

DOI: 10.13288/j.11-2166/r.2016.08.006

思路与方法

基于项目反应理论的中风病证候要素评价量表的条目筛选

张杨¹, 高颖^{2*}, 马斌², 钟海珍³, 柳伟伟⁴, 王琪⁴

(1. 首都医科大学附属北京朝阳医院, 北京市朝阳区工体南路 8 号, 100020; 2. 北京中医药大学东直门医院; 3. 南京医科大学第二附属医院; 4. 中国军事科学院生物信息统计咨询中心)

摘要 目的 基于项目反应理论对中风病证候要素评价量表条目进行筛选与优化。方法 收集 410 例中风病患者临床信息, 在前期研究所形成的条目池基础上, 采用两参数的 Logistic 项目反应模型分别计算各条目的项目信息函数, 以删除项目信息函数值 < 0.1 的条目为标准进行条目筛选。结果 内风证分量表中项目信息函数值 ≥ 0.1 的条目数过少, 故将其条目按项目信息函数值 < 0.01 进行删除, 共删除近 48 h 内病情加重或波动、瞳神异常 (瞳孔缩小、散大或不等)、口噤、项强、肢体僵硬状态、舌短缩、弦脉 7 个条目; 按照筛选标准内火证删除滑脉 1 个条目; 痰湿证删除恶心呕吐、胖大舌或齿痕舌、厚苔、腻苔 4 个条目; 血瘀证删除肌肤不仁、涩脉 2 个条目; 气虚证删除大便或小便失禁、缓脉或弱脉、细脉 3 个条目; 阴虚证删除耳鸣、腰膝酸软、弦脉、细脉 4 个条目。结论 形成了含 76 个条目的基于项目反应理论的中风病证候要素评价量表 (初表), 其中内风证 8 个条目、内火证 24 个条目、痰湿证 12 个条目、血瘀证 8 个条目、气虚证 15 个条目、阴虚证 9 个条目。

关键词 项目反应理论; 中风病; 证候评价量表; 条目筛选

项目反应理论也称潜在特质理论或潜在特质模型, 是一种现代心理测量理论, 其意义在于可以指导项目的筛选和测验编制^[1]。中医证候具有“内实外虚、多维界面、动态时空”的特点^[2], 不同四诊信息对同一证候要素轻重程度的测评意义并不等价, 因此很难满足经典测量理论真分数与观测分数间存在简单线性关系的假定, 而项目反应理论假定观察变量与其潜在特质间存在非线性关系, 可能更符合中医证候评价量表条目的筛选。因此, 本研究基于项目反应理论对中风病证候要素评价量表条目进行筛选与优化, 报告如下。

1 资料与方法

1.1 中风病证候要素评价量表条目池的形成

“证候评价”是一种对证候轻重程度的标准化、量化的评定^[3]。中风病证候要素评价量表分别由内风证、内火证、痰湿证、血瘀证、气虚证

及阴虚证 6 个分量表构成, 测评方式为每个分量表单独计分^[4]。首先通过 3 轮专家问卷调查优化条目池, 然后根据每条四诊信息本身的特点及可操作性原则, 采用 2~4 级不等的分级形式进行条目备选答案设计, 再结合专家经验分别计算出各条目权重完成按权赋分, 最终形成含 97 个备选条目的《中风病证候要素评价量表备选条目表》, 其中内风证 15 个备选条目、内火证 25 个备选条目、痰湿证 16 个备选条目、血瘀证 10 个备选条目、气虚证 18 个备选条目、阴虚证 13 个备选条目。

1.2 研究对象

1.2.1 诊断标准 参照 1995 年中华医学会第四次全国脑血管病学术会议通过的《各类脑血管疾病诊断要点》^[5] 中脑梗死、脑出血诊断标准。

1.2.2 纳入标准 符合上述诊断标准; 经头颅 CT 或 MRI 确诊; 年龄 35~80 岁; 病程 ≤ 3 个月; 患者签署知情同意书。

1.2.3 排除标准 蛛网膜下腔出血; 严重失语或构音障碍影响问诊信息采集; 合并有肝、肾、造血

基金项目: 国家科技重大专项 (2009ZX09502-028)

* 通讯作者: gaoying973@126.com, (010) 84013209

系统、内分泌系统等严重疾病及骨关节病；精神障碍或严重痴呆；伴重症感染者。

1.3 一般资料

病例来源于 2010 年 9 月至 2012 年 1 月全国脑病重点专科协作组单位的 12 家临床研究中心，包括北京中医药大学东直门医院、北京中医药大学东方医院、天津中医药大学第一附属医院、长春中医药大学附属医院、浙江中医药大学附属第一医院、安徽中医学院第一附属医院、江门市五邑中医院、佛山市中医院、陕西省中医院、湖北省中医院、河南省中医院、河南中医学院第一附属医院的 410 例住院或门诊患者，其中男 245 例，女 160 例（另有 5 例患者性别信息填写不全）；年龄 86 岁 ~ 31 岁，平均（64.20 ± 10.74）岁。

