

好的流行病学实践和肺癌病因探讨

杜应秀¹ Joseph M. Wu²

一、肺癌病因探讨的简短回顾:肺癌是一种典型的与环境因素及生活方式有关的疾病,近三、四十年来,世界许多地方特别是大城市中肺癌发病率的持续增加,引起了广泛的注意。目前虽已提出了不少与肺癌病因有关的危险因素,但仍存在许多有待探讨的问题。

吸烟对肺癌的重要性是世所共知的,但仍有令人费解之处。例如:为什么中国、日本、韩国男性吸烟率相当高,但肺癌患病率却低于吸烟率较低的男性白人^[1]。中国城市居民吸烟率与农村居民相似,但肺癌患病率城市远高于农村^[2]。Alavanja^[1]还认为严重吸烟者中大约只有 10% 的人患肺癌而不是 100%。

被动吸烟与肺癌的关系更是莫衷一是,有着完全相反的报道。其中较难解释的是,吸烟主要是诱发鳞癌,而主流烟雾和侧流烟雾的主要成分基本一致,被动吸烟如能诱发肺癌,亦应是鳞癌,但认为被动吸烟能诱发肺癌的报道则多为腺癌。被动吸烟如能诱发腺癌,其潜隐期一定会比主动吸烟诱发鳞癌的时间更长,但通常腺癌患者都比鳞癌患者年轻^[3]。

大气污染与肺癌的关系早已受到注意,有人甚至提出肺癌发病率与大气中的 BaP 浓度有剂量-效应关系。但是大气污染情况瞬息万变,人们接触污染大气的情况也大不相同,用统计方法所计算出来的结果是否能真正符合实际情况很难肯定。况且居住于同一污染地区的居民,其肺癌发病情况并不相同。室内空气污染在中国曾被强调^[4],特别是认为烧煤及烹调油烟所致的室内空气污染是女性肺癌的主要原因。但是烧煤所造成的室内空气污染,在中国的存在已有数百年历史,而女性肺癌却是近三十年来才迅速增加。东西方国家室内空气污染情况大不相同,但女性肺癌的情况却极其相似。

室内放射性污染在西方被认为是肺癌的危险因素之一。但一般居室的放射性污染都很恒定,很难解释肺癌发病率的急速上升。某些职业接触与肺癌的关系比较明确,已有十多种工业毒物或职业过程被确定为与肺癌有关^[5]。这主要是由于病因明确,采用一般流行病学探讨危险度的方法即可确定。但是一般城市居民的肺癌职业归因危险度则很难估计。

肺癌是一种长潜隐期且多因素的疾病。在一些已知的因素中,其危险性的重要性对不同种族,性别的人群来说,已很难判断。至于一些未被了解的危险因素,情况当更复杂,加以肺癌的发病机理尚未完全清楚。这就提出了一个重要问题,应该用什么方法来探讨肺癌的病因,才能得出较明确的结论。

二、好的流行病学实践的重要原则:近年来许多流行病学家提出,用流行病学的方法探讨疾病的病因,必须遵循好的流行病学实践(Good Epidemiological Practices, GEPs)的原则^[6]。GEPs 的根本目的就是在于确保研究工作的质量。而且在研究过程中的各个环节,都必须给予很好考虑。现将 GEPs 的基本原则和内容概述于下:

1. 假说及目的:一个研究计划,首先必须明确研究的目的。对于所提出的理论依据及假说,应该给予充分的描述,并指出已被接受的观点,存在争论的观点和有待探讨的问题。

2. 研究对象:研究对象是根据研究目的来确定的。此时必须注意到样本的同质性及随机性。应该根据研究的性质确定样本的同质范围,并避免对样本进行主观的选择,这是避免混杂或干扰因素的主要措施。此外,应该强调的是,对照组的选择与观察组具有同等重要的意义,不可忽视。样本的大小及观察时间的长短,取决于疾病的性质、发病率的高低和潜隐期的长短。低发病率和长潜隐期疾病,必须有较大的样本例数及较长的追踪观察时间。调查中还应注意研究对象的意愿、信心,以及对其隐私的保密,这是取得可靠资料的重要因素。

