

医学综合评价方法 及其应用

孙振球 主 编
王乐三 副主编

Microsoft Excel - 例5-1数据文件与结果

宋体 12 B / U

文件(F) 编辑(E) 视图(V) 插入(I) 格式(O) 工具(T) 数据(D) 窗口(W) 帮助(H)

I16 f

	A	B	C	D
1	医院质量的综合评价			
2	评价指标	组合权重 C_i	一医院	二医院
3				
4	日均门诊人次	0.1980	1.2430	1.2400
5	日均住院人数	0.0990	1.3585	1.3710
6	实际病床使用率	0.0545	1.0742	1.0725
7	急诊占总诊疗人数百分	0.0882	1.1260	0.6840
8	病床利用效率值	0.1089	1.1703	1.1234
9	病房有效率值	0.2912	1.0247	1.0022
10	病房病死率	0.1602	1.2500	0.8696
11	综合评分指数GI		1.1646	1.1670
12	医院质量顺位		3	2
13				

就绪

H < > H \目标数据表\权重系数\综合评价表\目标树/

Chemical Industry Press



化学工业出版社
现代生物技术与医药科技出版中心

医学综合评价方法及其应用

孙振球 主 编

王乐三 副主编

图书在版编目 (CIP) 数据

医学综合评价方法及其应用/孙振球主编—北京:
化学工业出版社, 2005. 12
ISBN 7-5025-8147-2

I. 医… II. 孙 III. ①综合评价-方法②综合评价-应用-医学 IV. ①F224. 12②R

中国版本图书馆 CIP 数据核字 (2005) 第 157962 号

医学综合评价方法及其应用

孙振球 主 编

王乐三 副主编

责任编辑: 邵桂林 蔡 红

责任校对: 王素芹

封面设计: 关 飞

*

化 学 工 业 出 版 社
现代生物技术与医药科技出版中心 出版发行

(北京市朝阳区惠新里 3 号 邮政编码 100029)

购书咨询: (010)64982530

(010)64918013

购书传真: (010)64982630

<http://www.cip.com.cn>

*

新华书店北京发行所经销

北京云浩印刷有限责任公司印刷

三河市延风装订厂装订

开本 787mm×1092mm 1/16 印张 15½ 字数 330 千字

2006 年 2 月第 1 版 2006 年 2 月北京第 1 次印刷

ISBN 7-5025-8147-2

定 价: 35.00 元

版权所有 违者必究

该书如有缺页、倒页、脱页者, 本社发行部负责退换

《医学综合评价方法及其应用》编写人员

主 编 孙振球

副 主 编 王乐三

编写人员 (以姓氏汉语拼音为序)

杜方冬	胡 明	胡国清	胡平成
胡宇峰	黄刊迪	黄志平	李杏莉
刘明芝	史静琤	孙振球	王乐三
王一任	熊国强	许林勇	杨土保
易 东	曾小敏		

前 言

为适应医学科学领域中,尤其是医学科学研究、医学科学教育和医学科学管理中评价与决策的需要,我们在广泛参阅国内外综合评价理论及其应用相关文献的基础上,对《医用综合评价方法》进行了全面修改与着力补充,编著了《医学综合评价方法及其应用》。

本书简明阐述了综合评价的概念和一般程序;结合实例,介绍了综合评分法、综合指数法、秩和比法、层次分析法、TOPSIS法、模糊综合评判法、数据包络分析法、人工神经网络分析法、功效系数法、密切值法、灰色关联分析法、决策树法、普通相关法和优序法等当前较为常用的多种综合评价方法及其医学应用;从综合评价角度,介绍了当前常用的一些统计分析方法,包括各种多元线性回归分析法、正交设计分析法、Meta分析和生活质量测评等方法及其医学应用;在目前尚无综合评价专有计算机软件的情况下,尝试编制了有关计算程序。

