



高等中医院校研究教程

中医临床研究设计 与SAS编程统计分析

胡立胜 主编
周 强

學苑出版社



新世纪中国医药研究丛书

中医临床研究设计与 SAS编程统计分析

张洪海 主编
王 磊 副编

人民卫生出版社

中医临床研究设计 与 **SAS** 编程统计分析

胡立胜 主编
周 强

学苑出版社

内容简介

本书是一本把临床流行病学和循证医学的原理应用到中医临床研究设计和统计分析的方法学专著。共有 25 章,包括中医证候或体质的诊断、治疗、预后和病因研究方法。既考虑到中医学传统的特色,又考虑到其与西医学的结合。供中医及中西医结合的研究生、年轻医生及临床科研工作者参考。

主编简介

胡立胜,男,1944 年 1 月 15 日生于江苏。1970 年 7 月毕业于中国协和医科大学 8 年学制医学系医学专业(北京);1978 年 5 月起任主治医师;1979 年 - 1982 年就读中国医学科学院寄生虫病研究所流行病学专业硕士研究生,获医学硕士学位(上海);1983 年 - 1986 年就读中国协和医科大学基础医学院流行病学专业博士研究生,获医学博士学位(北京);1989 年 9 月 - 1991 年 11 月在美国加利福尼亚大学(旧金山分校)医学院内科学系做博士后研究学者。1998 年 1 月至今,任北京中医药大学中医科研设计与循证医学研究中心主任、教授、博士研究生导师、执业医师(北京)。从事教学、科研和临床 30 余年。

图书在版编目(CIP)数据

中医临床研究设计与 SAS 编程统计分析/胡立胜、周强主编. - 北京:学苑出版社,2004.5

ISBN 7 - 80060 - 080 - 7

I. 中… II. ①胡…②周… III. 中医临床 - 计算机辅助设计 IV. R53

中国版本图书馆 CIP 数据核字(2004)第 036798 号

出版发行:学苑出版社

社 址:北京市丰台区南方庄 2 号院 1 号楼 100078

网 址:www. book001. com

电子信箱: xueyuan@ public. bta. net. cn

销售电话: 010 - 67675512、67602949、67678944

印 刷 厂: 河北省高碑店市鑫宏源印刷厂

开本印张: 850 × 1168 16 开本 18.25 印张

字 数: 380 千字

版 次: 2004 年 5 月北京第 1 版

印 次: 2004 年 5 月北京第 1 次印刷

印 数: 0001--1000 册

定 价: 30.00 元

前 言

中医临床医学是一门经验性很强的科学。长期以来，临床研究方法主要停留在传统经典文献的体会和诠释、临床观察和医案报告、师徒的口传心授和个人经验积累等。这些方法过去和现在都对中医临床医学的发展发挥了积极作用，而且迄今一直占据主导地位，成为研究方法的主要方面。但是，上述研究方法毕竟有它的局限性和盲目性，由此得出的许多结论常常带有片面性，有些甚至是错误和荒谬的。（参考：胡镜清，赖世隆. 中医临床研究方法指南——概论. 国家中医药管理局科技教育司，1999，1页）。

如何吸取、应用现代科技（包括西医学）的理论、方法和手段促进中医学的发展，这是极富有意义的挑战性课题。要达到此目的，需要方法学的突破及思路创新。

临床流行病学（设计、测量、评价）、循证医学、相应的医学统计方法和计算机技术的快速兴起、发展、普及，使临床研究的方法学在 1980-90 年代进入临床流行病学（DME）时代，从 90 年代末又发展到循证医学（EBM）时代。临床流行病学或循证医学，正是临床研究方法学的突破及思路创新，尤其是随机双盲对照临床试验，是堪与显微镜的发明媲美的里程碑式的贡献，必将成为未来 50 年临床医学最瞩目的成就。

虽然中医学的基础理论（对生命规律的认识）、临床思维和实践特点都与西医学有显著差异，但科学是有共性的，作者力图找出中、西医学的交叉点，将 DME 或 EBM 的原理、方法渗透进入中医学，整合到一起，这就是本书的任务和特点。

