

基于专家调查法的针刺治疗缺血性中风恢复期疗效评价体系研究^{*}

王升强¹, 梁伟雄^{2△}

(1. 广东省东莞市太平人民医院, 广东东莞 523905; 2. 广州中医药大学 DME 中心, 广东广州 510405)

摘要: 临床试验中, 疗效指标的选取直接决定了试验的质量。一套完整的疗效评价体系能够较为全面客观的反映研究的各个方面, 从而能够为循证医学提供高质量的依据。目前我国中医药临床试验领域疗效评价体系的建立还处于初步阶段, 本研究采用专家调查法, 在文献分析的基础上, 建立了针刺治疗缺血性中风恢复期临床疗效评价体系, 以期后续研究提供一定的借鉴和指导。

关键词: 专家调查法; 针刺; 缺血性中风; 疗效评价体系; 协调系数

中图分类号: R246.6 **文献标志码:** A **文章编号:** 1000-2723(2014)03-0042-05

中医药领域一直缺少高质量的临床试验研究, 除了研究设计不合理外, 疗效评价尤为关键^[1]。恰当的疗效评定指标是临床试验的重要环节, 采用不同的疗效指标甚至可以得出有效或无效的截然相反的结论^[2]。根据国家中医药管理局的《中医临床研究发展提纲》(1999-2015)中有关论述, 建立系统的中医药防治疾病的诊治规范和疗效评价体系已经被列入计划, 尤其是疗效评价体系备受重视。国内很多学者认为中医药疗效评价应该结合自身特点, 建立包括病、证、生存质量在内的多维度多属性疗效评价体系^[3]。

在疗效评价体系建立中, 具体评价指标的筛选至关重要。由于疾病的复杂性和中医药治疗的多效应, 我国中医药临床试验中疗效评价指标繁多, 差异较大, 研究者在选取疗效评价指标时比较随意。笔者既往的一项研究发现: 在 83 篇国内发表的针刺治疗缺血性中风随机对照研究文献中, 采用的疗效评价指标高达 34 种^[4]。基于上述原因, 本研究拟在文献分析的基础上, 采用专家调查法, 建立针刺治疗缺血性中风恢复期疗效评价体系, 以期借鉴和参考。

1 文献分析

本次研究选取 1999-2012 年间发表在 CBM、

CMCC、CNKI 以及国内中医药领域较权威的中国中医药科技文献数据库(TCMLRS), 检索词: 针刺(包括电针、头皮针), 中风或脑卒中。对所有题录进行合并, 共搜索文献 937 篇, 在剔除重复部分及并发症、中风先兆、综述、经验总结、理论探讨等部分后, 得到随机对照临床试验共有 164 篇, 经比较发现有 7 篇为重复发表文献, 同时排除明显有不合理随机方法(如按照入院或者门诊顺序随机分组)文献 10 篇, 排除设立了不合理对照措施(如未经证实的阳性对照)文献 18 篇, 最后实际选择 129 篇。所有入选的文献经过查阅, 将其疗效评价指标按照病、中医证候以及生存质量 3 个维度归类。

2 专家咨询

2.1 研究小组的建立

研究之前建立了组织小组, 主要用于专家的选定、专家咨询问卷的编制以及问卷的回收、整理。

2.2 专家的选择

根据相关文献^[5], 专家一般以 15~50 人为宜, 本研究选择具有丰富的针刺、康复、神经内科临床工作经验或从事中医药疗效评价及卫生统计领域工作, 并具有中级以上业务职称的国内专家学者, 此次咨询的专家总数共 15 人。

^{*} 基金项目: 国家“十一五”计划科技重大专项(2008ZX09312-021)

收稿日期: 2014-03-12

作者简介: 王升强(1979-), 男, 安徽怀宁人, 医学博士, 主治医师, 主要从事中西医结合临床康复及中医药疗效评价研究。

△通信作者: 梁伟雄, E-mail: liangwx@yeah.net

2.3 备选指标的选择和问卷的制定

本研究共进行两轮专家咨询。课题小组在文献分析和专家会议(头脑风暴)的基础上设计了第一轮咨询问卷,在问卷中首先向专家介绍了针刺治疗缺血性中风恢复期疗效评价现状和专家调查法的背景和要求。根据目前中西医疗效评价体系的研究现状,在病、证、生存质量3维疗效评价的基础上,结合WHO对于疾病健康状态病理、病损、残损、残障4个水平的定义,将针刺干预缺血性中风恢复期临床疗效评价体系分成神经功能、生存质量、终点指标、中医证候及生物理化指标等5个维度(一级指标),并在此基础上构建了评价指标体系的备选指标池。

