

440385

医用综合评价方法

孙振球 田风调等 编著



中国科学技术出版社

医用综合评价方法

孙振球 田凤调

刘 健 陈 科 编著

杨瑞璋 王广义

中国科学技术出版社

内容提要

本书编写受中国卫生统计学会的委托，在总结概括有关专题研讨会资料信息的基础上，参阅了大量国内外有关综合评价理论及应用实践的文献；阐述了综合评价的概念及一般程序，并结合实例，介绍了综合指数，综合评分，秩和比等目前常用的10余种综合评价方法及其医学应用；可供医学院校有关教学科研人员，各级医疗卫生行政机构和业务机构的有关管理人员和各级各类医务人员参考。

* * *

医用综合评价方法

孙振球 田凤调等 编著

责任编辑 邓俊峰 刘健
封面设计 刘健

*

中国科学技术出版社出版(北京海淀区北石桥路32号)

各地新华书店经销

湖南医科大学印刷厂印刷

*

开本 787×1092 毫米 1/32 印张 8.25 字数 194千字

1994年2月第1版 1994年2月第1次印刷

印数 1—2000册 定价 7.00元

目 录

第一章 概 述

一、多因素综合评价的概念	(1)
(一)综合评价的意义	(1)
(二)综合评价的任务与内容	(2)
(三)进行综合评价的基本条件	(4)
(四)综合评价的一般步骤	(11)
(五)综合评价的常用方法	(11)
二、评价因子的选择	(12)
(一)凭经验选择评价因子	(13)
(二)用单因素分析法挑选评价因子	(13)
(三)用多元相关分析法挑选评价因子	(14)
(四)用多元回归分析法挑选评价因子	(14)
(五)用逐步回归法挑选评价因子	(14)
(六)用岭回归方法挑选评价因子	(15)
(七)用 AIC 信息量准则挑选评价因子	(16)
(八)用指标聚类法挑选评价因子	(17)
三、评价因子的权重估计	(17)
(一)专家评分法	(18)
(二)模糊数学综合评价的逆方法	(31)
(三)统计计算法	(34)
四、多因素综合评价方法的回顾与展望	(35)

第二章 综合指数法

一、基本概念	(39)
(一)意义	(39)

(二) 分类	(39)
(三) 基本步骤	(40)
二、应用实例.....	(42)
(一) 在因素分析中的应用	(42)
(二) 在环境评价中的应用	(48)
(三) 在营养状况和生长发育状况评价中的应用	
.....	(50)
(四) 在社会医学与卫生事业管理中的应用	(52)
(五) 在临床医学中的应用	(57)

第三章 综合评分法

一、基本概念.....	(58)
(一) 意义	(58)
(二) 各评价指标诸等级分值的确定方法	(58)
(三) 综合评价总分计算方法	(61)
二、应用实例.....	(66)
(一) 在儿童少年生长发育评价中的应用	(66)
(二) 在临床医学中的应用	(68)
(三) 在卫生事业管理中的应用	(72)

第四章 秩和比法

一、基本概念.....	(74)
(一) 秩和比的定义与计算	(74)
(二) 秩和比法	(75)
二、应用实例.....	(82)
(一) 多指标综合评价—在社会医学与卫生事业 管理中的应用	(82)

(二)综合效应的测定—在环境卫生评价和劳动 卫生评价中的应用	(91)
三、几个有关问题的讨论	(98)
四、小结	(113)

第五章 交叉积差法(Cpd 分析)

一、基本概念	(115)
(一)Cpd 的定义与计算	(115)
(二)常用的 Cpd 分析方法	(117)
二、应用实例	(117)
(一)单因素 Cpd 分析及其在医学中的应用	(117)
(二)多因素 Cpd 分析及其在医学中的应用	(136)

第六章 模糊综合评判法

一、基本概念	(158)
(一)模糊子集的概念和模糊矩阵	(158)
(二)模糊综合评判模型	(168)
二、应用实例	(170)
(一)在卫生事业管理中的应用	(170)
(二)在环境评价中的应用	(178)
(三)在临床医学中的应用	(181)

第七章 多元统计分析方法 在综合评价中的应用

一、多元线性回归	(186)
(一)基本概念	(186)
(二)在医学中的应用实例	(187)

二、判别分析	(191)
(一)基本概念	(191)
(二)在医学中的应用实例	(195)
三、聚类分析	(199)
(一)基本概念	(199)
(二)在医学中的应用实例	(206)
四、主成份分析	(209)
(一)基本概念	(209)
(二)在医学中的应用实例	(210)