本书是中南大学湘雅医学院(原湖南医科大学)、湖南中医学院和第三军医大学一批长期从事流行病学、卫生统计学、生物统计学以及医学信息管理科学教学与科学研究的正、副教授和在读博士研究生在综合评价及其医学应用领域中的经验总结,也是他们辛勤劳动的结晶。在本书编写规范中,力图内容全面,层次清晰,文字简洁,通俗易懂;注重基本概念与应用实例,简化操作步骤,并通过各种软件及自编程序实现有关计算,以使读者易于理解、易于操作、易于接受、举一反三。

本书可供医药院校有关教学科研及管理人员、各级医疗卫生行政和业务机构的有关管理与专业技术人员和各级各类医疗卫生工作者作为参考书使用;也可供各级各类医药院校本科生与硕士研究生作为教材使用。

中南大学湘雅公共卫生学院硕士研究生韦玉华、周绍英、邓艳霞、郭艳芳、张磊、夏强、郑磊、杨金星、莫苗芳、彭丽霞、李凌等参与了本书的校稿工作,谨此表示衷心的感谢。

受本人能力与水平所限,本书中不足之处在所难免,希望广大读者和医学界同仁指正。

孙振球

2005年6月于长沙

目 录

第一章 综合评价概论	1
第一节 综合评价的基本概念	1
一、基本概念	1
二、进行综合评价的基本条件	2
三、综合评价的一般步骤	6
四、综合评价的常用方法	6
第二节 评价指标的选择	7
一、凭经验选择评价指标	7
二、用单因素分析法挑选评价指标	8
三、用多元相关分析法挑选评价指标	8
四、用多元回归分析法挑选评价指标	8
五、用逐步回归法挑选评价因子	8
六、用岭回归法挑选评价指标	9
七、用 AIC 信息量准则挑选评价指标	9
八、用指标聚类法挑选评价指标	10
第三节 评价指标的权重估计	10
一、专家评分法	10
二、模糊数学综合评价的逆方法	18
三、统计计算法	19
第二章 综合评分法	20
第一节 基本概念	20
一、意义	20
二、各评价指标诸等级分值的确定方法	20
三、综合评价总分计算方法	22
第二节 应用实例	25
一、在儿童少年生长发育评价中的应用	25
二、在临床医学中的应用	26
三、在卫生事业管理中的应用	27
第三章 综合指数法	29
第一节 基本概念	29
一、指数的分类	29
二、基本步骤	30
三、综合指数评价法的用途	31
第二节 应用实例	31
第四章 秩和比法	36

第一节 基本概念	36
一、秩和比与秩和比法	36
二、秩和比法的基本步骤	37
三、常用的编秩技巧	39
第二节 应用实例	41
第五章 层次分析法	49
第一节 基本概念	49
第二节 应用实例	50
第六章 TOPSIS 法	53
第一节 基本概念	53
一、基本思想	53
二、基本步骤	53
第二节 应用实例	54
第七章 多元线性回归分析法	57
第一节 基本概念	57
一、多元线性回归模型	57
二、应用多元线性回归进行综合评价的步骤	57
三、多元线性回归中指标的数量化	60
第二节 应用实例	60
第八章 主成分分析法与因子分析法	63
第一节 基本概念	63
一、综合评价的主成分分析实现步骤	63
二、综合评价的因子分析实现步骤	64
第二节 应用实例	65
第九章 判别分析法	70
第一节 基本概念	70
第二节 应用实例	72
第十章 聚类分析法	75
第一节 基本概念	75
一、相似系数	75
二、类间相似系数的计算	76
三、应用注意事项	77
第二节 应用实例	77
一、系统聚类	77
二、动态样品聚类	81
三、有序样品聚类	81
第十一章 正交设计分析法	83
第一节 基本概念	83
一、正交表	83
二、试验步骤	86

