维普资讯 http://www.cqvip.com

重庆医科大学学报 1992年第17卷第1期

⑥ 9a,17(1) 26-30

COX模型在肺癌手术预后因素分析中的应用

预防医学系 王昌玲 周燕荣 钟晓妮 金亚萍 曾 军
第一医院 殷毓琼 李良彬

R 734.207

提 要 本文对191例肺癌病人的术后随访资料，应用Cox回归模型从23个可能影响患者预后的因素中，确定对预后有影响的因素有：病理类型，肿瘤大小、淋巴结转移、手术方式、切端癌浸润及併发症。其中，对患者生存影响最大的不利因素是淋巴结转移，而有利因素是病理类型为鳞癌及采用肺叶切除手术。并通过按危险指数分组的生存曲线估计了模型的拟合度及样本人群的生存率。由此，阐明了将Cox回归模型用于多因素预后分析的优越性，并为临床医师估计患者预后、评价和选择治疗方法提供了科学依据

关键词 Cox回归模型；原发性支气管肺癌；多因素分析 肺癌 肿瘤
中图分类号 O242

在临床随访研究中，治疗效果的评价多采用生存期来衡量。对生存期的统计分析，传统的方法常采用寿命表法、列联表或多元线性回归。但是，寿命表及列联表不能同时考虑多个因素对生存期的影响，无法进行多因素分析，而多元线性回归又不能处理截尾数据。因此，这些传统的分析方法在实际应用上均有很大的局限性。

1972年英国统计学家D.R.Cox提出了一个解决多因素生存分析的统计模型——Cox回归模型[1]（也称比例危险模型），使生存分析技术有了重大进展。应用Cox模型进行生存分析国内已有报导[2-4]，但是应用Cox模型研究肺癌病人术后生存规律国内尚未见报告。本研究采用Cox模型对影响肺癌病人术后预后的多种因素进行分析，以揭示肺癌病人术后预后与各因素之间的关系，为估计患者预后，评价及选择治疗方法提供科学依据，并探讨Cox模型的应用效果。

Cox模型的结构、参数估计及检验：

假定病人的生存时间为t，同时病人具有p个与生存时间有关的变量（因素）x，$x=(x_1, x_2, x_3\cdots\cdots x_p)$。如果记特征为x的病人其在时刻t的危险函数为$h(t, x)$，则Cox回归模型的结构为：

$$h(t, x)=h_0(t)\exp(\beta_1x_1+\beta_2x_2+\cdots\cdots+\beta_px_p) \quad (1)$$

（1）式是Cox回归方程，它以危险函数$h(t,x)$作为应变量。危险函数$h(t,x)$表示生存时间已达到t的人，在时刻t的瞬时死亡（失效）率。所以，直观上看，对固定的时刻t，如果危险函数$h(t,x)$大，则说明在t时刻后立即死亡的可能性大。反之、危险函数$h(t,x)$小，则说明在t时刻后立即死亡的可能性小，继续活下去的可能性大。$h_0(t)$为基准危险函数，相当于各因素均处于零状态时的危险函数，为一不确定值。β是Cox模型中的参数，称为回归系数。$\beta=(\beta_1,\beta_2\cdots\cdots\beta_p)$。它可以描述各因素对生存影响的程度。Cox模型的参数估计采用最大偏似然法（即将数据代入一个偏似然函数中去，求出使该函数达到最大值的回归系数估计值）。对估计出来的参数及回归方程效果的检验，一般可采用wald检验或似然比检验[5-8]。

维普资讯 http://www.cqvip.com

资料及方法

我们收集到重庆医科大学第一附属医院胸外科1979年1月——1988年12月手术治疗的原发性支气管肺癌（本文简称肺癌）病人293例，其中除去由于病变广泛不能切除而仅作剖胸探查，手术后一个月内死亡以及资料不完整的病例外，可以做分析的有191例。191例患者均经病理检查证实为肺癌，以手术的日期作为观察起点，随访预后并按统一标准记录可能影响预后的各项指标。记录的项目包括：性别，年龄、病理类型、病理组织学级别，临床类型、肺癌分期、肿瘤大小、淋巴结转移、手术方式、切端癌浸润、併发症，其他肺部疾病、肺功能、其他治疗情况以及死亡日期和死亡原因。各指标经数量化（指标及其数量化见表1）以后，全部资料输入微型电子计算机，采用stata生存分析软件系统完成各项分析。

