

肺癌发病危险评估方法

吴海云 潘平 何耀 刘岚 曾强 朱广卿 危成琦 黄建始

【摘要】 目的 建立一种适合于健康管理系统及社区预防保健中使用的个体肺癌发病危险的评估方法。方法 分析我国近 20 年来肺癌发病危险因素及发病率和病死率的资料,并结合国外的研究结果,由多学科专家小组讨论,达成共识,得出肺癌发病的主要危险因素及其相对危险度,按照哈佛癌症风险指数中使用的计算公式,计算个体肺癌发病的相对风险和绝对风险。结果 建立了一种可通过病史及生活方式问卷资料评估个体肺癌发病风险的计算方法。结论 该方法可用于评估个体肺癌发病风险,提示个体不同危险因素对其肺癌发病风险的影响。

【关键词】 肺肿瘤; 危险因素; 风险性评估; 数学模型

A risk model for predicting lung cancer WU Hai-yun*, PAN Ping, HE Yao, LIU Lan, ZENG Qiang, ZHU Guang-qing, WEI Cheng-yun, HUANG Jian-shi* Institute of Geriatric Cardiology, the General Hospital of Chinese People's Liberation Army, Beijing 100853, China
Corresponding author: WU Hai-yun, Email: why46301@163.com

【Abstract】 Objective To develop a mathematical model used in health management system and community medical care as a tool for health risk appraisal to predict the lung cancer risk of Chinese population. Methods Data of lung cancer in Chinese of the past 2 decades was collected and analyzed, group consensus was performed among multidisciplinary researchers to identify major risk factors and their relative risk. The equation of the Harvard Cancer Risk Index was used to compute the relative and absolute lung cancer risk. Results A mathematical model for the prediction of lung cancer risk was developed by way of medical history and lifestyle questionnaire. Conclusions This model can be used in Chinese for the risk estimation of lung cancer, beneficial to inform users of the major risk factors for lung cancer.

【Key words】 Lung neoplasms; Risk factors; Risk assessment; Mathematical model

肺癌是世界范围内引起死亡的最常见的恶性肿瘤^[1]。在我国,肺癌是城市居民的头号癌症杀手,在农村其发病率也在迅速上升^[2]。目前,大多数肺癌患者获得诊断时已属晚期,致使其 5 年生存率极低。同时,肺癌又是目前病因最为明确,预防最有办法的癌症之一。现在有确凿的证据表明,80% 以上的肺癌是由吸烟引起的^[3]。其他的一些可预防的因素在肺癌的发病中也起一定作用。因此,早期筛查高危人群,通过健康教育和生活方式改变,控制危险因素,从而减少肺癌的发生,具有十分重要的公共卫生学意义。健康风险评估是研究致病危险因素与慢性疾病发病率及病死率之间数量依存关系及规律

的技术^[4]。通过健康风险评估,可以确定各种危险因素对某一疾病发生所起的作用强度,识别高风险人群,明确预防的重点,有利于帮助个体认识健康危险因素,强化个人的健康促进行为,制订个体化的健康干预措施,并评价这些措施的效果^[5]。我们采用由多学科专家小组讨论,达成共识的方法,依据我国近 20 年来肺癌流行病学资料,确定成年人肺癌发病的主要危险因素及其相对危险度,并应用哈佛癌症风险指数工作小组提出的计算公式^[6],建立个体肺癌风险评估计算模型,以用于健康管理系统及社区预防保健和健康教育。

资料与方法

1. 组成多学科专家小组:成立的专家小组成员中,包括流行病学研究人员,医学统计学研究人员,肿瘤临床医生及从事健康管理和健康教育的专业人员。小组成员分次座谈,查找、选择和确定我国近 20 年来肺癌发病的相关流行病学资料,在充分讨论的基础上,对预测变量的选取、数据的取舍和计算模

作者单位:100853 北京,解放军总医院老年心血管病研究所(吴海云、何耀、曾强);东方脉搏健康管理研究所(潘平);航天中心医院体检中心(刘岚);解放军空军总医院肿瘤科(朱广卿);北京市中西医结合医院内科(危成琦);中国医学科学院北京协和医学院清华大学医学部(黄建始)