1 广东省广州医学院预防医学教研室 510182

2 纽约医学院生物化学及分子生物学系

3. 资料收集:

(1) 调查表格的设计。调查表格设计总的要求是:目的明确、格式清晰、项目排列合理并有利于问题的分析和统计处理。描述性资料应予数量化。表格中所选用的项目完全取决于调查目的。主要内容不能有遗漏,无关内容不要列入,以免增加调查的困难。对调查所得的资料将用计数资料处理或计量资料处理,在调查表格设计时就应给予考虑。

被调查者在被调查时的合作情形,在调查表格中应有记录。

(2) 质量控制:要从流行病学研究中得出正确的结论,必须注意质量控制。质量控制就是排除和控制各种各样的偏误和混杂因素。以下是质量控制的一些原则:

①大数法则:事物的客观规律存在于大量的事实之中,换句话说,要有一定的数量才有质量,我们不能将偶然现象作为客观规律。因此调查时必须有足够的例数。一般来说可根据初步数据进行例数估计。

②同质性:同质是指事物性质相同,具体来说在事物(名词)前加上更多的形容词时,其同质性程度也愈高。在分类上同质性的样本才能归为一类。至于研究组与对照组之间除所欲比较的因素外,其他因素是否齐同,应进行均衡性检验。

③取样的随机性:在一个同质性群体中取样,必须遵循随机原则,一个选择性样本所造成的误差,往往很难用统计的方法加以校正。

④可比性评价:对资料进行可比性分析,是控制混杂偏误的重要方法,所以在对资料进行分析之前,应先进行可比性的探讨。

⑤标准化:标准化处理也是控制混杂因素的有用方法。这里包括率的标准化,联系程度的标准化,死亡比标准化,比例死亡比,以及进行配对分析等。

⑥偏误来源:总的来说,偏误有如下的来源:如询问偏误,回答偏误,选择性偏误,意愿偏误,错误分类偏误,回忆偏误,以及出版偏误等等,这些都必须仔细地给予控制和排除。

4. 环境暴露量测定:探讨与环境因素有关的疾病,必须进行暴露量的测定。目前常用的方法有三种,即问卷法,固定点采样法和佩带个人采样器法。其中佩带个人采样器法,最能正确反映暴露量的情况。暴露量的测定包括采样及实验室分析。在采样时应正确操作仪器。实验室分析必须尽可能排除“系统误差”及控制“偶然误差”,除了严格操作要求外,还可通过对数据的处理来对误差进行估计,并恰当地表示误差的程度。如方法的灵敏度(最小检出浓度),准确度(回收率)和精密度(变异系数)等。

环境因素暴露的确定,通常是先探讨两者有无相关关系,在确定有相关关系后,再进一步探讨两者是否因果关系,然后进而探讨是否存在剂量-效应关系。剂量-效应关系可有三种情况^[7]。第一是线性关系,第二是达到阈值后再呈线性关系,第三是某种因素在低浓度时可能有保护作用,当浓度高达一定程度后才有效应的线性关系。这些问题在分析时必须注意。

5. 环境暴露标记物测定:在环境流行病学中,环境暴露标记物的测定非常重要。它常用以表示不良暴露的程度,对疾病的易感性,或疾病的早期警告。标记物可分为三类^[8],即敏感性标记物,暴露性标记物和效应标记物。换言之,各种标记物不但反映了机体的敏感性,而且指示了暴露的程度和污染物对机体的作用机制。目前,危险因素流行病学正在日益受到来自流行病学本行及以外的批评的时候,使我们不得不想到在将来的研究中,必须更多地采用测定暴露标记物的方法,即分子流行病学的方法。

6. 数据处理:通过流行病学调查将会获得大量的数据,正确处理数据是取得正确结论的关键。不同性质的资料,其数据统计处理的方法各有不同,此处不予叙述。但在进行统计处理时,以下一些重要问题必须加以考虑:

(1) 数据的审核:流行病学调查所得的大量数据,在输入电脑进行分析之前,必须认真审核。对于明知不正确的数据当然应该放弃,对于可疑的数据在追查之后,仍无法确定数据是否正确时,宁可放弃,以免在正确数据中混入不可靠的数据。

(2) 对阳性结果的数据进行处理固然重要,对阴性结果也应探讨其原因。在追踪性研究中,对病例失访的原因,也应加以说明。

(3) 相对危险度的探讨是流行病学研究的重要目的。相对危险度可依 RR 值划分为高 (>10), 中 (2~9)

和低 (<2)三种情况^[9]。对 RR>2 的危险因素,通常较易为人们所接受,但对低危险度因素来说则要慎重考虑,此时除了分析危险因素本身的意义(即是否具有微弱的联系)之外,还应审查研究设计或研究方法是否合理和正确。

(4) P-trend 检验:对所得的相对危险度,是否具统计学意义可进行 P-trend 检验,例如对一组群体进行研究时,发现其相对危险度具统计学意义。但对群体进行分层分析时,暴露量的增加,相对危险度虽也有增加,但未达统计学意义,此时进行 P-trend 检验,就可以判明它们之间是否存在剂量-效应关系。

(5) Meta 分析^[10]:对某一疾病相似的流行病学研究,在世界各地可能都已进行过,要对各地研究进行比较、评价和综合,可以采用 Meta 分析。其主要作用是用统一评价的方法,解决单个研究结果之间的矛盾或未能明确的问题,改进对作用-效应的估计,并达到提高统计检验的效能。

(6) 合并分析(pooled analyses)^[11]:Meta 分析亦可用于低危险性因素分析,不过一个单独的低危险性因素分析资料,往往不会被接受,此时可采用合并分析。即研究者们预先协商好,用相似的研究方案、调查内容和调查方法,然后分别在各地单独进行,最后进行合并分析,这样往往能够得到最好的效果。

7. 数据的解释:对数据的解释是流行病学中最困难的部分。正如 Sir William Petty^[12]所指出:大部分的数据处理,只需四年级的数学水平,而对数据的解释则需要敏锐的观察力和丰富的知识。现在已有了能快速运算的电脑,流行病学家和统计学家也设计了许多分析资料的软件可供使用,但问题是这些软件的设计有无结合或建立于发病机理的基础之上,或者只是一种数学上的“游戏”。数据分析必须坚持严肃的客观态度,不能受领导者、资助单位,以及社会上各种各样的影响和压力所左右。当然还应有勇于承认错误的态度。

疾病病因的探讨必须有流行病学资料、暴露量的资料以及生物监测的资料,并将三者结合和相互验证,然后进一步判断其是否能合理地解释发病机理。即结论必须符合生物学的合理性(biological plausibility)。

8. 研究者:一个规模较庞大的研究计划,不可能由少数几个人来完成,必须集合各方面有关专家的通力合作。因此使参加者对研究目标有统一的认识,非常重要。对调查方案的统一认识,方法上的统一,对调查表格上各种项目填写方法的统一,是避免偏性的重要措施之一。研究领导机关对研究者的学术水平、素质及职业道德进行审查,是确保研究质量的重要措施。

三、肺癌病因研究的今后方向:尽管人们对肺癌的病因已有所了解,但还远远没有解决。根据目前的情况,今后的研究应以什么为重点?

1. 揭开肺腺癌之谜:常见的肺癌细胞类型有四种:鳞癌,大、小细胞癌和腺癌。临床学家常将之归为小细胞癌和非小细胞癌(包括鳞癌、大小细胞癌和腺癌)两类。由于肺癌的细胞类型与病因有密切关系,因此流行病学家^[13]则将之分为 Kreberg I (包括鳞癌和大、小细胞癌)和 Kreberg II (包括腺癌和肺泡癌)。并认为前者与环境因素有关,后者则为机体的异原性(heterogeneous)。