问卷内容包括各位专家对针刺干预缺血性中风恢复期疗效评价各指标的重要性判断,专家对指标的熟悉程度和判断依据以及专家的基本情况(如年龄、职务职称、专业以及从事工作年限等内容)。此外,对专家问卷中设置了开放性的问答,以便专家对课题和备选指标提出自己开放的意见和建议。

第二轮咨询问卷将第一轮统计结果以最大值、最小值、中位数、4分位数的形式附在第二轮问卷上,同时将各个专家第一轮问卷的评估结果一并附上作为参考,还将第一轮统计结果以及各项统计指标的意义向专家作一说明,并征询每位专家在看完第一轮的结果后是否改变自己的预测。

2.4 专家赋值依据

在调查表中,指标重要程度依据李克5分量表法(5-Likert scale):很重要(5分),重要(4分),一般重要(3分),不太重要(2分),不重要(1分)^[6],考虑到每位专家对不同问题的权威程度不同,需了解专家对每个指标的熟悉程度和判断依据,熟悉程度分为5个等级:不熟悉、不太熟悉、一般、较熟悉、很熟悉,熟悉程度系数分别是0.2,0.4,0.6,0.8,1.0。判断依据按常规分为理论分析、实践经验、同行了解、直观选择4类,影响程度分别为大、中、小,分别赋予不同的量化值。见表1。

表1 咨询表中判断依据影响程度的赋值

判断依据赋值	大	中	小
理论分析	0.3	0.2	0.1
实践经验	0.5	0.4	0.3
国内外同行的了解	0.1	0.1	0.1
直观了解	0.1	0.1	0.1

2.5 指标筛选标准

指标筛选以同时满足重要性赋值均数 ≥ 3.5 且变异系数 <0.20 两项指标为标准,同时结合专家意见,经过课题组集体评议后进行指标筛选结果的确认。

2.6 统计分析

2.6.1 统计指标

首先对专家的性别、年龄、职务或职称、从事专业的年限等个人特征进行描述性分析,以了解专家的基本情况。另外的统计指标有:

(1)专家积极系数:专家对本项研究关心合作的程度,专家积极系数越高,说明专家对本项研究关心合作的程度越高,本项研究结果的可信度越高。采用问卷回收率表示。

(2)专家权威程度^[7]:用专家权威系数(Cr)来表示,专家权威系数一般由两个因素决定,一个是专家对问题作出判断的依据,用Ca表示,一个是专家对指标的熟悉程度系数,用Cs表示。权威程度=(判断系数+熟悉程度)/2,即 $Cr=(Ca+Cs)/2$ 。专家的权威程度以自我评价为主。专家的权威程度与预测精度呈一定的函数关系,一般来说,预测精度随着专家权威程度的提高而提高。

(3)专家意见集中程度:专家对各指标的相对重要性意见集中程度,采用指标的重要性赋值均数表示。

(4)专家意见的协调程度:用以说明参与专家对每项指标的评价是否存在较大分歧,采用变异系数和专家协调系数表示。

2.6.2 统计方法

所有数据分析在SPSS13.0软件平台上进行,定性资料用卡方检验或者秩和检验,定量资料符合正态分布用t检验,不符合正态分布用秩和检验。所有显著性检验均采用双侧检验的P值, $P<0.05$ 被认为有显著性意义。

3 研究结果

3.1 专家基本情况

3.1.1 一般资料

本研究共邀请专家15名,专家的选择领域较广,包括针灸科、中西医康复科、神经内科临床医生,中医药疗效评价研究领域及卫生统计学专家,平均年龄为 (46.29 ± 6.38) 岁,平均从事专业年限 (23.45 ± 7.61) 年。见表2。

表 2 咨询专家一般资料

项目	人数	百分比/%
年龄(岁)	<40	3
	40~	6
	50~	6
从事专业年限(年)	10~	6
	20~	7
	≥30	2
		13.3
职称	正高	8
	副高	5
	中级	2
学历	博士	9
	硕士	4
	本科	2
		13.3