第八章 其他综合评价方法

一、优序法	(215)
(一)基本概念	(215)
(二)在医院管理中的应用实例	(215)
二、普通相关法	(217)
(一)基本概念	(217)
(二)在儿童生长发育评价中的应用实例	(217)
三、层次分析法	(221)
(一)基本概念	(221)
(二)在医院管理中的应用实例	(221)
四、灰色判别分析法	(226)
(一)基本概念	(226)
(二)在学生体质评价中的应用实例	(226)
五、综合评价图	(231)
(一)基本概念	(231)
(二)常用评价方法及其应用实例	(231)
参考文献	(239)

第一章 概 述

一、多因素综合评价的概念

(一) 综合评价的意义

本世纪以来,尤其是近30年来,科学技术的发展十分迅速,科研成果的应用周期与新技术的生命周期越来越短。据专家估计,单项新技术的生命周期每隔20年约缩短50%。例如:本世纪20年代,一项新技术的生命周期约40~50年;40年代约20~25年;50~60年代约12~15年;70~80年代约8~9年。

随着科学技术的普遍进步,医学科学当然也得到相应发展。基础医学、生物医学和现代信息科学理论与技术的日益进步,大大促进了预防医学及临床医学的发展。在医学科学的发展过程中,必然要淘汰一些过时的、低劣的方法或技术,改进一些有缺陷的,然而仍有应用价值的方法和技术,发展和创新一些高效的方法和技术。例如:在面对着临床新药物,新疗法,新技术如雨后春笋般涌现的时候,我们如何判断其效果的真伪?如何识别其利弊?另一方面,我们还必须认识影响人们生产生活的各种条件的现状,从而不断改善不利条件,发展有利条件,以求提高人们的健康素质。例如:在影响人们健康的诸多因素中,有必要区分什么是有利因素?什么是不利因素?什么是健康?什么是不健康?如此种种,给广大卫生工作者,包括卫生管理工作者,提出了一个如何对现有的方法和技术,以及人们生产生活条件及健康状况的优劣程度等进行客观评价的问题。

所谓评价，即价值的确定，是通过对照某些标准来判断测量结果，并赋予这种结果以一定的意义和价值的过程。测量结果本身，并无什么意义可言，必须加以分析、整理和评价之后，才有其实际意义和价值。例如身高 150cm，体重 40kg，仅就这两数字而言，没有什么实际意义。而当与某一年龄的生长发育标准进行对比时，就能看出其意义与价值了。因此，测量结果仅能反映现状，只有通过评价之后，才能对现状的意义加以判断。

考虑单一因素的评价易于实现，只要分别依据该因素给研究对象一个评价等级或分数，依等级或分数高低，便可排出优劣顺序；对于复杂的状况，因为同时受到多种因素的影响，必须结合考察多个有关因素，依据多个有关指标进行评价，这就是所谓综合评价。

有人曾指出，综合评价是对一个复杂系统的多个指标进行总评价的特殊方法。例如，某个儿童的营养状况评价，就是综合分析食物种类、数量、配比、吸收、疾病情况，以及身体发育、形态、机能、智力、遗传等有关因素后，而得到的总的印象，或总体的评价。当然，综合评价不同于多个指标分析的简单相加，而是在掌握有关历史资料的基础上，将各种有关因素的信息集中，依据其内在联系进行适当加工提炼，并密切结合医疗卫生工作实践，用数理统计方法或数学方法制订出恰当的评价模型，以谋求对评价对象的类别或优劣等级进行较为客观的判断，为医疗卫生工作决策提供依据。

（二）综合评价的任务与内容

医疗卫生工作的理论和实践是一个广泛的领域，因而有关的综合评价必然涉及到各个方面，有着十分丰富的内容。

依评价目的，可分为临床评价、卫生评价、管理评价与技

术评价等等。

临床评价包括诊断性试验和方法评价,用以评估某种诊断手段的应用价值,通常结合考察其敏感性、特异性与准确性进行综合评定。例如心电图运动试验对诊断冠状动脉狭窄的应用价值评定;疗效评价,用以评估各种临床治疗药物或疗法对某种或某类疾病的治疗效果,往往根据选定的多个疗效指标,对其科学性及实用性进行综合评定。例如内科疗法与外科疗法对颈总动脉病变所致一过性脑缺血的疗效评价;预后及转归评价,用以评估某些临床措施或病程中呈现的某些征象对疾病预后和转归的影响等等。

卫生评价包括环境评价,用以对生活环境或生产环境的优劣进行评估。例如大气质量、水质、土质污染程度的评价;营养评价,用以评估群体或个体营养状况,以及某些食品的营养价值等。例如,婴幼儿营养状况的评价;生长发育评价,用以对不同发育阶段的儿童及青少年体格发育与行为智力发育状况进行评价。