1.4 调查方法

由经过统一培训的脑病科医师完成临床信息采集，填写《中风病证候要素评价量表备选条目表》。采集时点包括入组当天、入组第 4 天。

1.5 统计学方法

采用 Mplus 7.0（由 LindaMuthen 和 BengtMuthen 合作研发，下载地址：<http://www.statmodel.com/>）统计软件，选择两参数的 Logistic 项目反应模型^[8]，分别计算内风证、内火证、痰湿证、血瘀证、气虚证、阴虚证分量表中各条目的项目信息函数，以删除项目信息函数 < 0.1 的条目为标准进行条目筛选。

2 结果

表 1 示，由于内风证中仅头晕、目眩两条目的

表 1 《中风病证候要素评价量表备选条目表》各分量表条目的项目信息函数

证型	条目	项目信息函数	证型	条目	项目信息函数	证型	条目	项目信息函数
内火证	躁扰不宁	0.44374	内风证	近 48 h 内病情加重或波动	0.00028	气虚证	神疲	0.19077
	面红或目赤	0.79302		目偏不瞬	0.01740		乏力	0.27814
	口唇深红	0.33645		目珠游动	0.01409		面色 _㿅 白	0.38451
	呼吸气粗	1.42671		瞳神异常（瞳孔缩小、散大或不等）	0.00119		口张	1.04869
	谵语	1.27237		口噤	0.00818		呼吸微弱	0.80971
	心烦	0.58993		项强	0.00315		语声低微	0.30545
	口臭	0.28444		手足或下颌颤动	0.06133		气短	0.40622
	发热	0.13565		抽搐	0.02773		手或足肿胀	0.18465
	口干	0.83885		肢体僵硬状态	0.00055		肢体松懈瘫软	0.10582
	渴喜冷饮	1.03149		步履不稳（共济失调）	0.01601		肢体发凉	0.25782
	口苦	0.29066		头晕	0.96448		心悸	0.44722
	头胀痛	0.13149		目眩	9.40562		自汗	0.17293
	黄痰	0.38937		舌短缩	0.00027		大便或小便失禁	0.08499
	腹胀	0.20733		舌颤	0.01028		大便不成形	0.14879
	呃逆	0.36940		弦脉	0.00364		胖大舌或齿痕舌	0.13948
	便干便秘	0.43201	痰湿证	神情呆滞	0.53282		缓脉或弱脉	0.06385
	小便黄赤	0.76480		但欲寐	0.66314		沉脉	0.14249
	芒刺舌	0.76307		头昏	0.10429		细脉	0.04906
	舌红	0.30070		头部闷重	0.52108	阴虚证	两颧潮红	0.80397
	舌干	0.73273		肢体沉重	0.54114		心烦	0.46684
	黄苔	0.36182		痰量	0.25739		手足心热	1.22755
	燥苔	0.76769		流涎	0.71406		盗汗	0.54317
	数脉	0.50394		纳呆	0.34494		耳鸣	0.08987
	疾脉	0.49562		口黏腻	0.28857		腰膝酸软	0.06673
	滑脉	0.01060		恶心呕吐	0.01447		便干便秘	0.35613
血瘀证	面色晦暗	1.33266		大便不成形	0.13545		不寐	0.22081
	口唇色暗	1.15618		胖大舌或齿痕舌	0.07533		舌干	0.45580
	肌肤润泽状况	0.57717		厚苔	0.05643		少苔	0.24790
	头刺痛	0.22235		滑苔	0.13907		弦脉	0.00849
	痛有定处	0.11211		腻苔	0.04548		细脉	0.00800
	肌肤不仁	0.08254		滑脉	0.13681		数脉	0.33051
	舌质暗	0.13692						
	舌下脉络	0.27261						
	涩脉	0.07878						
	结脉或代脉	0.37089						

项目信息函数 > 0.1 ，故调整删除标准，将内风证项目信息函数 < 0.01 的条目删除，包括近 48 h 内病情加重或波动、瞳神异常（瞳孔缩小、散大或不等）、口噤、项强、肢体僵硬状态、舌短缩、弦脉；删除内火证中项目信息函数 < 0.1 的条目为滑脉；删除痰湿证中项目信息函数 < 0.1 的条目，包括恶心呕吐、胖大舌或齿痕舌、厚苔、腻苔；删除血瘀证中项目信息函数 < 0.1 的条目，包括肌肤不仁、涩脉；删除气虚证中项目信息函数 < 0.1 的条目，包括大便或小便失禁、缓脉或弱脉、细脉；删除阴虚证中项目信息函数 < 0.1 的条目，包括耳鸣、腰膝酸软、弦脉、细脉。