表1 观察指标及其数量化

指标	数量化	指标	数量化
1.性别(男、女)	x_1 女 0；男 1	8.淋巴结转移	x_{13} 有 1；无 0
2.年龄	x_2 <45 0；45～59 1；≥60 2	9.手术方式（肺楔形切除、肺叶切除、袖式肺叶切除、全肺切除）	x_{14} 肺楔形切除 1；否则 0 x_{15} 肺叶切除 1；否则 0 x_{16} 袖式肺叶切除 1；否则 0
3.病理类型（鳞癌、腺癌、小细胞、大细胞、其他）	x_3 鳞癌 1；否则 0 x_4 腺癌 1；否则 0 x_5 小细胞 1；否则 0 x_6 大细胞 1；否则 0	10.切端癌浸润	x_{17} 有 1；无 0
4.病理组织学级别（高分化、中分化、低分化、未分化）	x_7 高分化 1；否则 0 x_8 中分化 1；否则 0 x_9 低分化 1；否则 0	11.併发症	x_{18} 有 1；无 0
5.临床类型（中心型、周围型）	x_{10} 中心型 1；周围型 0	12.其他肺部疾病	x_{19} 有 1；无 0
6.肺癌分期	x_{11} Ⅰa 0；Ⅰb 1；Ⅱ 2；Ⅲ 3	13.肺功能	x_{20} 正常或基本正常 0；稍减低 1；显著减低 2
7.肿瘤大小	x_{12} <3cm 0；3cm～ 1；5cm～ 2	14.其他治疗（放射治疗、化学治疗、中草药治疗、未治疗）	x_{21} 有放疗 1；否则 0 x_{22} 有化疗 1；否则 0 x_{23} 有中草药 1；否则 0
		15.术后生存期	x_{24} 取实际生存期（天）
		16.指示变量	x_{25} 死于肺癌 1；否则 0

维普资讯 http://www.cqvip.com

计算结果

将191例随访资料作Cox模型拟合。计算模型中不包括任何伴随变量以及模型中包含某个（类）伴随变量时的最大似然函数的自然对数值和似然比统计量。结果发现：肺癌分期(x_{11})、淋巴结转移(x_{13})、在$\alpha=0.005$水平显著。年龄(x_2)、病理类型(x_3～6)、肺瘤大小(x_{12})、在$\alpha=0.05$水平显著。手术方式(x_{14}～16)、在$\alpha=0.1$水平显著。说明上述这些单指标对生存期有影响。又根据最大似然函数的自然对数值的大小，经似然比检验进行变量筛选，考察指标综合作用时对预后的影响。最终纳入Cox模型对预后有显著性意义的因素有：病理类型(x_3～6)、肿瘤大小(x_{12})、淋巴结转移(x_{13})、手术方式(x_{14-16})、切端癌浸润(x_{17})以及併发症(x_{18})[7]见表2。其余因素没有显示与预后有关。最终模型为：

表2　191例肺癌病例的Cox模型拟合结果

因素x	回归系数β	标准误SE	Wald值	P值*	RR
x_3	−0.9869	0.5157	−1.914	<0.05	0.37
x_4	0.1143	0.4852	0.236		
x_5	−0.5294	0.6694	0.791		
x_6	0.7877	0.8503	0.926		
x_{12}	0.4244	0.1484	2.859	<0.01	1.53
x_{13}	1.2730	0.2360	5.394	<0.01	3.57
x_{14}	0.8269	0.5319	1.555	<0.1	2.29
x_{15}	−0.5409	0.3407	−1.588	<0.1	0.58
x_{16}	−0.5659	0.4487	−0.126		
x_{17}	0.8094	0.4176	1.938	<0.05	2.25
x_{18}	1.6111	0.5443	2.960	<0.01	5.01

* 单侧Wald检验

$$h(t,x)=h_0(t)\exp(-0.9869x_3+0.1143x_4-0.5294x_5+0.7877x_6+0.4244x_{12}+1.2730x_{13}+0.8269x_{14}-0.5409x_{15}-0.5659x_{16}+0.8094x_{17}+1.6111x_{18})$$

经似然比检验 $\chi^2=64.22$ $(p\ll0.005)$ 差异有显著性意义，表明Cox回归方程成立。同时，对可能存在的交互作用进行了检验，结果未发现有交互作用。

由于特征为x的病人的危险函数h(t,x)(即预后好坏)可通过线性关系式：

$$I(x)=\beta_1x_1+\beta_2x_2+\cdots\cdots\beta_px_p$$

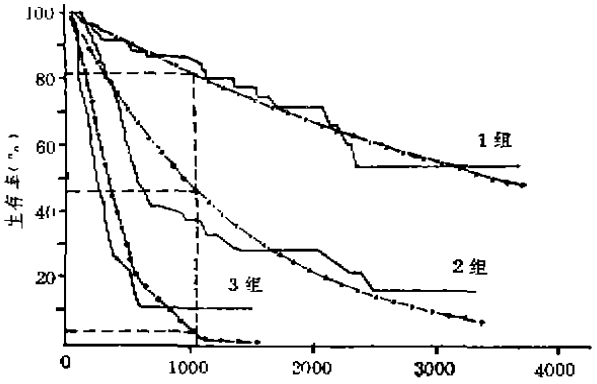
直接反映，因此，称I(x)为危险指数(预后指数)。本例：

$$I(x)=(-0.9869x_3+0.1143x_4-0.5294x_5+0.7877x_6+0.4244x_{12}+1.2730x_{13}+0.8269x_{14}-0.5409x_{15}-0.5659x_{16}+0.8094x_{17}+1.6111x_{18})$$

