

中医临床研究软指标评价体系的初步构建^{*}

□朴海垠 谢雁鸣^{**} (中国中医科学院 北京 100700)

摘要:通过对中医临床研究中中医软指标的筛选、赋权、量化及用适宜统计方法进行归纳和整理,比较其各自优缺点,并分析其适宜应用范围,探索了建立符合中医软指标自身特点的、科学的中医软指标评价体系的方法。考虑到中医软指标的主观性、非线性等复杂特性,中医软指标的构建宜结合定性研究和定量研究的范式,借鉴多学科理论成果。

关键词:中医软指标 评价体系 评价方法

在中医临床研究领域中,软指标主要是指可重复性差、不能准确度量的指标。如果将临床指标分为硬指标和软指标两大类,那么硬指标包括实验室检查指标、影像学指标、生物学指标等可以通过相关仪器或检测方法精确测量得到精确数值的客观指标,而软指标则是指除此之外,通过医生的观察采集或患者的自我报告而获得的无法用仪器或其他手段直接进行精确测量的主观指标,比如疾病的症状,大多数体征,人的情绪和心理状态等。中医临床实践内容和诊疗思维的复杂性、非线性以及作为实践和研究主体和客体的人的主观性令中医软指标具有主观性、非线性、多维性、动态性、模糊性特点。中医软指标的特性决定了在研究中只有选择适合其特点的测量、评价方法,才能达到客观评价的目的,最终使中医药临床干预的效果能够得

到更为真实的反映。

所谓评价指标体系,是指为完成一定研究目的而设计的由若干个有相互联系的指标组成的指标群。指标体系的建立不仅要明确指标体系由哪些指标组成,还要确定指标间的相互关系,即指标结构。中医软指标评价体系就是为达到对中医软指标进行评价的目的而设计的一组指标群,其中包括软指标群及其数值、权重。因此中医软指标评价体系宜包括中医软指标的筛选、设计、赋权(权重)、量化(赋值)及评价。

一、中医软指标的筛选

中医软指标的筛选是建立中医软指标评价体系的第 一步,筛选的方法既包括临床流行病学调查所获得的直接证据,也包括数理统计方法和相关领域专家经验的支持。通过这些过程的不断演进和去粗取精,加之适当的检验,从而达到指标的优化。中医软指标包括各种无法

收稿日期:2006-03-13

修回日期:2006-11-28

* 国家科技部科技基础性工作专项资金课题(2001DEA20010):“中药材标准及相关中医临床疗效评价标准”之子课题—“绝经后骨质疏松症临床疗效评价标准的研究”,负责人:谢雁鸣;国家中医药管理局课题(04-05LP8888):“绝经后骨质疏松症中医临床疗效评价方法学研究”,负责人:谢雁鸣。

** 联系人:谢雁鸣,主任医师,博士生导师,研究方向:老年病科研究,Tel: 010-64014411 转 2402, E-mail: caigong4407@vip. sohu. com。

以数值精确度量的症状和体征。根据不同的研究目的,首先应在指标群组成上进行专业角度的初选,初选指标应尽量全面,可以借用统计学中的综合法和分析法来辅助进行。综合法是对已存在的指标群按一定的标准进行聚类,使之体系化;分析法是将度量对象和目标划分成若干部分、侧面(即子系统),并逐步细分,直到每一部分和侧面都可以用具体的指标来描述、实现。此时所得到的指标初选结果就基本较为详尽。

初选所得的指标群虽然全面,但难免存在重复、冗余的指标,在实际的测量和评价中,容易影响其它重要指标的评价效能,而使评价结果失真。这时需要使用一些方法来进行指标的精确。心理学经典测验理论中的项目特征曲线,可以以每个项目(题目)得分与总分之间的回归线来进行项目分析,然后结合研究目的和受试者的特征来增删题目,通过项目分析可以排除指标体系建立过程中的主观臆断,增加题目的客观性^[1]。另外,还可以通过计算“辨识度”和“冗余度”两个统计系数来分别判断指标的区分能力与重复程度^[2]。

二、中医软指标的赋权

确定了指标组成后,就需要对具体的指标进行赋权。对指标的赋权是指标在研究对象整体中的相对重要程度的量化,其大小直接影响评价的结果。合理地确定、分配指标权重是指标量化的关键。指标权重的计算方法通常可以归结为两大类:主观赋权法和客观赋权法。

1. 主观赋权法

主观赋权法根据决策者对各指标的主观重视程度赋权,常见的有专家咨询法(Delphi)、层次分析法(AHP)、相关系数法等。主观赋权法能够反映研究者的意志,但结果具有很大的主观随意性,少有严密的数学处理,故在科学性和可信度往往有限。