胡立胜

北京中医药大学中医科研设计与循证医学研究中心主任

周 强

卫生部中日友好医院

2003 年 2 月 5 日于北京

目 录

第一章 制订中医证候宏观辨证量表:统一辨证标准	(1)
第一节 现时中医证候宏观标准产生的背景	(1)
第二节 中医证候宏观辨证量表的制订方法	(1)
第三节 制订方法举例	(2)
第二章 中医证候宏观辨证量表的表面效度评价:Kappa 值法	(5)
第一节 定量观察(多分类变量)的一致性评价方法	(5)
第二节 定性观察(二分类变量)的一致性评价方法	(7)
第三章 中医证候宏观辨证量表的重测	
信度评价(R 值法)及反应度评价	(8)
第一节 R 值法的设计和 SAS 统计分析	(8)
第二节 量表反应度的评价	(10)
第四章 病证结合的观察性描述性研究:横断面调查	(12)
第一节 横断面调查的概念和特点	(12)
第二节 横断面调查的目的、用途	(12)
第三节 人群抽样调查设计及随机抽样的 SAS 程序	(13)
第四节 横断面调查需要的样本含量估计方法	(17)
第五节 抽样调查容易产生的主要偏倚及其控制措施	(20)
第六节 横断面调查资料的 SAS 程序统计分析(一):描述性统计	(20)
第七节 证候的“三间”(时间、空间、人间即人群特征)分布	(26)
第五章 证候与微观指标之间的相关和回归分析	(28)
第一节 直线相关研究	(28)
第二节 直线回归研究	(29)
第三节 大样本资料的直线回归研究	(31)
第四节 非线性回归研究	(33)

第六章 为中医证候筛选微观指标的多元线性逐步回归分析	(36)
第一节 用途	(36)
第二节 应用条件和资料类型	(36)
第三节 例题	(36)
第四节 SAS 程序	(37)
第五节 运行结果摘要	(38)
第六节 多重回归的决定系数	(39)
第七节 多重回归分析样本含量的估计	(39)
第七章 同病多种证候与多项微观指标之间典则相关研究	(41)
第一节 概念和主要用途	(41)
第二节 例题	(41)
第三节 SAS 程序	(42)
第四节 运行结果	(42)
第八章 某病多种证候与多项基因指标之间的对应分析:证候的基因本质研究	(45)
第一节 对应分析的概念和工作假设	(45)
第二节 设计类型	(45)
第三节 SAS 程序和运行结果	(46)
第九章 中医体质诊断研究方法概述	(49)
第一节 体质假设与验证	(49)
第二节 未知体质分类标准条件下的探索性研究设计类型	(50)
第三节 已知体质分类标准条件下的研究设计类型	(50)
第十章 中医体质与分类指标之间的数量依存关系:Logistic 回归分析 ...	(52)
第一节 适用条件	(52)
第二节 例题和 SAS 程序	(52)
第三节 运行结果的摘要及解释	(54)
第十一章 体质或证候的指数计量诊断	(56)
第一节 适用条件	(56)
第二节 几个概率	(56)

第三节	例题	(57)
第四节	利用指数之和进行计量诊断及鉴别诊断	(59)
第十二章	体质或证候的判别分析	(61)
第一节	概念、用途和应用条件	(61)
第二节	两个指标的两类判别	(61)
第三节	多指标的多类判别	(64)
第十三章	体质或证候的指标聚类分析	(68)
第一节	指标聚类分析的适用条件	(68)
第二节	指标聚类分析的原理	(68)
第三节	例题	(68)
第四节	数据集	(70)
第五节	SAS 程序	(71)
第六节	输出结果的摘要及解释	(75)
第十四章	建立体质(或证候)分类标准的因子分析	(80)
第一节	概念和用途	(80)
第二节	系统默认值大于 1, 提取公因子	(80)
第三节	指定特征根值大于 2, 以提取 7 个公因子	(89)
第十五章	诊断研究及其循证医学评价	(98)
前言	(98)
第一节	诊断性试验研究设计	(98)
第二节	评价诊断性试验的指标	(100)
第三节	诊断性试验指标的含义	(101)
第四节	怎样建立参考值(正常值)	(103)
第五节	诊断性试验的循证医学评价原则	(106)
第十六章	病或证候预后研究:生存或死亡的 COX 比例 风险回归模型分析	(108)
第一节	COX 比例风险模型的用途、条件和模型结构	(108)
第二节	设计类型和创建数据文件	(109)
第三节	SAS 程序	(111)
第四节	主要运行结果	(113)