对每一个病人都根据其特征由上式算得了I(x)值。

为估计模型拟合度，将191个危险指数按其大小分成了三个（危险）组。在三个组预后差异有显著性意义的基础上，对每个危险组的生存数据用Cox模型和以往通常采用的Kaplan—Meier法估计了生存率并且作出图形。从图中可见，Cox模型生存曲线单调下降，并且与Kaplan—Meier法估计的生存曲线的趋势基本一致。因此，可以认为模型的拟合度比较好[7]

维普资讯 http://www.cqvip.com



图：不同危险组的生存曲线

"——" kaplan—meier估计的生存曲线

"-·-·-·" Cox模型生存曲线

1组	$I(x)\leqslant 0$	n=84	d=18
2组	$0<I(x)\leqslant 1.5$	n=81	d=45
3组	$I(x)>1.5$	n=26	d=19

讨　论

在预后分析中，以往常用的分析方法是单因素分析法。单因素分析肺癌病人术后生成率的影响因素，丁嘉安[8]等的2048例分析结果显示：手术方式和范围、癌肿类型和分期、淋巴结转移均显著影响术后生成率。黄志英[9]等的1286例肺癌中获随访的822例分析结果，预后以鳞癌最好，肺门或纵膈淋巴结有转移与无转移的存活率有显著性差异，中心型与周围型肺癌的预后无明显差异。本研究通过对191例肺癌病人术后的随访资料，采用Cox模型进行影响预后的多因素分析，除得到与上述单因素分析基本相同的结论外，还发现肿瘤大小、併发症、切端癌侵润三个因素也对预后有影响。由于应用Cox模型作预后因素分析不但在多种因素共存的情况下综合分析了各因素对预后的影响，而且合理处理了资料中的截尾数据，充分利用了资料提供的信息。因此，仅通过相对少的观察对象就获得了比较可靠的、合符实际的结论，提高了研究的效率。

应用Cox模型进行预后分析，可以定量地筛选出影响预后的主要因素，并根据回归系数的正负号判断出对预后有利或不利的因素，根据wald值的大小比较各因素对预后的影响强度。从表2可见，对肺癌病人生存影响最大的不利因素是淋巴结转移(x_{13})，其次依次是併发症(x_{18})、肿瘤大小(x_{12})、切端癌侵润(x_{17})以及采用肺楔形切除术(x_{14})。而对生存有利的因素是肺癌的病理类型为鳞癌(x_3)及采用肺叶切除术(x_{15})。这些结论为估计病人预后，评价治疗方法提供了依据。

通过按危险指数分组的生存曲线（见图），可以估计出样本人群的生存率。本例三个危险组的三年生存率分别为81%、46%和6%。还可用此图估计样本外人群的生存率，即用Cox模型对病人的预后进行预报。

同时，根据某一因素的回归系数β，也很容易计算出该因素的相对危险度RR。例如，一名鳞癌病人与一名非鳞癌病人（死亡）的相对危险度为：

$$RR=\frac{h(t,x)}{h_0(t)}=\exp(\beta_3 x_3)$$

$$=\exp(-0.9869)=0.37\text{（倍）}$$

表明鳞癌病人死亡的危险度比非鳞癌病人下降了63%，故鳞癌病人的生存也相应较长。

同理，可计算出每一因素的相对危险度。（见表2）。用Cox模型计算相对危险度反映某一因素对肺癌病人术后生存的影响程度不仅方便、直观，而且不必有发病率很低的假定。因此，它比Logistic回归更适合于流行病学资料的分析。总之，Cox模型用于多变量生存分析具有较多的优越性，是一种值得推广的统计分析方法。

（本文作图得到曾庆同志的帮助，特此致谢）

参　考　文　献

[1] Cox DR. J R Statist Soc 1972; B.34: 187

[2] 郭小伟，等. 上海第一医学院学报 1985;

维普资讯 http://www.cqvip.com

12(4):289
[3] 沈其君，等.中国卫生统计 1988；5(2)：30
[4] 施 榕，等.中国卫生统计 1989；6(1)：34
[5] 倪宗瓒主编.医学统计学 第一版 北京人民卫生出版社 1990：733
[6] 骆福添，等.中国卫生统计 1987；4(3)：39
[7] 骆福添，等.中国卫生统计 1987；4(4)：59
[8] 丁嘉安，等.中华医学杂志 1988；68(1)：23
[9] 黄志英，等.中华结核和呼吸杂志 1988,11(6)：330

SURVIVAL ANALYSIS OF POSTOPERATED PATIENTS WITH LUNG CANCER BY COX REGRESSION MODEL

Wang Changling, et al
Department of Health Statistics

This paper presents the application of Cox regression model in survival analysis of lung cancer for the first time in China. By using Cox model, a multivariate analysis was made on 191 patients with lung cancer treated by operation. Among the twenty three factors, only pathological sort, the size of tumor, lymph node metastases, type of operations, infiltra-ing across the end of tumor, and complications showed a high significance with prognostic relationship. Of those factors influecing postoperative survivaltime, lymph node metastases, is detrimental for patients, squamasum carcinoma and lobectomy are preferable. From survival curves of groups of different harzad groups, fit in the model and survival rate in the interclass was calculated, Therefore it showed advantages using Cox regression model in multivariate survival analysis

KEY WORDS Cox regression model; Lung cancer; Multivariate analisis