考生具有不同的知识状态及不同的认知结构”这一现象无法做出解释,不能剖析考生的认知结构存在什么样的问题,故该评价方式现已不能满足考试需求。认知诊断作为一种新型评价方式,通过分析考生心理内部加工过程,不仅能够得到宏观的分数,还能评估考生的认知结构,并提供具有诊断功能的反馈报告,挖掘考生不足之处并提出补救性的学习建议。这种多维度、细颗粒度的反馈信息方式,突破了“分数”维度评价的限制,遵循“诊断-反馈-补救”为一体的目标,实现以评促学,以评促教,稳步提高考生中医理论水平^[23]。本研究基于认知诊断分析结果编制了考生分数报告,该报告可以帮助考生在后续学习中针对自身薄弱环节有目的地进行改进,体现了个性化教育以及因材施教的素质教育,有利于新时代中医药卫生人才培养。

4 小结

本研究在现代测量理论框架下,针对中医师岗位胜任力第二阶梯考核理论考试,探索了基于认知诊断分析、评价中医师考试的路径。研究通过数据科学建模驱动分数估计,制作诊断式分数报告模板,验证认知诊断模型可行性,并绘制了目标属性的知识谱图。研究结果体现了认知诊断分析在处理复杂考试数据和进行精细化评估方面的优异表现,也为医学教育评价领域提供了重要的理论和实践参考。在理论层面,本研究将前沿的教育测量学与认

知心理学理论与中医知识、技能和属性相结合,构建了具有中医特色的认知诊断理论,有助于反馈教育和教学。在应用层面,本研究为中医师岗位胜任力第二阶梯考核提供了分数报告,为考生和医学教育工作者提供了有针对性的改进信息,并为命题工作提供了指南,推动了具有认知诊断特性的标准化评价形式的发展。未来可以在此基础上,进一步获得各类中医药考试底层数据,利用认知诊断进行综合分析,甚至可以基于 CDM 的新方法探索自适应测试,推动构建中医药类教育考试的评价体系与标准体系。

参考文献

- [1]新华社.中共中央 国务院印发《深化新时代教育评价改革总体方案》[EB/OL]. (2020-10-13)[2024-03-29]. http://www.gov.cn/zhengce/2020-10/13/content_5551032.htm.
- [2]李振,周东岱.基于学科知识图谱的智能化认知诊断评估方法[J].现代教育技术,2022,32(11):118-126.
- [3]徐敢,罗卫花.以医师资格考试制度为轴心的医教考协同发展政策探索:以中医执业医师资格考试为例[J].中医教育,2024,43(2):35-40.
- [4]国家中医药管理局,教育部,人力资源社会保障部,等.国家中医药局 教育部 人力资源社会保障部 国家卫生健康委《关于加强新时代中医药人才工作的意见》[EB/OL]. (2022-04-08)[2024-03-29]. https://www.gov.cn/zhengce/zhengceku/2022-06/23/content_5697348.htm.
- [5]汪大勋,高旭亮,蔡艳,等.一种广义的认知诊断 Q 矩阵修正新方法[J].心理科学,2019,42(4):988-996.
- [6]刘慧,李亚男.基于认知诊断理论的国际中文阅读测验 Q 矩阵的构建[J].华文教学与研究,2023(4):68-76.
- [7]范婷婷,孙波,曾用强.英语阅读测试中不同认知诊断模型比较研究[J].中国考试,2023(5):53-61.
- [8]BI X, WANG Y, LI M, et al. Validation of the Memorial Sloan Kettering Cancer Center nomogram for predicting non-sentinel lymph node metastasis in sentinel lymph node-positive breast-cancer patients [J]. Onco Targets Ther, 2015, 8: 487-493. doi: 10.2147/OTT.S78903.
- [9]温忠麟,黄彬彬,汤丹丹.问卷数据建模前传[J].心理科学,2018,41(1):204-210.
- [10]IACONANGELO C. Uses of Classification error probabilities in the three-step approach to estimating cognitive diagnosis models [D]. New Brunswick NJ: Rutgers University, 2017.
- [11]WANG W, SONG L, CHEN P, et al. Attribute-level and pattern-level classification consistency and accuracy indices for cognitive diagnostic assessment [J]. J Educ Measur, 2015, 52: 457-476. doi:10.1111/jedm.12096.