3.1.2 地域分布

地域分布相对较广, 具有较好的地域代表性(见表 3)。

表 3 咨询专家地域分布情况

地区	人数	构成比/%
天津	3	20
安徽	5	33.3
广东	6	40
香港	1	6.7

3.1.3 专家研究领域分布

咨询专家组中 15 位专家来自针灸学、康复医学、神经内科、临床流行病学及卫生统计学 5 个不同研究领域。见表 4。

表 4 咨询专家研究领域分布情况

研究领域	人数	构成比/%
针灸学	5	33.3
康复医学	5	33.3
神经内科	2	13.3
临床流行病学	2	13.3
卫生统计学	1	6.7

3.2 问卷结果

3.2.1 专家积极系数和权威程度

两轮专家咨询均收回问卷 11 份, 全部符合填写要求, 问卷回收率 73.3%, 专家积极系数为

73.3%。

专家的权威程度由专家的权威系数决定, 权威系数为判断系数和熟悉程度系数的算术均值。专家对指标的判断系数 0.86, 专家对指标的熟悉程度为 0.77, 而专家的权威系数为 0.82, 显示专家的权威程度较高。

3.2.2 专家意见的集中程度及变异系数

3.2.2.1 一级指标(见表 5)

表 5 各一级指标(维度)集中程度及变异系数

指标	集中程度(得分均数)		变异系数	
	第一轮	第二轮	第一轮	第二轮
神经功能	4.10±0.74	4.46±0.52	0.18	0.13
生存质量	4.50±0.52	4.73±0.47	0.11	0.10
终点指标	4.30±0.79	4.55±0.52	0.18	0.11
中医证候	3.70±0.68	4.18±0.41	0.18	0.10
理化指标	2.80±0.63	3.09±0.70	0.23	0.23

3.2.2.2 二级指标(见表 6~9)

表 6 神经功能维度

指标	集中程度(得分均数)		变异系数	
	第一轮	第二轮	第一轮	第二轮
残损水平	4.10±0.74	4.36±0.51	0.18	0.13
残疾水平	4.50±0.53	4.73±0.47	0.12	0.10

表 7 终点指标维度

指标	集中程度(得分均数)		变异系数	
	第一轮	第二轮	第一轮	第二轮
病死率	4.10±0.74		0.18	
致残率	4.60±0.52	4.73±0.47 [▲]	0.11	0.10 [▲]
复发率	4.20±0.63	4.36±0.51	0.13	0.12
不良反应发生率	4.00±0.67	4.10±0.54	0.17	0.13

注: ▲为病死率+致残率

表 8 中医证候维度

指标	集中程度(得分均数)		变异系数	
	第一轮	第二轮	第一轮	第二轮
患者自评量表	4.40±0.52	4.46±0.52	0.12	0.12
患者他评量表	4.00±0.67	4.18±0.41	0.17	0.10

表 9 理化指标维度

指标	集中程度(得分均数)		变异系数	
	第一轮	第二轮	第一轮	第二轮
TCD	2.80±0.63	2.91±0.70	0.23	0.24
梗死面积变化	2.20±0.63	2.64±0.51	0.29	0.19
其他生化指标	2.40±0.52	2.55±0.52	0.22	0.20

3.2.2.3 专家意见的协调系数

采用 Kendall 协调系数 W 检验(Kendall's coefficient of concordance W)的分析方法^[8]。协调系数 W 反映了全部专家对所有指标的评价意见的一致性,取值范围在 0~1 之间, W 越大,表示专家的意见越统一,协调程度越高。经协调系数得出卡方值(Chi-Square),然后计算 P 值,根据 P 值大小判断数据之间是否具有 consistency。

经统计,第一轮对所有指标的评价意见协调系数 W 为 0.40,卡方值 44.00, $P<0.01$,第二轮协调系数 W 为 0.73,卡方值 79.92, $P<0.01$,显示专家对各指标的评分具有一致性。