2. 客观赋权法

客观赋权法依据客观信息(如决策矩阵)进行赋权,有主成分分析、因子分析、熵值法等,这些方法多基于数理统计、模糊数学、灰色系统理论、运筹学、系统工程、经济学等建立模型,均有其优缺点^[3]。客观赋权法具有较强的数学理论依据,可以避免评价结果的主

观随意性,但是同时又不能体现决策者的意愿。

3. 主客观综合赋权法

近年来,针对多指标研究中指标权重的确定所存在的困难,主观赋权法和客观赋权法单独使用都存在一定的局限,不少学者以各种方法把主、客观权重结合起来得到指标的最终权重。田氏^[4]提出的秩和比法(RSR)利用反映各指标实际水平之间的分比(SR)和反映各指标在评价中的经验权重系数(W'),来确定综合权重系数。又如陶氏^[5]提出的综合权重赋值法等,都是主观赋权法和客观赋权法的综合运用的实例。

主客观综合赋权的优点是可以将研究者给出的主观权重偏好信息与客观的决策矩阵信息进行有机地结合,使确定的权重同时反映主观程度和客观程度,有利于弥补各自的局限。但从这方面的文献中也可以看出,主客观综合的赋权方法往往需要使用大量数理公式进行具体计算,其理论不易快速掌握,且操作较为繁杂,显然这是它的缺点。

目前:中医临床软指标研究中主观赋权方法的使用较为常见,且中医软指标评价体系亦属于多指标系统,也应更多尝试采用主客观综合赋权方法,以适应其多指标、多维度的评价需要。另外,由于软指标本身所具有的较强的主观色彩,在重视专家经验的同时亦不应忽视患者的感受,许氏^[6]所提出的个体化权重系统的思路亦可为中医软指标研究所借鉴。

三、中医软指标的量化

在临床研究中,中医软指标的量化目前多采用线性量化方式。这里针对中医软指标非线性特点,介绍几种可供参考的软指标量化的非线性方法。

美国医学结局研究组(Medical Outcomes Study, MOS)研制的健康状况问卷(SF-36),其中躯体疼痛条目7、8和一般健康状况条目1的答案设置了非等距的计分方式(表1、表2、表3)。也就是说根据大量人群的调查结果,症状或答案的各个等级之间的距离往往并不相等,通常的等距计分方式反映的是一种理想化的线性状态,这种仅仅从计算和赋值方便的角度来量化的方法并不具有很强的科学性。

表1 身体疼痛状况问卷调查

答案	计分
根本没有疼痛	6.0
有很轻微疼痛	5.4
有轻微疼痛	4.2
有中度疼痛	3.1
有严重疼痛	2.2
有很严重疼痛	1.0

表2 疼痛影响正常工作状况问卷调查

答案	计分
根本没有影响	6.0
有一点影响	4.75
有中度影响	3.5
有较大影响	2.25
有极大影响	1.0

表3 总体健康状况问卷调查

答案	计分
非常好	5.0
很好	4.4
好	3.4
一般	2.0
差	1.0

表4 线性症状语言简明量化表

语言评价	量化值
无症状	0
症状较轻	0.25
中等症状	0.5
症状较重	0.75
症状极重	1

注:若将无症状量化为0,症状极重量化为1,传统的线性量化方法将中等症状量化为0和1的平均值0.5,即对称中心。这是一种视各症状等级之间为等距的量化方法。

表5 非线性症状语言简明量化

语言评价	量化值
无症状	0
症状较轻	0.236
中等症状	0.618
症状较重	0.854
症状极重	1

注:M指数症状语言非线性量化方法则使用黄金分割(0.618)来定义“非对称中心”,符合人体感觉-反应的复杂的非线性关系特点。

缪兴中等^[7]以M物理模型中提取的大众化特征指数M指数,建立了症状语言描述的数字对应关系—M指数症状语言线性和非线性简明量化表(表4,表5),其量化理论简洁明了,具有很强的可操作性,其科学性和客观性值得临床实践的检验。

另外,在某些适宜的中医软指标的量化中,可以借鉴现代医学的疼痛强度评价方法中广泛使用的视觉模拟评分法(VAS)或数字评分法(NRS),其量化的结果可以直接由测量的具体数值表示。视觉模拟评分法采用一条10cm长的直线或尺,两端标明有0和10的字样。0端代表无痛,10端代表最剧烈的疼痛。让病人在直线或尺上标出自己疼痛的相应位置。然后用尺子测量出疼痛强度的数值。数字评分法以0~11或0~101点数字来代表疼痛强度,0表示无痛,疼痛增强时增加点数,10或101点表示最剧烈疼痛,由患者选择对应自身情况的数字来计分^[8]。