通信作者:吴海云,Email: why46301@163.com

型所应用的取值达成共识。

2. 流行病学资料的获取:采用“肺癌”、“危险因素”、“流行病学”、“发病率”、“病例对照”等中文词及其组合,通过查找关键词、主题词、摘要和题目,检索国内主要数据库,包括 CNKI 中国学术期刊网全文数据库,万方数据库,并采用“lung cancer”、“Chinese”、“risk factor”、“case-control”等英文词,检索 Medline,结合手工检索及文献追溯的方法,查找近 20 年来国内外主要学术期刊上发表的中国人肺癌相关流行病学研究资料。同时,结合政府部门、学术机构和专题研究小组发表的专题报告和统计报告,如《中国卫生年鉴》,卫生部发布的《中国居民营养与健康状况调查报告》等。选择资料时,尽可能选用基于我国人群研究、样本较大的重大项目或课题研究资料。若国内资料缺失或不充分,则由小组成员参考国外相关研究资料,讨论确定。

3. 建立计算模型:模型采用的数学计算公式为哈佛大学风险指数工作小组提出的计算公式:

$$RR = \frac{RR_{P_1} \times RR_{P_2} \times \dots \times RR_{P_n}}{[(P_1 \times RR_{G_1}) + (1 - P_1) \times 1.0] \times [(P_2 \times RR_{G_2}) + (1 - P_2) \times 1.0] \times \dots \times [(P_n \times RR_{G_n}) + (1 - P_n) \times 1.0]}$$

公式中,RR 为被预测个体患某一疾病与其同性别年龄组一般人群比较的相对风险。RR, 指个体中存在的危险因素的相对危险度;P 为其同性别年龄组人群中暴露于某一危险因素者的比例;RR_G 为由专家小组对某一危险因素(包括不同分层)的相对危险度达成共识的赋值^[6]。

计算出个体患病的相对风险后,可与其同性别年龄组一般人群比较,参照哈佛大学风险指数工作小组制订的标准^[7](表 1),表达为低于一般人群或高于一般人群等 5 个等级。若乘以其同性别年龄组一般人群某病的发病率,即可算出个体患病的绝对风险值。

4. 开发评估软件:由专家小组成员与软件公司合作,开发基于计算机运算的评估软件。

表 1 被预测个体与同性别年龄组一般人群比较的风险大小

相对危险	危险水平
0-0.5	显著低于一般人群
0.5<0.9	低于一般人群
0.9<1.1	相当于一般人群
1.1<2.0	高于一般人群
>2.0	显著高于一般人群

结 果

经讨论达成共识的肺癌发病危险因素及其相对危险度见表 2。若某项危险因素资料缺失,则按不

存在某项危险因素处理,但在提供风险评估报告时予以说明。

表 2 肺癌危险因素及其相对危险度

危险因素	相对危险度
吸烟	
被动吸烟(不吸烟女性)	1.3
已戒烟	2.0*
吸烟指数 < 100	1.8
吸烟指数 100-199	2.6
吸烟指数 200-299	4.2
吸烟指数 300-399	5.8
吸烟指数 ≥ 400	8
吸烟斗或旱烟	4.6
职业接触史	
无职业接触史	1.0
石棉	9.0
硅粉尘	2.6
煤烟或焦油	2.2
空气污染(大城市生活)	1.3
肺癌家族史	1.6
既往病史	
肺结核史	2.6
慢性支气管炎	2.4
肺炎病史	2.0
蔬菜水果摄入量 < 400 g/d	1.4

注:*若戒烟年数 ≥ 10,则相对危险度为 1

代入上述公式,该男性与同性别年龄组一般人群比较,其患肺癌的相对风险为:

举例:男性,46 岁,吸卷烟 20 年,每天吸烟 16 支;无职业性粉尘接触史;生活在北京。父母及兄弟姐妹无患肺癌者,无慢性肺病及肺炎史,每日蔬菜水果摄入量 400 g。则其输入计算公式的相应值见表 3。

$$RR = 5.8 \times 1.3 / [(2.0 \times 0.014) + (1 - 0.014)] \times [(1.8 \times 0.07) + (1 - 0.07)] \times [(2.6 \times 0.11) + (1 - 0.11)] \times [(4.2 \times 0.14) + (1 - 0.14)] \times [(5.8 \times 0.16) + (1 - 0.16)] \times [(8.0 \times 0.12) + (1 - 0.12)] \times [(4.6 \times 0.05) + (1 - 0.05)] \times [(9.0 \times 0.006) + (1 - 0.006)] \times [(2.6 \times 0.024) + (1 - 0.024)] \times [(2.2 \times 0.01) + (1 - 0.01)] \times [(1.3 \times 0.14) + (1 - 0.14)] \times [(1.6 \times 0.12) + (1 - 0.12)] \times [(2.6 \times 0.04) + (1 - 0.04)] \times [(2.4 \times 0.04) + (1 - 0.04)] \times [(2.0 \times 0.06) + (1 - 0.06)] \times [(1.4 \times 0.56) + (1 - 0.56)] = 0.64$$