3 讨论

项目反应理论是外显变量为分类变量、潜在变量为连续变量的潜在变量模型，它采用非线性模型将受试对象对项目的反应（外显变量）与其潜在特质（潜在变量）连接起来，可以用于编制测验量表、题库建立、条目筛选、计算机适应性测验等领域^[9]。它通过项目信息函数、区分度参数（a）和难度参数（b）并结合项目特征曲线图进行量表条目的筛选，剔除信息量较低的条目^[8-11]。本研究结果变量是定性的，研究目的是希望对量表中的 97 个条目进行筛选，故可选用项目反应理论。由于参数及项目特征曲线的函数形式不同，项目反应理论有多种不同的模型，如 Normal 模型和 Logistic 模型，Logistic 模型与 Normal 模型（模型中引入了较复杂的积分运算）相比形式上更为简洁，更具数学模型的特点，并且模型的项目参数和能力参数的估计比较方便，估计精度更高。本研究中的四诊信息是由医生采集完成的，不存在猜测度的问题，或者说猜测参数（机遇参数）为 0。因此选择两参数 logistic 模型是较为合理的。

项目信息函数是用来说明各个项目对不同能力水平受试对象所能提供信息的多少，可以用来编制测验、诊断项目好坏。在两参数模型中，项目信息函数被定义为：

$$I_i(\theta) = \frac{P_i^*(\theta)^2}{P_i(\theta) * (1 - P_i(\theta))} \quad (i=1, \dots, n)$$

其中， θ 为能力水平， $I_i(\theta)$ 为项目 i 在能力 θ 上所提供的信息， $P_i(\theta)$ 为能力是 θ 的受试对象答对某一个项目的概率， $P_i^*(\theta)$ 为在 θ 点上的 $P_i(\theta)$ 值的导数^[9]。

项目区分度参数 a，即变化的斜率（ θ 取值为

b 处的切线斜率）。项目区分度用来描述项目对受试对象潜在特质的区分能力，a 值越大，代表项目特征曲线越陡，说明项目对受试对象的区分度就越高，能将不同能力水平的受试对象区分开；a 值越小，代表项目特征曲线越平坦，则不能很好地区分受试对象。理论上 a 的取值在正负无穷之间，但在实际中通常介于 0.3 和 2 之间^[12-14]。

本研究的量表是他评量表（供临床医师使用），故不需使用难度参数（b）进行条目筛选。项目信息函数是综合说明每个项目所提供信息量的客观指标，其大小将直接反映出受试能力水平的估计精度，在一定程度上受项目区分度的影响。为分别探索项目信息函数与区分度参数（a）对中风病证候要素评价量表条目筛选的影响作用，本研究先以项目信息函数值作为量表条目筛选的原则，即以删除项目信息函数值 < 0.1 的条目为标准来进行条目筛选。

本研究中内风证分量表仅有头晕、目眩两条条目的项目信息函数值 > 0.1 ，筛选结果不是很理想，主要原因考虑与纳入患者的病程及病情有关，根据中风病证候演变规律可知，内风证主要集中在发病早期或病情波动时，恢复期及病情稳定的患者内风证相对较少，对于该分量表还有待补充发病早期的患者来进一步完善。舌脉诊是中医辨证论治体系的重要组成部分及特色所在，但本研究中各分量表涉及舌脉的相关条目被删除较多，原因可能在于舌脉条目的量化临床上相对较困难，对于证候要素程度变化的判断，舌脉条目并不是特别敏感的指标。此外，项目反应理论需满足单维性假设与局部独立性假设，即组成某个测验的所有项目都是测量同一潜在的特质，并且对某个被试能力而言，项目间是相互独立无相关存在的，各证候要素分量表是否均能满足上述两个假设条件，目前尚无定论。

综合上述条目筛选结果，形成了含 76 个条目的基于项目反应理论的中风病证候要素评价量表，其中内风证 8 个条目、内火证 24 个条目、痰湿证 12 个条目、血瘀证 8 个条目、气虚证 15 个条目、阴虚证 9 个条目。基于项目反应理论，中风病证候要素评价量表条目筛选后的量表条目明显减少，测评时间也随之相应缩短，在一定程度上有利于提高临床可操作性。所形成的量表尚需通过心理测量学的方法（克朗巴赫系数法、区分度法及相关系数法等）进行信度、效度、反应度测评以及进一步的临床验证。中医证候评价量表的研究还应回归到