四、中医软指标的结构设计

目前,大多数中医临床研究设计当中,软指标的分级计分采用的多为无、轻、中、重四级,或无、轻、中、较重、极重五级标准。在这种等级量化方法中,软指标在各个维度的内容同时被包含于每个等级的表述当中,这种方法由于操作简单而易于临床研究的设计与实施。但在实际研究中随着临床干预的进行,软指标在不同维度上的动态变化有时并非呈现平行的关系,因此针对不同病证的具体特点,其中医软指标在某些维度的必要细化是很有意义的,宜根据不同维度的具体特征分别设计相应的指标。

中医软指标的维度一般包括:

1. 性质程度

症状、体征的性质特征,出现的明显性和严重性程度。

2. 持续时间

症状、体征每发生一次持续时间长短。

3. 发作频率

症状、体征发作的次数多少,根据不同疾病有不同情况。

4. 相关因素

其中又包括伴随条件、诱发因素和缓解条件。

有的症状、体征本身无明显特征,可通过伴随条件来判断病情;诱发因素可以反映症状、体征对特定刺激的反应程度;缓解条件反映症状、体征对不同程度缓解条件的反应。有些症状、体征包含的维度较多,有些则较单一,宜针对不同研究目的和方向而制定具体的指标维度。

五、中医软指标的评价方法

基于中医软指标多维性、动态性和非线性特点,中医软指标的测量需要在不同的维度多次重复进行,因此所得数据呈现多终点、多时点的特点。因此综合多终点、多时点的统计分析方法能够简化对分析结果的解释,并提高检验功效,但也要防止掩盖特定时点、特定方面的差别。

1. 横向(同一时点)资料的分析方法

由于中医软指标在横向上具有显著的多维性,因此对同一时点资料的分析应注意不同多变量分析方法的不同适应性。

(1) Hotelling T² 检验或多变量方差分析。

此法将每一个指标当成一个变量来进行分析^[9]。如果测定中条目(变量)较多,一般需要先降维,把多个变量综合为少数几个主要方面的指标后再作分析。可以用直接累加和加权累加两种方法进行指标的综合。此法较为简单而便于操作,适用于多数分析。

(2) O'Brien 的参数及非参数综合检验法。

O'Brien 法实际是将单变量的秩和检验推广到多变量情形。非参数综合检验方法,适合于多变量渐近正态分布资料。总的说来,两法均具有很高的综合性,且各具自己的特点。前者的适用性较广,但未考虑权重,检验效率也不太高;后者检验效率高,考虑到权重,但对变量分布有所要求。但当统计量值在参数和非参数界值附近时,两法结论可能不一致,宜慎用^[10]。

(3) TOPSIS 法。

此法通常不用于进行组间比较,但通过对变量的距离变换而构成综合统计量(接近度),可以达到组间检验的目的,其中在形成决策矩阵时需注意应将各组资料混合进行,而不宜分组处理^[11]。

另外,还有扩展的模糊综合评判法等。

2. 纵向(不同时点)资料的分析方法

纵向研究设计一般需要描述总体的平均增长趋势或不同个体之间增长趋势的差异。由于中医软指标在纵向上呈现动态非线性变化的趋势,而纵向测定资料不能采用一般常规统计方法处理的一个重要原因就是同一个体不同时点的测定值间存在着高度相关性(自相关)^[12]。因此在资料分析中应充分注意应用适合方法进行分析。传统的纵向分析方法包括探索性统计描述分析、重复测量方差分析、时间序列分析等方法。另外,还有一些经过发展而形成的新型方法或较为复杂但对复杂资料分析较为适合的纵向资料统计方法。

(1) 序贯主成分方差分析。

是将主成分分析与方差分析思想结合起来,首先寻找主成分,再对主成分进行序贯的比较。但此法只能对一个具体的变量进行分析,因此仅适合对疗效的某一方面得分(或总得分)进行单独的分析^[13]。

(2) Markov 过程法。

是将连续型的变量划分为离散状态后,按纯不连续 Markov 过程处理。此法不但可分析各状态的变化情况(转移概率),而且可分析任意两状态间转移的影响因素。这对深入的分析疗效的纵向变化及影响因素是非常有用的。其关键是状态的合理划分及建立伴随变量与转移强度的关系模型^[14]。