第五节	计算预后指数	(114)
第十七章	中医临床试验的原理	(116)
第一节	医学实验与直观经验	(116)
第二节	群体概念和群体特征	(117)
第三节	什么是临床试验	(118)
第四节	临床试验的意义	(119)
第五节	临床试验三大步骤:设计、测量、评价	(120)
第六节	临床试验的主要类型及其优缺点	(122)
第七节	临床试验四大原则	(124)
第八节	失访和不依从的处理	(127)
第九节	影响中医临床试验研究质量的因素及控制	(128)
第十八章	中医药随机对照临床试验设计	(141)
第一节	中医药临床研究的特点	(141)
第二节	中医药临床研究设计的难点	(142)
第三节	中医药治疗性和预防性临床试验(随机对照试验 RCT) 研究的主要环节、设计模式和设计原则	(143)
第四节	单中心、分层、区组随机分组 SAS 程序设计	(152)
第五节	多中心、不分层、区组随机分组的 SAS 程序设计	(153)
第六节	多中心、分层、区组随机分组的 SAS 程序设计	(157)
第十九章	中医临床试验的统计学基础与统计方法的正确选择	(160)
第一节	医学统计学基本概念	(160)
第二节	样本含量的估计方法	(162)
第三节	假设检验	(165)
第四节	统计学应用贯穿于临床试验研究的整个过程	(166)
第二十章	中医临床试验的 SAS 编程统计分析	(169)
第一节	两组可比性(均衡)分析	(169)
第二节	总疗效分析	(177)
第三节	症状缓解分析	(185)
第四节	检测数据(计量资料)分析	(192)
第五节	临床等效性或非劣效性检验	(202)
第六节	剂量-效应关系分析:两步假设检验	(204)

第七节 中医药临床试验不良反应因果关系的判断与统计分析	(205)
第八节 临床试验研究质量的评价	(210)
第二十一章 中医临床试验统计分析计划与报告	(216)
统计分析计划	(216)
一、题目	(216)
二、目的	(216)
三、随机分组	(216)
四、有效性及安全性评价	(216)
五、试验数据的统计分析	(217)
统计分析报告	(219)
一、病例分组情况报告	(219)
二、可比性分析报告	(220)
三、临床疗效分析报告	(222)
四、安全性分析报告	(228)
第二十二章 中医临床试验病例报告表	(232)
第二十三章 中医成组序贯临床试验设计与分析	(247)
第一节 成组序贯试验的设计要求	(247)
第二节 计数资料的成组序贯分析方法	(248)
第三节 计量资料的成组序贯分析方法	(249)
第二十四章 中医与循证医学	(252)
第一节 什么是循证医学?	(252)
第二节 为何突然关注循证医学?	(252)
第三节 怎样实践循证医学?	(253)
第二十五章 中医文献的循证医学荟萃分析(Meta分析)	(259)
第一节 Meta分析概述	(259)
第二节 计数资料(OR值)Meta分析的统计分析方法	(260)
第三节 计量资料(均数之差)Meta分析的统计分析方法	(265)
第四节 循证医学系统评价专用软件—RevMan软件应用例题： 中药治疗慢性前列腺炎随机对照试验的系统评价	(268)

第一章 制订中医证候宏观辨证量表：统一辨证标准

第一节 现时中医证候宏观标准产生的背景

- 1.来源于古代医家的论述、医案的记载；
- 2.现代教科书中的描述；
- 3.学术机构所制定；
- 4.政府部门组织专家所制定；
- 5.来源于一定范围的临床流行病学研究。