亦即该男性肺癌发病风险为其同性别年龄组一般人群的 0.64 倍,文字上可表达为其肺癌发病风险低于一般人群。我国该男性所在年龄组一般人群肺癌发病率为 32/10 万,其今后 5 年肺癌发病的绝对风险为:5 × 0.64 × 32/10⁵ = 102.4/10⁵,但应考虑该年龄组中,肺癌发病风险随年龄增加会显著增加,故最后所得值还应以其年龄组的增长率校正。该年龄

段每年肺癌发病率增加 10%，最终计算的该男性 5 年肺癌发病的绝对风险为 $102.4/10^5 \times (1 + 10\%) = 16^5/105$ ，即 0.165%。

表 3 举例男性输入计算公式的相应值

危险因素	RR _i	RR _c	相应危险因素 人群暴露率(%)
吸烟			
已戒烟	1.0	2.0	0.01
吸烟指数 < 100	1.0	1.8	0.07
吸烟指数 < 199	1.0	2.6	0.11
吸烟指数 < 299	1.0	4.2	0.14
吸烟指数 < 399	5.8	5.8	0.16
吸烟指数 ≥ 400	1.0	8.0	0.12
吸烟斗或旱烟	1.0	4.6	0.05
职业接触史			
石棉	1.0	9.0	0.00
硅粉尘	1.0	2.6	0.02
煤烟或焦油	1.0	2.2	0.01
空气污染(大城市生活)	1.3	1.3	0.14
肺癌家族史	1.0	1.6	0.12
既往病史			
肺结核史	1.0	2.6	0.04
慢性支气管炎	1.0	2.4	0.04
肺炎史	1.0	2.0	0.06
蔬菜水果摄入量 < 400 g/d	1.0	1.4	0.56

注:RR_i:个体中存在的危险因素的相对危险度;RR_c:由专家小组对某一危险因素(包括不同分层)的相对危险度达成共识的赋值

其中吸烟是可降低的危险因素。若该男性戒烟,则其风险可降到一般人群的: $0.64 \times 2/5.8 = 0.22$ 倍,其今后 5 年内肺癌发病风险可降为 0.057%,即可下降约 2/3。

讨 论

近年来,健康管理在我国无论作为一种预防和控制疾病,促进个体和群体健康的理论思想,或是作为一种实践策略,都已引起人们越来越广泛的关注。所谓健康管理,即是对个体或群体的健康进行全面监测、分析、评估,提供健康咨询和指导,以及对健康危险因素进行干预的过程^[8]。对个体而言,健康管理主要包括 3 个环节:(1)收集疾病风险相关临床资料,测量危险因素;(2)进行健康风险评估;(3)在风险评估基础上,采取相应干预措施,预防或减少未来疾病的发生^[5]。其中,健康风险评估被普遍认为是进行健康管理的基础和核心环节。

20 世纪 60 年代起,美国国立心肺和血液研究所在 Framingham 心脏研究基础上,确定了心血管病的许多独立危险因素,建立了一系列危险因素分数表,通过对大样本人群的随访验证、多种基于计算机

技术的统计方法(logistic 回归、Cox 比例风险模型等)及预测变量的调整优化,相继开发了用于评估冠心病及脑卒中发病危险的数学预测模型^[9]。1985 年,Wasson 等^[10]提出了“临床预测”的原则和方法,许多学者其后对此进行了完善和补充^[11]。20 世纪 90 年代初,美国 Arizona 大学发展了一个数量化计分的癌症危险度评价工具^[12]。许多国际著名的医学研究机构,包括哈佛大学、梅奥医院、密西根大学等,还分别开发了用于多种恶性肿瘤(如肺癌、胃癌、乳腺癌、大肠癌)、糖尿病、骨质疏松症等发病风险的评估模型^[1,13]。欧洲及日本、韩国等亚洲国家,亦分别开发了适合本国人群的疾病风险评估工具。这些评估方法,随着互联网技术的快速普及,近年来得到了飞速发展。目前,在欧美国家,健康风险评估不仅在政府机构、学术团体、大学及医疗机构及其网站上广泛应用,也被许多提供健康服务的商业性机构广泛采纳。