(3) 衍生变量法。

也称为响应特征分析(response feature analysis),衍生变量(derived variable)或综合测度量(summary measures)是根据研究目的及所关心的问题的侧重点不同而综合出来的原始观察值中没有的新变量^[12],之后可用常规的统计方法进行处理,给出一个总的效应大小的度量,这样克服了传统的组间比较方法的不足;有助于保证分析的侧重点,且结果易于解释;适合各时点统计量的计算中例数不一样的情况。其缺点容易导致信息的损失,而且要求每个衍生变量的值都来自同分布的总体,有时难以保证^[15]。

(4) 伪多变量方差分析与轮廓分析。

伪多变量方差分析(quasi-MANOVA)是把同一变量各时点的测定值作为不同的变量来处理,再借用

多变量方差分析来处理重复测量资料的方法。该法的缺点是只能得出处理组间比较的结果,而无法得到不同时间点以及交互作用的结果。轮廓分析是按时间顺序连接某变量各时点测定值的均值,比较不同组间的变动规律是否一致。两法都将纵向变化作为一个整体来处理,而不分析各测定点间的相互关系,从而不用考虑自相关问题,而且可借用已有的方法来处理纵向资料,但均要求满足多元正态性和协方差阵齐性^[16]。

(5) 潜变量增长曲线模型。

潜变量增长曲线模型(latent growth curve model)用潜变量来描述总体的平均增长趋势和依时间变化的情况,模型同时考虑因素的平均值和方差。其优点在于对变量之间直接的和间接的影响关系都可以进行分析,其局限性是往往要求比较大的样本容量,而且对于所有个体的评估要求测试时间间隔相同,因此仅适用于在某几个固定时间点观测得来的纵向资料^[17]。

(6) 多层线性模型。

多层线性模型(hierarchical linear model)所使用的多层分析法通过考虑测量水平和个体水平不同的差异,不仅明确表示出个体在不同测量点的变化情况,而且可以说明个体之间存在的差异。它的缺点是,首先其参数估计方法较传统方法要复杂得多,而且也不能处理变量之间间接的影响关系和复杂的观测变量与潜变量之间的关系^[18]。

3. 综合评价方法

中医软指标属于多指标复杂系统,其研究方法不宜局限于指标的横向或纵向研究,而应有机结合各个维度、各个时点的指标特征进行时、空的综合研究,以进行有效的临床评价。目前各学科领域所使用的综合评价方法有秩和比法(rank sum ratio, RSR),数据包络法(data envelopment analysis, DEA)^[19], BP神经网络(BP neural network, BPNN)^[20]。还有基于多种方法集成的综合评价方法,如模糊人工神经网络^[21],多层次灰色评价^[22]等。

六、中医软指标评价体系的初步构建

通过本研究,初步构建了中医软指标的筛选,优化,设计,赋权,量化,评价的中医软指标评价体系。中

医软指标的评价宜从中医软指标的筛选,优化,设计,赋权,量化以及相关统计方法选取的过程实现评价的相对客观化和规范化。科学的中医软指标评价,应当通过中医软指标评价的各个步骤中合理方法的使用来实现。

中医软指标的评价方法,可以多方面广泛借鉴心理学和社会学等领域的主观指标研究成果,将定量和定性研究各自的优势和特点有机结合,这些对于进行更加科学和深入的中医软指标研究具有十分重要的意义。针对软指标在不同维度和阶度的特性,及其非线性动态发展的特点,中医软指标的量化、赋权和统计方法的选取宜从软指标自身的特点出发,在今后的研究中注重尝试和探讨各种体现其主观性、非线性、多维性、动态性等复杂特性的方法。并在充分利用现代数理统计和计算机技术的发展成果的基础上,寻求切合中医临床实践需求和符合中医软指标研究特性的具有较强针对性的研究方法,以最终达到对软指标体系的科学建立和应用取得全面的把握。在中医软指标研究的每一环节,都应注重其复杂事物的重要特征,同时在中医基本理论的正确指导下,恰当地结合使用相关现代科学研究成果,方能保证中医软指标评价体系的科学建立。

参考文献

- 董奇,申继亮. 心理与教育研究方法. 杭州:浙江教育出版社,2005. 166~185,433~437.
- 苏为华. 论统计指标体系的构造方法. 统计研究,1995,2:63~65.
- 游海燕. 预测、评价指标体系构建研究现状述评. 数理医药学杂志,2005,18(3):265~267.
- 田凤调. 利用RSR确定权重系数的研究(二)——五种权数的比较. 中国卫生统计,1992,9(4):13~18.
- 陶菊春,吴建民. 综合加权评分法的综合权重确定新探. 系统工程理论与实践,2001,(8):43~48.
- 许军,陈和年,等. 自测健康评价的权重研究. 中国行为医学科学,2000,9(3):161~163.
- 缪兴中,孙同明,缪金等. M指数症状语言的线性和非线性简明量化表. 第三届国际传统医药大会,2004.
- 段红光,倪家骥. 疼痛的定性定量测量. 辽宁医学杂志,1996,10(1):1~2.
- 郑小华,石原秋. 影响乳腺癌病人生命质量的因素分析. 中国社会