3.2.3 结果分析

3.2.3.1 第一轮专家咨询

Delphi 法专家的选择是咨询成败的关键^[8]。本研究选择的专家涉及多个相关领域,这些保证了专家咨询结果的可靠性。

第一轮问卷调查中,从最能反映指标重要性的集中系数来看,神经功能、生存质量、终点指标、中医证候等一级指标均显示了较好的优势,理化指标集中系数 2.80 ± 0.63 ,提示专家普遍认为在针刺干预缺血性中风恢复期临床疗效评价中没有必要采用理化指标。在一级指标的变异系数中,理化指标的变异系数为 0.23,显示专家在对其重要性评估时波动较大,其他一级指标均显示了较理想的变异系数,提示专家的协调程度较高。在二级指标方面,除了理化指标下属的血流变、梗死面积变化、生化指标等没有显示理想的集中系数外,其他指标集中系数、变异系数均较为满意。

3.2.3.2 第二轮专家咨询

第一轮咨询很多专家提出了意见和建议,有一部分专家认为,按照国际惯例,可以将终点指标维度病死率以及残疾率合并,便于统计分析;有的专家认为,脑功能成像在理化指标中作用应该引起重视。综合以上结果,理化指标这一级指标在集中系数及变异系数上没有达到满意的结果,经过课题小组商议,将在第二轮咨询中予以说明;同时在第二轮专家咨询表中合并二级指标死亡率以及残疾率。

在反映重要性程度的集中程度方面,与第一轮一样,理化指标的重要性仍然不能令人满意,未能达到筛选标准。同时,在其下属的各二级指标集中系数也不能达到筛选标准,提示参与咨询的专家普

遍认为理化指标在评价针刺干预缺血性中风恢复期疗效时不重要,就脑功能成像在针刺临床中的运用问题,大多数专家认为目前的条件该指标还不能普及,但可以作为研究针刺疗效机制时采用。

在反映专家协调程度方面,除理化指标及其下属三级指标变异系数相对较高以外,其他一、二级指标均令人满意。第一轮协调系数 0.40,第二轮上升为 0.73,两轮经显著性检验 $P<0.01$,提示专家对指标评估的一致性较好,咨询结果可取。

3.3 结论

经过两轮咨询,并经课题小组商议,理化指标不符合筛选标准,将其及其下属二级指标予以剔除,初步确定了神经功能、生存质量、终点指标以及中医证候 4 个一级指标(维度),神经功能包括残损水平、残疾水平 2 个二级指标,终点指标包括病死率+残疾率、不良事件发生率、复发率 3 个二级指标,中医证候包括患者自评量表及患者他评量表两个二级指标(如图 1)。

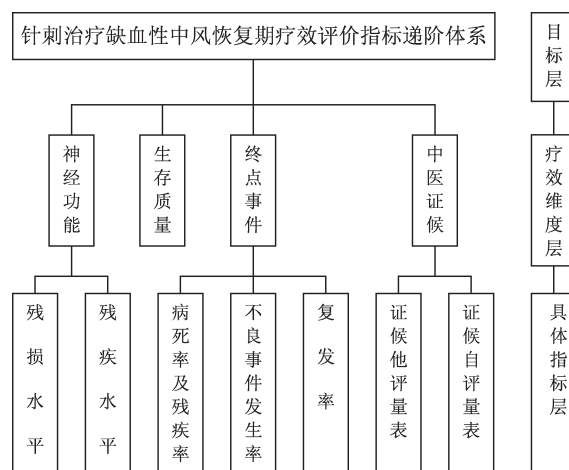


图1 针刺干预缺血性中风恢复期疗效评价指标体系

4 讨论

近年来,国内外学者出于对针刺科学性的疑虑,将研究的焦点集中在针刺的有效性上,一些大样本、多中心的临床试验研究显示,针刺治疗中风、偏头痛、纤维肌痛疗效不明显^[9]。目前已经完成的 Cochrane 系统评价中没有一个有足够的说服力的结论表明针刺有效^[10],这可能与缺乏合适的疗效评价体系有关,现有的体系难以彰显中医药治疗的优势。针刺是一项复杂干预,除了其独特的多途径、多靶点生物学干预以外,患者的身体机能和心理状态等个体化因素也能影响其疗效,如何建立一套科学

合理的评价体系全面客观的评价其疗效是摆在我们面前的一项重要任务。基于此,国内不断有方法学专家提出中医药领域应该开展综合评价研究,并提出了一系列建设性的理论和方法^[11-12],目前国内已经开展和正在开展的几项临床研究均显示了中医药领域采用综合评价体系的可行性^[13-14],本研究在文献分析的基础上,通过对不同领域权威专家的咨询,得出的针刺治疗缺血性中风恢复期疗效评价体系也一定程度的体现了其合理性和权威性,期望能为后续临床研究提供借鉴和依据。