现就第2、3、4项与第5项结合，制定中医证候宏观辨证量表的方法，介绍如下。

第二节 中医证候宏观辨证量表制订方法

制订方法的选择：

- 1.利用现有的四诊项目组合法 + 临床流行病学调查证候加权积分法：

(1) 国标或行标规定的主要症状都具备（如无主要症状，则要求任意3项症状）+ 舌苔 + 脉象。合计至少5项阳性才能辨为某证候。根据此项规定可计算出每个证候诊断的最低加权积分临界值。优点：宏观辨证统一化、规范化、量化。

(2) 现场调查证候加权积分步骤：

① 每项症状按严重度或频次制定0分、2分、4分、6分的得分量表：0分 = 无某项症状，2分 = 轻症状，4分 = 中症状，6分 = 重症状；都要用文字描写各分值的具体内容。这是症状和体征（四诊）的定量方法。

② 舌、脉象如果不易打分，也可按无 = 0分，有 = 6分；

③ 证候加权积分（属计量资料）= 每项四诊得分* 每项四诊对证候的权重。（权重 = 同一证候中每项四诊出现频率的百分构成比，可根据文献或现场调查计算）。

④ 证候加权积分作为证候疗效评价的定量依据。

⑤ 如果再以证候加权积分为基础转化为证候分级值，则可以把证候分级为0, 1, 2, 3级，分别代表某证候的无、轻、中、重。

2.利用古、今传统文献统计法：因为四诊资料和证候在原始文献（医案）中无定量值，只能读到有或无的记载。可以统计各种证候之每项四诊资料发生的频次，换算为频率（条件

概率),再用对数转换为诊断概率指数,制出指数计量诊断与鉴别诊断表。

3.只按照现代临床流行病学调查和统计分析:

特点:不受现有的证候诊断(宏观辨证)标准的限制,完全按照无偏的自然人群样本调查收集到的四诊资料进行统计分析。

优点:不但可以选出某证候的主要四诊资料(指标),而且可以得到每个指标对证候的权重。例如主成分及因子分析、指标聚类分析。

问题:将导致证候的重新分类,可能与现有的辨证体系不完全符合。

上述第1种方法可当作暂行“金”标准。最大优点是:

- 1.辨证统一;
- 2.辨证的结果可以重复,即重测信度(retest reliability)好;
- 3.症状和证候都可以定量,均可作为症状和证候疗效评价的依据。

第三节 制订方法举例

以《中医内科学》(张伯臾主编:第一版,上海:上海科技出版社,1986)对胃痛的辨证分型和四诊资料项目为基础。我们再按四诊资料的严重度或发生频次评定得分,拟定四诊资料量表,称为每项四诊资料得分。

证候加权积分 = 各项四诊资料得分 × 各项四诊项目权重的积和(代数和)。

加权的方法:某一个证候中,每一项四诊资料出现频率的百分构成比(根据临床流行病学调查数据计算)。

上述四诊资料得分是等级资料,证候加权积分是计量资料。

表1 胃痛各证候的宏观辨证量表(胃痛的宏观辨证“金”标准)

症状	得分				权重
	0分	2分	4分	6分	
1. 寒邪客胃					
1.1 胃脘疼痛	无	轻微	较痛能忍受	不能忍受	0.3
1.2 得温痛减,遇寒痛增	从不	偶尔	经常	总是	0.3
1.3 口和不渴	非常渴	很渴	有点渴	不渴	0.1
1.4 喜热饮	从不	偶尔	经常	总是	0.1
1.5 舌苔白	不是			是	0.1
1.6 脉弦紧	不是			是	0.1