在我国,20 世纪 80 年代末,陈坤等报道了基于模糊数学隶属度判别用于人群大肠癌筛查的数量化方法^[14]。其后,研究者们分别建立了我国人群冠心病、肺癌、大肠癌等的个体发病风险评估模型^[15]。另外,由流行病学及临床学者合作的课题组,提出了不同的国人缺血性心血管病发病风险的评估方法,开发了适合临床使用的简易评估工具^[16]。近几年来,国内多家商业性的健康管理机构还分别引进了国外的一些疾病风险评估系统,推动了国内疾病风险评估技术的发展。

迄今为止用于疾病风险评估的数学模型有多种。1972 年,Rothman 和 Keller^[17]提出了基于危险因素慢性病发病危险的计算方法,其后的许多计算方法,虽然形式不同,但其原理相似。用于建立评估模型的统计学技术,主要包括实证分析法、logistic 回归、比例风险模型、对数发病率、贝叶斯分析、马可夫模型及决策理论等^[18]。其基本步骤,一般包括:(1)选择危险因素,计算其组合的相对危险度;(2)确定人群归因危险分数;(3)获得不同年龄组人群基准发病风险的数据;(4)统筹考虑年龄特异性的其他致死性疾病的影响;(5)计算个体在一定时期内的绝对发病危险。实际使用中,可以采用评分相加的方法,根据对个体计算出的分值得出与其相对应的发病风险值。这种方法不需要较复杂的计算。但随着计算机和网络的日益普及,将个人数据直接输入计算公式,得出计算结果的方法,被越来越多地采用。建立评估模型的资料来源,可以是通

大型人群队列研究,经多因素回归分析得出某一疾病的相关预测变量,如 Framingham 心脏研究建立的心血管病风险模型。这是建立模型的基本方法。另外,也可以通过对多项研究资料的综合分析,确定相关的预测变量及其赋值。哈佛大学癌症风险指数工作小组提出的癌症风险指数采用的即为后一种方法。我国迄今尚缺乏肺癌发病的大型前瞻性队列研究,采用后一种方法更为可行。

国外已报道的肺癌发病危险评估模型有多种,如 Peto 等^[19]、Bach 等^[20]、Prindville 等^[21]及 Cassidy 等^[3]提出的模型。纳入的危险因素分别包括吸烟、戒烟、石棉接触史、FEV₁/FVC 值、痰细胞学检查及呼吸性病史。我们的计算模型中,纳入的肺癌发病危险因素均为目前已经大量流行病学研究,证据比较确实的因素,其中主要是吸烟,也纳入了多个可经病史和生活方式问卷获得的因素,包括职业接触、生活环境,既往肺病史及蔬菜水果摄入情况等。我们未纳入 FEV₁/FVC 值和痰细胞学检查结果,主要是考虑到这些数据在实际工作中难以获得。出于同样的考虑,我们的评估模型中,也未纳入一些新近提出的细胞学或生化学指标^[32]。摄入新鲜蔬菜少与肺癌发病的关系,其证据虽然相对较弱,但该项因素在健康教育中的实际意义较大,因此我们将其纳入评估模型中。

哈佛大学癌症风险指数中,只计算相对风险,而未计算绝对风险,主要是考虑到美国人群中种族背景差异很大,各人群的癌症发病风险相关较多^[4]。我们认为,我国人群的种族背景较单纯,人群间的遗传背景差异不大,因此,计算肺癌发病的绝对风险,在不同人群间不会引起很大的偏差。另外,给出绝对风险值,可以使风险的表达更为直观,在健康教育中会有更好的效果。

总之,我们通过分析我国近 20 年来肺癌发病危险因素及发病率 and 病死率的资料,并结合国外的研究结果,由多学科专家小组讨论,达成共识,得出肺癌发病的主要危险因素及其相对危险度,按照哈佛大学癌症风险指数中使用的计算公式,建立了一种可通过病史及生活方式问卷资料评估个体肺癌发病风险的评估模型,该模型可用于评估个体肺癌发病风险,提示个体不同危险因素对其肺癌发病风险的影响。

参 考 文 献

[1] Spitz MR, Hong WK, Anas CI, et al. A risk model for prediction of lung cancer. *J Natl Cancer Inst*, 2007, 99:715-726.