管理评价,包括卫生政策评价与卫生经济评价等,两者往往结合在一起,对医疗卫生政策,医疗卫生措施,医疗卫生单位管理水平,科研成果,科研方案等的优劣取舍进行综合评定。

医学技术评价中的“技术”指广义的技术,即医学科学用于实践的全过程,贯穿于从技术开发到商品化的各个阶段。在美国,技术评价的对象还包括有关法律的制订等等。总之,技术评价的对象既涉及到医学科学的广泛领域,也涉及到社会科学的广泛领域,其任务不仅在于对医学科研方案或医疗卫生措施从技术性能和经济效益等方面进行评价,还要着眼于社会,照顾周围环境,对其有利与不利影响进行综合评价。

按评价方式分，可分为事前评价、中间评价和事后评价。

事前评价，是在制订某项医疗措施计划时进行评价，这时还未开展大量的试验研究工作，还缺乏来源于实践的数据，主要是参考有关资料，汇集各方面意见，通盘考虑方案中的各种问题，制订切实可行的方案。这种评价具有预测性质，属探索性评价。

中间评价，是在大量进行试验研究工作之后进行的，着重验证设计或方案的正确性与可行性，研究暴露出来的问题，并采取必要的措施或对策，以决定在原计划或方案中应保留的部分，应改进的部分及应摒弃的部分。

事后评价，是在试验研究工作全部完成以后进行的，属于推广应用前的评价，着重全面审查研究成果，并与同类成果或技术在科学性、先进性、实用性、经济性等各方面进行综合比较，以决定优劣取舍。

对某一具体事件的评价，既涉及到前一种分类的内容，又涉及到后一种分类的内容；既包含对整个系统的评价，也包含对可靠性、可行性方面的评价；而且对于同一事件，依不同对象不同目的往往有不同的评价标准，这就使得评价工作复杂化多样化。当然，还得依据评价目的、方式、资料性质及应用要求等，采取恰当的评价方法。

(三)进行综合评价的基本条件

1. 首先要有一个高质量的内容丰富的信息源 所谓信息，是关于主体和客体之间有关情况的消息，是一种提供确定因素，消除不确定因素的消息。缺乏这种消息，便无法认识事物间的相互联系，也无法探求事物的规律。这些消息，有反映现状的，有反映历史的，有定量的，也有定性的……。从某种意义上讲，综合评价就是信息管理的全过程，即信息的收集，处

理和分析的过程，只有在充分占有有关评价对象及其相关因素的信息的基础上，才有可能作出较为可靠的评价。信息越多，越真实，评价的准确性与可靠性越高。

(1) 信息的收集。又称数据的收集，是一项至关重要的基础工作，有的数据来自第一手材料，有的来自第二手材料。

第一手材料，包括以各种形式直接收集到的数据，主要来源于三个方面：各种统计报表，各种医疗卫生工作记录和报告卡，以及专题调查或实验，后者包括各种形式的抽样调查，普查和典型调查。获取第一手资料往往需时较长，费用较高，但较为可靠。

第二手材料，多为已经公布或发表的有关资料，易于获取，代价较低，同时数据精度也易于保证。其缺点是可能不适用于直接应用。因而有些作者主张在应用这些数据时，可进行恰当的修正或处理。

收集的数据，应符合以下要求：

1) 完整：内容全面，无遗漏；范围齐全；时间连续；
2) 准确：这是最重要也是最基本的要求。资料应准确反映实际情况，各项目之间无矛盾，各数字无不合理现象；

3) 及时：有明确的时间限制，从某种意义上说，信息的价值取决于提供信息的时间；

4) 适用：信息能反映本质问题，包括那些实际上能产生潜在影响的信息，有利于提高工作效率；

5) 经济：要以最低的费用获取最多的信息。

(2) 信息的处理。信息的可靠性直接影响到评价的客观性，由于多方面的原因，我们收集到的资料的可靠程度不可避免地会受到某些干扰。同时，考虑到资料使用的方便，还有必要对数据进行处理。

信息处理，即用各种技术手段对原始资料进行审核、汇总和存贮，使之条理化、系统化的过程。包含清除那些不合理的可疑值，从而改善资料质量的过程；包含定性资料、等级资料与定量资料间的相互转换，尤其是定性或等级资料定量化，从而适于进一步分析的过程；也包含由已知信息来推算有关的未知信息，从而获得新的信息的过程。