- [12]杨志明. 提高选择题区分度的若干方法[J]. 教育测量与评价, 2017(2): 5-10.
- [13]DE LA TORRE J. The generalized DINA model framework[J]. Psychometrika, 2011, 76(2): 179-199.
- [14]MA W, DE LA TORRE J. GDINA: an R package for cognitive diagnosis modeling [J]. J Statist Software, 2020, 93(14): 1-26.
- [15]DE LA TORRE J. An empirically based method of Q-matrix validation for the DINA model: development and applications[J]. J Educat Measur, 2008, 45(4): 343-362.
- [16]GEORGE AC, ROBITZSCH A. Validating theoretical assumptions about reading with cognitive diagnosis models [J]. Int J Test, 2021, 21(2): 105-129.
- [17]张怡. 基于认知诊断的学习测验开发、应用与启示:以 Tatsuoka 的分数减法测验为例[J]. 教育测量与评价, 2022(5): 71-82.
- [18]于胜君, 龚张斌, 项乐源. 基于岗位胜任力的中医执业医师资格考试改革方向的思考[J]. 中医药管理杂志, 2023, 31(12): 61-63.
- [19]涂冬波, 蔡艳, 戴海琦. 基于 DINA 模型的 Q 矩阵修正方法[J]. 心理学报, 2012, 44(4): 558-568.
- [20]喻晓锋, 罗照盛, 秦春影, 等. 基于作答数据的模型参数和 Q 矩阵联合估计[J]. 心理学报, 2015, 47(2): 273.
- [21]汪文义, 宋丽红, 丁树良. 基于探索性因素分析的 Q 矩阵标定方法[J]. 江西师范大学学报: 自然科学版, 2015, 39(2): 138-144.
- [22]DE LA TORRE J, DER ARK LAVAN, ROSSI G. Analysis of clinical data from a cognitive diagnosis modeling framework [J]. Measur Evaluat Counsel Develop, 2018, 51(4): 281-296.
- [23]赵娅玲, 郭晓梅, 杨若俊, 等. 中医执业医师资格考试通过率较低的原因及对策[J]. 中国中医药现代远程教育, 2021, 19(14): 192-194.

Analysis and Assessment of a Large-Scale Traditional Chinese Medicine Examination Based on Cognitive Diagnostic

JIANG Zhehan^{1,2}, XU Lingling², PAN Shucheng³, JIN A'ning¹

1. Chinese Medicine Qualification Center of the National Administration of Traditional Chinese Medicine, Beijing, 100029; 2. National Center for Health Professions Education Development, Peking University; 3. School of Nursing, Peking University

ABSTRACT Objective To analyze the scores of the second-tier assessment on theory of competency of Traditional Chinese Medicine (TCM) physicians based on cognitive diagnostic model, and to provide a reference for cognitive diagnostic studies of large-scale medical examinations. **Methods** Cognitive diagnostic analyses were conducted by selecting the responses of 16,310 candidates who took the theory examination of the second-tier assessment on the competency of TCM physicians in 2023. The degree of fit between the topological deterministic input noise and gate model (G-DINA model) and the examination data was examined; item parameter analysis and reliability analysis were conducted under the cognitive diagnostic framework; the candidates' mastery of cognitive attributes was comprehensively analysed at different levels. **Results** The mean score of 16,310 candidates in this study was (189.76 ± 40.86) , and a total of 9,994 candidates passed this examination, with a pass rate of 61.28%. The total number of questions in the examination was 300, and among the five modules, the frequency of assessment of TCM fundamentals was the highest (132 questions, 44%), followed by TCM clinic (87 questions, 29%), Western medicine fundamentals (60 questions, 20%), and TCM classics (12 questions, 4%) and medical humanities (9 questions, 3%) was less. The results of the model fit index showed that the area under the curve (AUC) of the G-DINA model was 0.98, indicating that the model fit the study data excellently. The results of item parameter analysis showed that all the discriminations of the cognitive diagnostic model for the 300 test questions were greater than 0, with an average discrimination of 0.25, and the overall performance of the question discriminations was good. The results of the reliability analysis showed that the test level classification accuracy was 0.75, and the attribute level classification accuracy was 0.83 to 0.99, indicating that the G-DINA model has high attribute classification accuracy. The results of the cognitive diagnostic analyses showed that the proportion of candidates who had passed the exam was higher for the cognitive attributes of each module compared to those who had not passed the exam. **Conclusion** The analysis of large-scale TCM examinations based on cognitive diagnosis can provide support for improving the quality of TCM talent education and cultivation.

Keywords qualification examination for traditional Chinese medicine practitioners; competence of traditional Chinese medicine practitioners; cognitive diagnosis; traditional Chinese medicine education; standardized test

(收稿日期: 2024-04-30; 修回日期: 2024-07-04)

[编辑: 邓媛]