第二节 结果分析及应用实例	89
一、极差分析法	89
二、方差分析法	93
三、多指标试验的综合分析	100
第十二章 模糊综合评判方法	103
第一节 基本概念	103
一、基本概念	103
二、模糊综合评判方法在综合评价中的应用	110
第二节 应用实例	112
第十三章 数据包络分析法	116
第一节 基本概念	116
一、数据包络分析法的意义	116
二、数据包络分析法的基本概念	118
三、数据包络分析法的经济学价值	122
四、DEA 模型的定义与算法	125
第二节 应用实例	137
一、分析步骤与方法	137
二、应用实例	139
三、数据包络分析的优点与不足	144
附 DEA 程序	145
程序 1 相对效率和投影的计算 (EP. m)	145
程序 2 规模收益分析和拥挤状态的确定 (SE. m)	146
程序 3 弱拥挤状态的确定 (SR. m)	148
程序 4 例 13-1 的计算程序 (Example1. m)	149
程序 5 例 13-2 第一组数据的计算程序 (Example2 _ 1. m)	150
程序 6 例 13-2 第二组数据的计算程序 (Example2 _ 2. m)	151
第十四章 人工神经网络分析法	152
第一节 基本概念	152
一、人工神经网络的意义	152
二、几种常用的人工神经网络模型	153
第二节 应用实例	156
一、人工神经网络在医院感染危险度综合评价中的应用	156
二、自组织神经网络方法	163
三、基于遗传算法的人工神经网络	165
第十五章 生活质量测评在综合评价中的应用	168
第一节 基本概念	168
第二节 生活质量评价	169
一、生活质量评价研究的设计	169
二、生活质量评价中的两个关键问题	171
第三节 生活质量评价工具的研制程序	172

第四节 高中生生活质量评价量表介绍	176
附 高中生生活质量测定量表 (HSS-QOLS)	178
第十六章 Meta 分析	182
第一节 基本概念	182
一、Meta 分析的意义	182
二、Meta 分析的步骤	182
三、Meta 分析的方法	185
四、综合效应量的检验	187
五、漏斗图和森林图	188
六、敏感性分析	188
第二节 分析步骤及应用实例	189
一、Meta 分析的软件	189
二、用 RevMan 软件进行 Meta 分析的步骤	189
三、建立了 Meta 分析新系统评价后的步骤	191
四、Meta 分析结果	193
五、建立系统评价后的操作过程说明	193
第十七章 综合评价的其他方法	197
第一节 密切值法	197
一、基本概念	197
二、基本步骤	197
三、应用实例	198
第二节 功效系数法	200
一、基本概念	200
二、基本步骤	200
三、应用实例	201
第三节 灰色关联分析法	203
一、基本概念	203
二、基本步骤	203
三、应用实例	204
第四节 决策树法	206
一、基本概念	206
二、基本步骤	207
三、应用实例	207
第五节 优序法	209
一、基本概念	209
二、基本步骤	209
三、应用实例	210
第六节 普通相关法	210
一、基本概念	210
二、基本步骤	211

三、应用实例·····	211
第十八章 常用综合评价方法的计算机实现·····	214
第一节 CES1.0 简介 ·····	214
一、CES1.0 界面 ·····	214
二、CES1.0 特点 ·····	214
三、CES1.0 运行环境 ·····	215
四、CES1.0 功能 ·····	215
第二节 应用实例·····	216
一、秩和比法应用实例·····	216
二、层次分析法应用实例·····	218
三、模糊综合评价应用实例·····	221
四、TOPSIS 法应用实例 ·····	222
五、密切值法应用实例·····	224
六、功效系数法应用实例·····	225
七、灰色关联分析应用实例·····	226
参考文献·····	232

第一章 综合评价概论

第一节 综合评价的基本概念

一、基本概念

随着决策科学的发展,评价科学也在不断发展。所谓评价 (evaluation), 是通过对照某些标准来判断观测结果, 并赋予这种结果一定意义和价值的过程。一般而言, 观测结果仅能反映现状, 只有通过评价之后, 才能对现状的意义加以判断。例如, 身高 140cm, 体重 35kg, 仅就这两个数字而言, 并没有什么实际意义, 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。