1) 对原始数据进行审核、汇总和贮存：审核的目的在于检查资料的完整性、准确性与有效性，从而确定信息源的可靠性。对于原始记录中的遗漏或重复，对各项目之间存在的矛盾与不合理现象，应立即补充，剔除，改正；在此基础上，根据研究设计中整理分析计划的要求进行合理的分组汇总，并以适当的方式，例如以数据软件的形式贮存。

2) 可疑值的清除：在一组观测值中有时出现少数过大或过小的极端值，使人怀疑是否发生了错误，这种数值称为可疑值，往往由过失误差造成。对可疑值应反复检查，寻找发生原因，以便加以纠正；若找不出任何原因，不能随意舍去，可增加观测次数，或用专门的统计方法决定其取舍。

确定可疑值取舍的统计方法有多种，兹介绍3种常用方法：Chauvenet法、Smirnov法和Grubbs法。它们的计算方法相同，仅界值不同，计算步骤如下：

① 计算观察值（包括可疑值在内）的均数 \bar{X} 与标准差 s ，并按下式计算 T 值。

$$T = \frac{|X_i - \bar{X}|}{s} \quad (1.1)$$

式中 X_i 为可疑值。

② 查 T_{α}, n 界值表作出判断。表中所列为不同的界值 T_{α}, n, α 为第一类错误的概率， n 为样本含量。若 $T < T_{\alpha}, n$ ， X_i 不

能舍去,若 $T \geq T_{\alpha, n}$, 则 X_i 可舍去。

③若有几个可疑值,则先判断离均差绝对值最大者,若可舍去,再舍去该值后重新计算 \bar{X} 与 S ,按上法继续判断,直到不能舍去时为止。

表 1.1 检验可疑值用 $T_{\alpha, n}$ 界值表

n	Chauvenet 法		
	($\alpha = \frac{1}{2} n$)	Smirnov 法 ($\alpha = 0.05$)	Grubbs 法 ($\alpha = 0.05$)
6	1.73	1.996	1.82
7	1.79	2.093	1.94
8	1.86	2.172	2.03
9	1.92	2.237	2.11
10	1.96	2.294	2.18
11	2.00	2.343	2.23
12	2.04	2.387	2.29
13	2.07	2.429	2.33
14	2.10	2.461	2.37
15	2.13	2.493	2.41
16	2.16	2.523	2.44
17	2.18	2.551	2.47
18	2.20	2.577	2.50
19	2.22	2.600	2.53
20	2.24	2.623	2.56
25	2.33	2.717	2.66
30	2.39		2.75
35	2.45		2.82
40	2.50		2.87
50	2.58		2.96
60	2.64		3.03
80	2.74		3.14
100	2.81		3.21

这些方法不适用于偏态分布资料,如一组观察值中可疑值个数较多,则应检查观察值是否服从正态分布。

例如,用分光光度计测得血中转氨酶标准管的光密度为 0.16、0.16、0.17、0.19、0.30。问 0.30 可否舍去?

算得: $\bar{X} = 0.19167$, $S = 0.05419$

$$T = \frac{|0.30 - 0.19167|}{0.05419} = 1.999$$

查 $T_{\alpha, n}$ 界值表: Chauvenet 法 $T_{1/2 \times 6, 6} = 1.73$, Smirnov 法 $T_{0.05, 6} = 1.1996$, Grubbs 法 $T_{0.05, 6} = 1.82$ 。今 $T > T_{\alpha, n}$, 故按任一方法的界值, 0.30 均可舍去。

3) 各种类型资料间的数量转换: 在进行综合评价时, 必然涉及到众多的评价指标, 有些是定量的, 有些是定性的或按等级分类的。为了满足某些评价方法对资料的要求, 在某些情况下, 需要把一种类型的指标转换成另一种类型的指标。

计量资料转换成计数资料的过程较易实现, 只要依据有关专业知识, 合理地制订不同分类间的数量界限, 并对各观察单位进行清点计数, 便可将一组计量资料转变成按性质或等级分类的计数资料。例如, 脉博数次数(次/分钟)本为定量指标, 如果规定脉博数少于 60 次/分为缓脉, 60~100 次/分为正常, 多于 100 次/分为速脉, 则该指标就转换成了有三个分类的计数资料, 可按缓脉、正常、速脉三个分类来清点观察单位数。

计数资料转换成计量资料的过程较为复杂。一方面, 某些定性指标本身的定量尺度难于把握; 另一方面, 对那些按顺序分类的资料, 如何给各等级以恰当的定量, 以致于这种定量能较为准确地反映各等级间的差别, 也是一个值得探索的问题。医学现象本身是一个极其复杂的过程, 目前很多医疗卫生评价指标都缺乏进行客观度量的有效手段, 例如, 头痛或腹痛程度, 就只能凭病人的主诉与医生的判断加以度量。与医学有关的社会、政治、文化等诸方面指标的客观度量就更为困难了。因此, 医学现象的数量化过程, 是计量医学发展中的一个问