(证候加权分)		可根据每人各项具体值计算)				
2. 饮食停滞						
2.1	胃脘疼痛	无	轻微	轻痛能忍受	不能忍受	0.15
2.2	脘腹胀满	从不	偶尔	经常	总是	0.15
2.3	噎腐吞酸	从不	偶尔	经常	总是	0.15
2.4	吐不消化食物	从不	偶尔	经常	总是	0.15
2.5	吐食或矢气后 痛减	从不	偶尔	经常	总是	0.10
2.6	大便不爽	从不	偶尔	经常	总是	0.10
2.7	舌苔厚腻	不是	是(只中 间)	中间+根	中间+根+ 侧	0.10
2.8	脉滑	不是	右关	右关+左关	右关+左关 +右寸	0.10
3. 肝气犯胃						
3.1	胃脘疼痛	无	轻微	较痛能忍受	不能忍受	0.20
3.2	胃脘胀闷	从不	偶尔	经常	总是	0.15
3.3	脘痛连肋	从不	偶尔	经常	总是	0.15
3.4	噎气频繁	从不	偶尔	经常	总是	0.10
3.5	大便不畅	从不	偶尔	经常	总是	0.10
3.6	每因情志而痛 作	从不	偶尔	经常	总是	0.10
3.7	舌苔薄白	不是			是	0.10
3.8	脉沉弦	不是			是	0.10
4. 肝胃郁热						
4.1	胃脘作痛	无	轻微	较痛能忍受	不能忍受	0.15
4.2	痛势急迫	从不	偶尔	经常	总是	0.10
4.3	烦躁易怒	从不	偶尔	经常	总是	0.10
4.4	犯酸嘈杂	从不	偶尔	经常	总是	0.10
4.5	口干口苦	从不	偶尔	经常	总是	0.20
4.6	舌红苔黄	不是			是	0.20
4.7	脉弦或数	不是			是	0.15
5. 瘀血停滞						
5.1	胃脘疼痛	无	轻微	轻痛能忍受	不能忍受	0.15
5.2	痛有定处而拒 按	从不	偶尔	经常	总是	0.15
5.3	痛有针刺感	从不	偶尔	经常	总是	0.15

5.4	食后痛甚	从不	偶尔	经常	总是	0.15
5.5	吐血便黑	从不	偶尔	经常	总是	0.15
5.6	舌质紫黯	不是			是	0.15
5.7	脉涩	不是			是	0.10
6. 胃阴亏虚						
6.1	胃脘疼痛	无	轻微	较痛能忍受	不能忍受	0.20
6.2	痛隐隐	从不	偶尔	经常	总是	0.20
6.3	口燥咽干	从不	偶尔	经常	总是	0.15
6.4	大便干结	从不	偶尔	经常	总是	0.15
6.5	舌红少津	不是			是	0.15
6.6	脉细数	不是			是	0.15
7. 脾胃虚寒						
7.1	胃脘疼痛	无	轻微	较痛能忍受	不能忍受	0.10
7.2	痛隐隐	从不	偶尔	经常	总是	0.10
7.3	喜温喜按	从不	偶尔	经常	总是	0.10
7.4	空腹痛甚	从不	偶尔	经常	总是	0.15
7.5	得食痛减	从不	偶尔	经常	总是	0.15
7.6	泛吐清水	从不	偶尔	经常	总是	0.10
7.7	纳差	从不	偶尔	经常	总是	0.05
7.8	神疲乏力	从不	偶尔	经常	总是	0.05
7.9	手足不温	从不	偶尔	经常	总是	0.05
7.10	大便溏薄	从不	偶尔	经常	总是	0.05
7.11	舌苔淡白	不是			是	0.05
7.12	脉虚弱或迟缓	不是			是	0.05

参考文献

- 1.张伯臾主编：中医内科学.第1版.上海：上海科技出版社，1986
- 2.赖世隆：临床流行病学在中医药研究的应用.摘自：王家良主编.临床流行病学.第2版.上海：上海科学技术出版社，2001，437页

第二章 中医证候宏观辨证量表的表面效度评价：Kappa 值法

中医宏观症状的采集和辨证通常根据面谈、问卷或量表等测量手段完成的，这些手段获得的资料被公认是软数据（soft data），应该进行效度评价。

效度又称真实性(validity)：测量指标及观测结果在多大程度上反映了事件的真实性。

常用的效度评价方法有：

- 1.表面效度（face validity）；
- 2.内容效度（agreement validity）；
- 3.平行效度（convergent validity）；
- 4.预测效度（predictive validity）；
- 5.结构效度（structure validity）。