单一因素的评价易于实现, 只要按一定的准则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; 但是在医疗卫生实际工作中, 对于复杂的状况, 因同时受到多种因素的影响, 必须综合考察多个有关因素, 依据多个有关指标对评价对象进行评价, 并排出优劣顺序, 这就是所谓综合评价 (synthetical evaluation)。

有人指出, 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, 就是综合分析其所摄入的食物种类、数量、配比和其吸收、疾病情况, 以及其身体发育、形态、功能、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分析指标的简单相加, 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, 用数理统计方法或生物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及其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。当代的医疗卫生工作者, 在总结经验、考核效果和进行科学研究的过程中, 经常会遇到综合评价问题。本章将结合实际, 着重介绍综合评价的基本思想、基本步骤和一些常用方法。

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。

根据评价手段, 综合评价可分为定量评价 (quantitative evaluation) 与定性评价 (qualitative evaluation)。定量评价较为客观、全面, 易为人们所接受。

根据评价的领域, 综合评价可分为临床评价 (clinical evaluation)、卫生评价 (health evaluation) 和管理评价 (administrative evaluation) 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; 包括疗效评价, 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, 对其有效性及安全性进行综合评定, 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；包括预后及转归评价，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。

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、水质污染程度、土质污染程度的评价；包括营养评价，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；包括生长发育评价，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；还包括疾病防治效果评价、生活质量评价等。

管理评价包括宏观管理评价与微观管理评价，前者又包括卫生状况评价与卫生实力评价，以及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、教育教学质量、科研成果、科研方案等的优劣取舍进行综合评定。

按照评价方式，综合评价可分为预评价（pre-event evaluation）、中期评价（interim evaluation）和终结评价（after-event evaluation）。

预评价，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，制定切实可行的方案，这种评价具有预测性质，属探索性评价。

中期评价，是在进行大量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，以决定在原计划或原方案中应保留的部分、应改进的部分及应摒弃的部分。

终结评价，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，并在科学性、先进性、实用性、经济性等各方面与同类成果或技术进行综合比较，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；既包括对整个系统的评价，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和多样化。

二、进行综合评价的基本条件

1. 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，有的数据来自第二手材料。

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（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。其缺点是可能不宜于直接应用,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评价的客观性,由于多方面的原因,人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、汇总和存储,使之条理化、系统化的过程。包含清除那些不合理的可疑值(离群值, outlier),从而改善资料质量的过程;包含缺失数据(missing data)数量与性质的评价,并按照一定原则进行缺失值估计,从而保证资料完整性的过程;包含定性资料、等级资料与定量资料间的相互转换,尤其是定性资料或等级资料定量化,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、汇总和存储: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)存储。

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(或离群值),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。

通常情形下,对单变量计量资料离群值的查找,可使用标准分法或图示法。若据某观测值计算的标准分 $Z \geq 3.29$,该值可以称为离群值;直方图、箱式图、P-P图和Q-Q图等也可发现那些潜在的离群值。对多变量离群值的查找,可计算马氏距离(Mahalanobis distance)、杠杆值(leverage value)与影响值(influence value)等,读者可以参考有关专著。对服从正态分布的单变量计量资料可以考虑下述3种较为简易的方法剔除离群值,即 Chauvenet法、Smirnov法和Grubbs法,他们的计算方法相同,仅界值不同,计算步骤有如下3步。

a. 计算观察值(包括可疑值在内)的均数 $\bar{X}$ 与标准差 S ,并按公式(1-1)计算 T 值:

$$T = \frac{|X_i - \bar{X}|}{S} \quad (1-1)$$

式中, X_i 为可疑值。

b. 查 $T_{\alpha,n}$ 界值表作出判断。表 1-1 中所列为不同的界值 $T_{\alpha,n}$, 其中, 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。

表 1-1 检验可疑值用 $T_{\alpha,n}$ 界值表

n	Chauvenet 法 ($\alpha=1/2n$)	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