- 医学,1992,5:31.
- 10 万崇华,方积乾,张玉祖,等. O'Brien 的非参数与参数综合法及其在生命质量资料分析中的应用. 中国卫生统计,1998,15(4):1~3.
 - 11 万崇华,方积乾,陈丽影,等. 同一时点生命质量资料分析的 TOPSIS 法. 中国医院统计,1998,5(3):134~136.
 - 12 Everitt RS. The analysis of repeated measures: a practical review with examples. The Statistician,1995,44(1):113.
 - 13 万崇华,陈丽影,等. 纵向生命质量资料分析的序贯主成分方差分析法. 昆明医学院学报,1999,20(3):52~56.
 - 14 万崇华,方积乾,等. 纵向生命质量资料分析的 Markov 过程法及其应用. 中华流行病学杂志,1999,20(3):162~165.
 - 15 万崇华,方积乾,等. 纵向资料分析的衍生变量法及其在生命质量资料分析中的应用. 中华医院管理杂志,2000,16(4):255~256.
 - 16 万崇华,方积乾,等. 纵向生命质量资料处理的伪多变量方差分析及轮廓分析法. 数理医药学杂志,2000,13(3):208~209.
 - 17 Duncan T E, Duncan S C, Strycker L A. An Introduction to Latent Variable Growth Curve Modeling: Concepts, Issues, and Applications. New Jersey, London: Lawrence Erlbaum Associates, 1999. 12~65.
 - 18 Bryk A S, Raudenbush S W. Hierarchical Linear Models: Applications and Data Analysis Methods. Newbury Park, CA: Sage Publication, 1992. 12~52.
 - 19 张音,徐勇男,蔡永宁. 医院——非线性复杂系统的综合评价. 中华医院管理杂志,1997,13(6):353~355.
 - 20 李建生,胡金亮,余学庆,等. 基于神经网络的中医证候量化诊断模型探索. 河南中医学院学报,2005,20(3):6~8.
 - 21 俞国燕,郑时雄,黄平. 复杂工程设计综合评价系统研究. 机械科学与技术,2001,20(1):4~9.
 - 22 胡笙煌. 主观指标评价的多层次灰色评价法. 系统工程理论与实践,1996,1:11~20.

A Method for Evaluating TCM Soft Indicators

Piao Haiyin, Xie Yanming

Clinic Basic Medicine Institute of TCM, China Academy of Chinese Medicine Sciences (CACMS), 100700 Beijing, China

This study explores to establish a scientific system for evaluating TCM soft indicators, through screening, weighting, measuring and comparing, using statistics. Taking into account the complex objective and non-linear nature of TCM, it is proposed that TCM soft indicator system shall be established, integrating both quantitative and qualitative approaches, and absorbing useful findings from other disciplines, such sociology and psychology.

Keywords: soft indicator, evaluation system, methodology

(责任编辑:张述庆,责任编辑:李澎涛,责任译审:邹春申)

香港中大有突破性发现: 植物提炼物质抗流感病毒

香港中文大学医学院最新研究有突破性发现:由两种植物提炼出来的物质 1-21-1-1 及 PYF1M2,每微克份量能有效抑制 5 成的人类流感和禽流感病毒在细胞内复制,抑制病毒的繁殖。香港中大计划与美国有关机构合作,研究制成抗流感药物的可行性。

香港中文大学医学院微生物学教授陈基湘 16 日指出,流感严重威胁人类健康,但现时抗流感种类不多,加上抗药性的问题,令到治疗流感的药物选择有限。

他表示,中大科研人员努力研制新的抗流感及禽流感病毒药物,发现由两种植物提炼出来的物质 1-21-1-1 及 PYF1M2,每微克份量能有效抑制 5 成的人类流感和禽流感病毒在细胞内复制,两种物质在抑制人类流感和禽流感病毒方面,其效用与“特敏福”类似。

香港中大医学院何鸿燊防治传染病研究中心主管沈祖尧指出,物质 1-21-1-1 及 PYF1M2 何时能制成疫苗推出市面应用,目前仍言之尚早。陈基湘也表示,目前该两种物质的研究只到细胞层面,预计在 2 至 3 年后才进入动物及人类临床测试阶段,中大会与美国有关机构合作,研究制出新的抗流感药物的可能性。(文摘)