题，也是评价科学发展中的一个问题，有待于探讨。当然，模糊数学的发生与发展为医学现象的数量化提供了一个重要的参考手段。本文只介绍目前常用的最简单数量化方法。

对于按性质分类的计数资料，如果研究的属性限于两类，则一般用 0 和 1 将其转换成 $(0, 1)$ 变量 x 。如性别，有男女之分，则可规定男（或女）为 $x = 0$ ，女（或男）为 $x = 1$ ；如果研究的属性多于两类，则一般转换成多个 $(0, 1)$ 变量，如血型，有 O 型、A 型、B 型、AB 型四个分类，则可规定 O 型为 $x_1 = 1$ ，其他型为 $x_1 = 0$ ，A 型为 $x_2 = 1$ ，其他型为 $x_2 = 0$ ，B 型为 $x_3 = 1$ ，其他型为 $x_3 = 0$ ，AB 型为 $x_4 = 1$ ，其他型为 $x_4 = 0$ 。但由于一个人的血型必然属于四型中的一型，故 x_1, x_2, x_3, x_4 四个 $(0, 1)$ 变量中只有三个独立，可任意去掉一个，如果去掉 x_4 ，用 x_1, x_2, x_3 三个独立变量来量化血型这个指标，O 型用 $x_1 = 1, x_2 = 0, x_3 = 0$ 表示；A 型用 $x_1 = 0, x_2 = 1, x_3 = 0$ 表示；B 型用 $x_1 = 0, x_2 = 0, x_3 = 1$ 表示；AB 型用 $x_1 = 0, x_2 = 0, x_3 = 0$ 表示。一般说来，一个有 m 种属性分类的定性指标，可转换成 $m - 1$ 个 $(0, 1)$ 变量。

对于按等级分类的计数资料，常按照各等级由低到高的排列顺序转换成取值由小到大排列的一个变量 x ，各等级间的数据差异应酌情而定，可取公差不等的等差级数，甚至可取等比级数。例如腹痛这个指标，如果分成不痛、轻度痛、中度痛和重度痛四个等级，则可规定不痛为 $x = 0$ ，轻度痛为 $x = 1$ ，中度痛为 $x = 2$ ，重度痛为 $x = 3$ 。

④ 常用的信息推算法：依据已掌握的信息及其与客观事物间的联系，我们可以间接推算出由于各种原因不能直接测定的所需信息。

① 平衡法：利用平衡关系式推算某些短缺资料。例如，在

某年人口数=上年人口数+同年出生数-同年死亡数+同年迁入人口数-同年迁出人口数的平衡关系式中,如果缺少某一项数据,都可由其他几项数据推算出来。设不知同年迁出人口数,其他数据均为已知,则可由以上关系式推出:同年迁出人口数=上年人口数+同年出生数-同年死亡数+同年迁入数-某年人口数。

②几何增长速度法:在已知历史数据按等比级数增长的条件下,可用此法推算两个历史数据之间所短缺的某些历史资料。设已知某事物的平均发展速度 $x = \sqrt[n]{a_n/a_0}$, 其中 a_n 为资料中最后一年的某指标量, a_0 为第一年的某指标量, n 为年份差数, 则可据此推算 a_0 至 a_n 年间任意一个年份的某指标量, 设为 a_t , 则 a_t 为:

$$a_t = a_0(1+x)^y \quad (1.2)$$

式中 y 为推算年份与 a_0 之差值。

③因素推算法:利用现象内部各构成因素之间的变化关系进行资料推算。例如,由某年龄组人口数=总人口数×该年龄组人口构成比,便可在已知某地总人口数及人口构成的历史资料的基础上,推算某地各年龄组人口数。

2. 要提倡现成历史资料的综合利用 不少医疗卫生单位长期以来在利用各种现成的历史资料方面存在着调查研究多,资料分析少;登记材料与表格多,科学结论少;单指标分析多,多指标综合评价少等现象,这使得现有的信息得不到充分利用,因此,提倡对现有历史资料的综合利用,甚至着手实现信息的区域性合作与国际合作,充分利用现有的计算软件,提倡专业工作者与卫生统计工作者进行广泛深入的协作,将使得通过各种途径而获取的各种信息发挥最大的作用,也将使得各种形式的综合评价模型更加稳定可靠。