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

例如, 用分光光度计测得血液中转氨酶标准管的光密度为 0.16、0.16、0.17、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 $T_{0.05,6} = 1.996$, 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 缺失数据处理: 数据缺失是资料分析中常见的问题之一, 可由各种原因引起, 如动物死亡、调查对象失访等, 其危害取决于数据缺失的类型、缺失值的数量以及缺失原因等。

在完全随机或随机缺失的情形下, 如果缺失值数量不大 (例如不足全体数的 5%), 对缺失的数据可予删除或进行缺失值估计。常用的缺失值估计方法有经验法、均值替代法、回归法、期望最大法 (EM 法) 等, 读者可以参考有关专著。

④ 各种类型资料间的数量转换: 在进行综合评价时, 必然涉及众多的评价指标, 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容易实现, 只要依据有关专业知识, 合理地制定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换成按性质或等级分类的计数资料。例如, 脉搏 (次/min) 本为定量指标, 如果规定脉搏少于 60 次/min 为

缓脉、60~100 次/min 为正常、多于 100 次/min 为速脉,则该指标就转换成了有 3 个分类的计数资料,可按缓脉、正常、速脉 3 个分类来清点观察单位数。

计数资料与等级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,以便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;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 4 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 4 个类型中的一型,故 x_1 、 x_2 、 x_3 、 x_4 4 个 (0, 1) 变量中只有 3 个独立,可任意去掉 1 个,如果去掉 x_4 ,则用 x_1 、 x_2 、 x_3 3 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 ,各等级间的数量差异应酌情而定,可取公差不同的等差级数,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 4 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 $x=3$ 。

⑤ 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,人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。

a. 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+同年迁入人口数-某年人口数”。

b. 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 $a_0 \sim a_n$ 年间任意一个年份的某指标量,设为 a_k ,则

$$a_k = a_0(1+x)^y \quad (1-2)$$

式中, y 为推算年份与 a_0 之差值。

c. 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各年龄组人口构成比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

长期以来，不少医疗卫生单位在利用各种现成的历史资料方面存在着调查研究多而资料分析少、登记材料与表格多而科学结论少、单指标分析多而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机软件，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。

三、综合评价的一般步骤

对某事件进行多因素综合评价的过程，实质上就是一个科学研究与决策的过程，原则上应当包括设计、收集资料、整理资料和分析资料等几个基本阶段，在实施中应着重注意以下几个基本环节。

- ① 根据评价目的选择恰当的评价指标 (evaluation indicator)，考察各指标间的内在联系，选择那些主要的、能反映事物本质的评价指标，这些指标应当明确、具体、可行、可靠；
- ② 根据评价目的确定各评价指标在对某事物评价中的相对重要性，或各评价指标的权重；
- ③ 合理确定各单个指标的评价等级 (evaluation grade) 及其界限；
- ④ 根据评价目的、数据特征选择适当的综合评价办法，并根据已掌握的历史资料建立综合评价模型 (evaluation model)，计算综合指标；
- ⑤ 确定综合指标的等级数量界限，在对同类事物综合评价的应用实践中，对选用的评价模型进行考察，并不断修改补充，使之具有一定的科学性、实用性与先进性，然后推广应用。

例如，评估新生儿缺氧状况的 Apgar 评分方法，就是综合评价方法成功应用的例证之一。首先根据医学理论与临床经验，选择心率、呼吸等 5 个体征作为评价指标，并赋予相等的权重；然后依据理论与实践，确定各单个指标三个评价等级的界限及 0、1、2 三个分值的评分标准，建立如表 1-2 所示的评分标准；最后确定以累加法累计某评估对象各指标评分，并确定正常、轻度缺氧、重度缺氧三个等级的数量界限。通过实践检验，该模型仍然是产科临床用以判断新生儿有无窒息及窒息程度的常用方法。