中医自身的规律上来,重视中医证候的特点,重视实用性与可操作性,使其源于临床并回归临床,引入前沿的循证医学临床研究方法,在应用实践中不断验证与完善。

参考文献

- [1] 余民宁. 项目反应理论 IRT 及其应用 [M]. 台北: 心理出版社, 2009: 1-50.
- [2] 郭蕾, 王永炎, 张志斌. 关于证候概念的诠释 [J]. 北京中医药大学学报, 2003, 26(2): 5-8.
- [3] 刘强, 高颖. 证候诊断量表的理论基础 [J]. 中华中医药杂志, 2010, 25(7): 989-992.
- [4] 高颖, 马斌, 刘强, 等. 缺血性中风证候要素诊断量表编制及方法学探讨 [J]. 中医杂志, 2011, 52(24): 2097-2101.
- [5] 中华神经科学会, 中华神经外科学会. 各类脑血管疾病诊断要点 [J]. 中华神经科杂志, 1996, 29(6): 379.
- [6] 陈炳为, 许碧云, 陈启光, 等. 两分类项目反应理论在中医证候中的应用 [J]. 中国卫生统计, 2011, 28(1): 16-18, 21.
- [7] 漆书青, 戴海崎. 项目反应理论及其应用研究 [M]. 南昌: 江西高校出版社, 1992: 129-215.
- [8] EMBRETSON SE, REISE SP. Item response theory for psychologists [M]. Mahwah: Lawrence Erlbaum Associates, 2000.
- [9] 高怀勇, 金桂林. 项目反应理论及其在计算机自适应测试中的应用 [J]. 西华师范大学学报(自然科学版), 2008, 29(1): 84-88.
- [10] KOSINSKI M, BAYLISS MS, BJORNER JB, et al. A six-item short-form survey for measuring headache impact: the HIT-6 [J]. Qual Life Res, 2003, 12(8): 963-974.
- [11] WEISS DJ. Adaptive testing by computer [J]. J Consult Clin Psychol, 1985, 53(6): 774-789.
- [12] 余嘉元. 项目反应理论及其应用 [M]. 南京: 江苏教育出版社, 1992: 1-157.
- [13] 杜文久. 高等项目反应理论 [M]. 重庆: 西南师范大学出版社, 2007: 52-277.
- [14] MUTHÉN LK, MUTHÉN BO. Mplus user's guide [M]. 6th Edition. Los Angeles: Muthén & Muthén, 2010: 60-61.

Items Selection of Stoke Syndrome Factor Evaluation Scale: Based on Item Response Theory

ZHANG Yang¹, GAO Ying², MA Bin², ZHONG Haizhen³, LIU Weiwei⁴, WANG Qi⁴

(1. Beijing Chaoyang Hospital, Capital Medical University, Beijing 100020; 2. Dongzhimen Hospital, Beijing University of Chinese Medicine; 3. Second Affiliated Hospital of Nanjing Medical University; 4. Biological Information Statistics Consultancy Centre, China Academy of Military Science)

ABSTRACT Objective Selecting and optimizing items of Stoke Syndrome Factor Evaluation Scale based on item response theory. **Methods** Clinical information of 410 cases of stroke was collected. Logistic item response model of two parameters was adopted to calculate item information function of each item respectively, based on the item pool formed in the previous study. Items were selected based on the standard of deleting items with item information function value smaller than 0.1. **Results** There were too little amount of items with item information function value greater than or equal to 0.1 in sub-scale of endogenous wind syndrome. Consequently, its items were deleted on the standard of item information function value smaller than 0.01. Totally 7 items were deleted, including items of aggravation or unstable condition within the current 48 h, abnormal pupil god (myosis, mydriasis or anisocoria), lockjaw, stiff neck, rigidity of the limbs, ankyloglossia, and string-like pulse; One item (slippery pulse) was deleted in sub-scale of endogenous fire syndrome according to the selecting standard; Four items (nausea and vomiting, enlarged tongue or tooth-printed tongue, thick fur, and greasy fur) were deleted in sub-scale of phlegm-dampness syndrome; Two items of numbness of skin and muscle, and uneven pulse were deleted in sub-scale of blood stasis syndrome; Three items of incontinence of urine or feces, moderate pulse, weak pulse or thready pulse were deleted in sub-scale of Qi deficiency syndrome; And four items of tinnitus, soreness and weakness of waist and knees, string-like pulse and thready pulse were deleted in sub-scale of Yin deficiency syndrome. **Conclusion** The Stroke Syndrome Factor Evaluation Scale (a preliminary scale) is formed based on item response theory. It contains 76 items, among which there are 8 items of endogenous wind syndrome, 24 items of endogenous fire syndrome, 12 items of phlegm-dampness syndrome, 8 items of blood stasis syndrome, 15 items of Qi deficiency syndrome, and 9 items of Yin deficiency syndrome.

Keywords item response theory; stoke; evaluation scale of syndrome; item selection

(收稿日期: 2015-05-02; 修回日期: 2015-12-11)

编辑: 焦爽