- [2] 杨功焕. 中国人群死亡因素及其他危险因素流行水平、趋势和分布. 北京:中国协和医科大学出版社, 2005.
- [3] Cassidy A, Duffy SW, Myles JP, et al. Lung cancer risk prediction: a tool for early detection. *Int J Cancer*, 2007, 120:1-6.
- [4] 龚幼龙. 社会医学. 北京:人民卫生出版社, 2000: 103-119.
- [5] Greenland P, Grundy S, Pasternak BC, et al. Problems on the pathway from risk assessment to risk reduction. *Circulation*, 1998, 97:1761-1762.
- [6] Kin DJ, Rockhill B, Golditz GA. Validation of the Harvard Cancer Risk Index: a prediction tool for individual cancer risk. *J Clin Epidemiol*, 2004, 57:332-340.
- [7] Saka M, Yoshida K. Health risk appraisal using internet technology. *Methods Inf Med*, 2005, 44:310-314.
- [8] Golditz GA, Atwood KA, Ermons K, et al. Harvard report on cancer prevention volume 4: Harvard cancer risk index. *Cancer Causes Control*, 2000, 11:477-488.
- [9] Wilson PW, D'Agostino RB, Levy D, et al. Prediction of coronary heart disease using risk factor categories. *Circulation*, 1998, 97:1837-1847.
- [10] Wasson JH, Sox HC, Neff RK, et al. Clinical prediction rules: applications and methodological standards. *N Engl J Med*, 1985, 313:793-799.
- [11] Laupacis A, Sekar N, Stiell IG. Clinical prediction rules: a review and suggested modifications of methodological standards. *JAMA*, 1997, 277:488-494.
- [12] Brown SL. Quantitative risk assessment of environmental hazards. *Annu Rev Public Health*, 1985, 6:6247-6267.
- [13] Gail MH, Brinton LA, Byer DP, et al. Projecting individualized probabilities of developing breast cancer for white females who are being examined annually. *J Natl Cancer Inst*, 1989, 81:1879-1886.
- [14] 陈冲, 俞维萍, 朱益民. 胃癌个体危险度及人群筛查数量化评价的研究. *中华流行病学杂志*, 2003, 24:310-315.
- [15] 郑颖颖, 蒋华. 肺癌个体健康危险因素评价的研究. *中国全科医学杂志*, 2003, 6:54-56.
- [16] 国家“十五”攻关“冠心病、卒中中综合危险度评估及干预方案的研究”课题组. 国人缺血性心脏病发病风险的评估方法及预后评估工具的开发研究. *中华心血管病杂志*, 2003, 31:893-901.
- [17] Rothman K, Keller A. The effect of joint exposure to alcohol and tobacco on risk of cancer of the mouth and pharynx. *J Chronic Dis*, 1972, 25:711-716.
- [18] Freedman AN, Seminaro D, Gail MH, et al. Cancer risk prediction models: a workshop on development, evaluation, and application. *J Natl Cancer Inst*, 2005, 97:715-723.
- [19] Peto R, Darby S, Deo H, et al. Smoking, smoking cessation, and lung cancer in the UK since 1950: combination of national statistics with two case-control studies. *BMJ*, 2000, 321:323-329.
- [20] Bach PB, Kattan MW, Thomsen MD, et al. Variations in lung cancer risk among smokers. *J Natl Cancer Inst*, 2003, 95:470-478.
- [21] Prindville SA, Byers T, Hirsch FR, et al. Sputum cytological atypia as a predictor of incident lung cancer in a cohort of heavy smokers with airflow obstruction. *Cancer Epidemiol Biomarkers Prev*, 2003, 12:987-993.
- [22] Phillips M, Albrecht N, Austin JH, et al. Prediction of lung cancer using volatile biomarkers in breath. *Cancer Biomark*, 2007, 3:95-109.

(收稿日期:2007-10-03)

(本文编辑:高健)