然而我们也要清醒的看到,任何一种方法都有其缺陷,本研究采用的专家调查法需要时间较长,容易使得专家产生“疲劳效应”而可能引起偏移^[15];由于无法预测意外事件,在临床运用时候,对于病情较急、预后较差的患者就存在缺陷;在具体实施过程中,无论是研究者和专家,都或多或少存在主观效应。因此在实际操作中,我们应该根据具体干预手段和治疗病种,合理选择综合评价方法,尽可能避免或者减少偏移,使其最大限度体现中医药干预的优势,为中医药临床及循证提供依据。

参考文献:

- [1] 王升强,谢仁明,梁伟雄. 针刺临床试验若干关键问题探讨[J]. 中国针灸,2011,31(5):451-454.
- [2] Jay Henson. The use of surrogate endpoints clinical trials[J]. Statistics in Medicine. 1989(8):403.
- [3] 赖世隆. 中医药临床疗效若干关键环节的思考[J]. 广州中医药大学学报,2002,19(4):245-250.
- [4] 王升强,贾琪,梁伟雄. 近十年针刺干预缺血性中风临床研究文献质量分析 [J]. 辽宁中医杂志,2010,37(12):2317-2319.
- [5] 曾光. 现代流行病学方法与应用[M]. 北京:北京医科大学中国协和医科大学联合出版社,1996:250-270.
- [6] 柯惠新,沈浩. 调查研究中的统计分析方法[M]. 北京:北京传媒大学出版社,2005:324.
- [7] 季新强,刘志民. Delphi 法及其在医学研究和决策中的应用[J]. 中国药物依赖性杂志,2006,15(6):422-426.
- [8] 程琮,刘一志,王如德. Kendall 协调系数 W 检验及其 SPSS 实现[J]. 泰山医学院学报,2010,31(7):487-490.
- [9] 石广霞,汪德瑾,刘存志,等. 针刺效应因素的分析[J]. 中医杂志,2009,50(10):942-944.
- [10] 李瑛,梁繁荣,余曙光. 临证医学与针灸临床疗效评价[J]. 中国针灸,2003,23(2):113-115.
- [11] 高凡珠,谢雁鸣,王永炎. 中医复杂干预与疗效综合评价[J]. 中国中医基础医学杂志,2010,16(6):527-529.
- [12] 谢仁明,王永炎,谢雁鸣,等. 中医临床疗效综合评价[J]. 福建中医药,2007,38(2):1-3.
- [13] 王阶,何庆勇,熊兴江. 冠心病心绞痛中医疗效综合评价体系的研究方法探讨 [J]. 世界科学技术-中医现代化,2009,11(11):11-14.
- [14] 陈磊,梁伟雄,吕志平. 生脉胶囊治疗慢性充血性心力衰竭临床疗效的多属性综合评价 [J]. 中华中医药杂志,2010,12(12):2336-2338.
- [15] Jane S, Catherine O'H, Patrick H, et al. Identifying appropriate tasks for the preregistration year: modified Delphi technique [J]. BMJ, 1999, 319: 224-226.

(编辑:岳胜难)

The Study on the Clinical Evaluations System on Ischemic Stroke Convalescence Treated by Acupuncture Based on Delphi Method

WANG Sheng-qiang¹, LIANG Wei-xiong²

(1. Taiping People Hospital of Dongguan City, Dongguan 523905, China;

2. The Centre of DME, Guangzhou University of Chinese Medicine, Guangzhou 510405, China)

ABSTRACT: The quality of clinical trials a direct result of how to select therapeutic effect index, a complete set of clinical evaluation system fully reflects all aspects of the trial, accordingly offer high quality basis for the evidence-based medicine. at present, we are in the primary stage in establishing clinical evaluation system of clinical trials about TCM, this study used Delphi method to establish clinical evaluation system of ischemic stroke convalescence treated by acupuncture based on document analysis, look forward to provide guidance for follow-up study.

KEY WORDS: Delphi method; acupuncture; ischemic stroke; evaluation system; cooperation index.