表 1-2 新生儿 Apgar 评分标准

体 征	出生后 1min 内 评分		
	0	1	2
心率	无	<100 次/min	≥100 次/min
呼吸	无	浅慢,不规则	佳,哭声响
肌张力	松弛	四肢稍屈曲	四肢屈曲,活动好
反射	无	皱眉	啼哭
皮肤颜色	青紫、苍白	躯干粉红、肢端紫	全身红润

注：以累加法累计总分，8~10 分为正常；4~7 分为轻度缺氧；0~3 分为重度缺氧。

四、综合评价的常用方法

广义地讲，目前常用的多种医学统计学方法及其衍生的方法似乎都可用于进行综合评

价, 因为任何统计指标都综合了一定的有关信息。例如, 平均预期寿命这个统计指标就综合了某地某年居民健康状况、卫生状况、经济文化状况以及社会政治因素等多方面的信息, 或者说, 这个指标可用于对某人群上述几方面的状况进行综合评价。此外, 如多维列联表分析方法、析因试验设计分析方法、正交试验设计分析方法等, 都可综合多个因素对某一结果进行综合评价。

近年来, 随着电子计算机的发展而发展起来的多元统计分析方法, 如多元回归和逐步回归分析、判别分析、logistic 回归分析、因子分析与主成分分析、聚类分析、时间序列分析等, 已经在很多疾病的诊断、治疗、预后估计、危险因子分析以及少年儿童生长发育分析等方面得到成功应用, 无疑可作为综合评价的方法加以运用。近 30 年来, 随着模糊数学的发展而发展起来的模糊多元分析方法, 如模糊聚类、模糊判别、模糊综合评判等方法, 也大大丰富了综合评价方法学的内容。

此外, 在医疗卫生工作实践中, 人们还采用了一些较为简单、快速、实用而具有非参数色彩的综合评价方法, 如综合指数法、综合评分法、秩和比法、包络分析法、TOPSIS 法、密切值法、功效系数法、交叉积差法、综合图形法、优序法、普通相关法、灰色模型法、层次分析法、决策树法等, 尽管这些方法在理论上还存在这样或那样的问题(如还不能有效地进行误差估计等), 但已经应用于医疗卫生的各个领域, 尤其在儿科领域中对各个发育阶段儿童生长发育评价、围产医学领域中的胎龄评估、预防医学领域中生活及生产环境评价与营养评价, 以及医疗卫生管理科学中医院工作效率、科研方案评估等方面取得了较大的成功, 展示了这些方法可观的应用前景。这些方法将在以后各章中分别加以介绍。

第二节 评价指标的选择

在对某事件进行评价时, 必然要综合考察诸多因素的影响。这些因素有些是可控的, 有些是不可控的; 有些是独立的, 有些是相互关联的; 有些对评价结果影响小, 有些对评价结果影响大。人们有必要对影响因素进行分析, 力图分清主次, 抓住主要因子, 剔除次要因子, 一方面使得建立的评价模型简单化, 能就事件的主流或本质进行评价; 另一方面可以节省计算量, 并有利于提高评价模型的精度与准确度。

选择评价指标的方法甚多, 本文仅介绍以下几种方法。

一、凭经验选择评价指标

凭经验选择评价指标, 即根据有关的理论和实践来分析各个影响因子对评价结果的影响, 挑选那些代表性与确定性好、有一定区别能力又互相独立的指标组成评价指标体系。

“系统分析法”是一种常用的凭经验挑选因子的方法, 这种方法从整体出发, 将与评价结果有关的诸因子按系统(或属性、类别)划分, 在对各系统的因子进行分析的基础上, 通过座谈或填写调查表的方法获得对各因子的专家评分, 确定其主次, 再从各系统内挑选主要的因子作为评价因子。在缺乏有关历史资料或因子难于量化时, 此法可较简便地确定评价因子集。此外, 尚可采用“文献资料分析优选法”, 即全面查阅有关评价指标设置的文献资料, 分析各指标的优缺点并加以取舍。