本章介绍使用 Kappa 值法对中医证候宏观辨证量表进行表面效度评价的方法。

Kappa 值是描述一致性的指标，K 越大，一致性越好， $k > 0.75$ 为好， $0.75 \geq k \geq 0.4$ 为较好， $k < 0.4$ 为差。

第一节 定量观察（多分类变量）的一致性评价方法

甲研究者使用中医证候宏观辨证量表独立地进行辨证，乙是专家的经验辨证。资料整理成表 1：

表 1 量表和专家各自独立辨证的结果（证候分度）

甲辨证的例数	乙辨证的例数				合计
	某证候 0 度	1 度	2 度	3 度	
某证候 0 度	26	6	3	0	35
1 度	11	29	10	1	51
2 度	5	12	30	7	54
3 度	0	1	5	28	34
合计	42	48	48	36	174

计算检验统计量 χ^2 和 Kappa 值:

$$\chi^2 = n [\sum (A_{ij})^2 / (m_i * m_j) - 1]$$

式中, i =行变量, j =列变量, n =对象例数, A_{ij} =第 i 行第 j 列的实际观察例数, m_i =第 i 行合计例数, m_j =第 j 列合计例数, 自由度 $df=(行数-1)(列数-1)=$

$(4-1)(4-1)=9$. 只有当 $\chi^2 > \chi^2_{0.05(9)}$, 即 $P < 0.05$ 时, 才计算 Kappa 值:

$$Kappa = (P_o - P_e) / (1 - P_e)$$

式中 P_o (符合率) = $\sum A_{ii}/n = (26+29+30+28)/174 = 0.65$;

P_e (期望符合率) = $\sum m_i * m_j / n^2 = (35 \times 42) + (51 \times 48) + (54 \times 48) + (34 \times 36) / 174^2 = 0.26$;

$$Kappa = (0.65 - 0.26) / (1 - 0.26) = 0.53$$

SAS 计算程序 (数据集名: ebml.Kappa) :

```
(1) libname ebml 'c:\sas\sasuser';
(2) data ebml.kappa;
(3) do i=0 to 3;
(4) do j=0 to 3;
(5) input f@@; output;
(6) end; end;
(7) cards;
(8) 26 6 3 0
(9) 11 29 10 1
(10) 5 12 30 7
(11) 0 1 5 28
(12);
(13) proc freq;
(14) tables i*j /chisq agree;
(15) weight f;
(16) run;
```

运行结果: Chi-square = 178.096, df = 9, p=0.001; simple Kappa = 0.529 (95%可信区间: 0.433, 0.626)

结果推断: 因为 Kappa 系数 = 0.53 > 0.4, 所以一致性较好。

第二节 定性观察（二分类变量）的一致性评价方法

如将证候分度简化为无和有两种二分类变量，则资料整理成表 2：

表 2 量表和专家各自独立辨证的结果（定性辨证）

甲辨证例数	乙辨证例数		合计
	是	非	
是	a	b	a+b
非	c	d	c+d
合计	a+c	b+d	n

$\chi^2 = [(ad-bc)\text{绝对值}-n/2]^2 \cdot n / (a+b)(a+c)(c+d)(b+d)$ 。自由度 $df=1$ ，

如果 $\chi^2 > \chi^2_{0.05(1)}$ ，即 $P < 0.05$ 时，认为具有关联性，这时才进一步计算符合率 P_0 。期望符合率 P_e ，Kappa 值：

$$P_0 = \frac{a+d}{n}$$

$$P_e = [(a+c)(a+b) + (c+d)(b+d)] / n^2$$

$$\text{Kappa} = (P_0 - P_e) / (1 - P_e)$$

参考文献：

- 1、方积乾主编. 医学统计学与电脑实验. 第 2 版. 上海：上海科学技术出版, 2001, 第 246、251、254 页
- 2、孙振球主编. 医学统计学. 北京：人民卫生出版社, 2002; 485-488 页