吸烟是男、女肺癌的重要因素,但多数研究者认为吸烟与腺癌无关,绝大部分的不吸烟者肺癌是腺癌。以往的报告表明男性肺癌以鳞癌为多见,腺癌次之;女性则以腺癌为多见,鳞癌次之。但是近年来^[14]的研究则发现男、女肺腺癌的比例均在上升,原因何在,值得探讨。

肺癌的四种细胞类型,不但在形态上有所不同,其细胞来源,生理生化机能亦有所不同,特别是分子生物学标记物上的差别,尤为引人注意。目前已证明肺腺癌细胞比之上皮细胞有着更多的分泌功能,是否内分泌障碍与肺腺癌有关,需要进一步深入探讨。

2. 室内空气污染:生活的现代化,使人们的生活中每天都要接触种类日益繁多的化学品。但由于人们大部分时间是在室内活动,因此室内空气污染问题日益突出,加以这种作用又属低浓度、长期限接触,所以问题变得更为复杂。

在各种重要的室内空气污染物中,烧煤所生成的污染物和烹调油烟与肺癌的关系已有较多有力的证据,而环境香烟烟雾(ETS)暴露与肺癌的关系则存在很大的争论。之所以存在争论:

第一,目前有关 ETS 与肺癌关系的研究,无论是接触史或暴露量的调查,大多用问卷法,对一个长潜隐期性疾病,许多偏误用问卷法是难于控制的。

第二,暴露量的测定即使是采用佩带个人采样器的方法也无法做到精确,而且监测指标的选择也不尽合理。例如已知 ETS 中至少含有 108 种化学物^[15],其中 4-Aminobiphenyl 和 2-Naphthylamine 为肯定致癌

物, Acetamine, BaP, Formaldehyde 和 N-nitrosodimethylamine 为极可能致癌物, 1,3-Butadine, Ideno[1,2,3-cd] pyrene, N-nitrosonornicotine 和 Nitrosopyrrolidine 为可能致癌物^[5]。目前只有有关 BaP 的报道, 其他化学物并未进行过研究。

第三, 目前还缺乏正确的生物标记物可供利用。虽然 D. Trichopoulos^[16]曾提出可用支气管上皮细胞癌前病变指征(EPPL)和 Reix index 来评价 ETS 暴露, 但作者所提供的数据却存在矛盾。

综上所述, 继续开展 ETS 与肺癌关系的研究, 实属必要。

3. 饮食/营养因素: 已有的报告指出, 多食用新鲜蔬菜、水果对肺癌有保护作用, 而高脂肪饮食, 特别是食物中富含饱和脂肪酸, 则为肺癌的危险因素之一。近来还有报道, 过多地额外补充维生素 A, 亦可增加肺癌的危险性。然而正如 Matanoski 所指出, 检测饮食的慢性作用, 必须对一个人的一生都进行调查, 而这是不可能的。流行病学家采用一个概略的方法, 即通过检测常用食物的消费频率来计算个人的暴露量, 但食物又会因季节和其他因素发生变化而影响结果。如果将食物转化为食物的成分, 如脂肪、维生素、矿物质和其他有害物质的成分, 这种转化又能产生新的误差。多种来源的误差, 将使弱影响联系的因素被掩盖或扩大, 从而得出对危险性的错误估计。可见我们必须对这类缺乏机制探讨的报告, 持谨慎的态度。最近有报道指出, 多食用新鲜蔬菜, 可以提高机体的抗氧化能力。这提示我们必须从发病机制上探讨饮食/营养因素与肺癌的关系。

4. 肺癌的分子流行病学: 前面已经说过, 分子生物学的研究能为我们提供机体对环境暴露的敏感性, 暴露接触的程度, 以及污染物对机体的作用机制的信息, 因此它已受到广泛的注意, 并取得飞速的发展。以下的一些标记物曾用于肺癌的研究, 特别有些还是与不同细胞类型的肺癌进行了联系, 显然这是有利于探讨不同细胞类型肺癌的发病机制。

肺癌细胞类型	P53 过度表达 (%)	P16 缺失 (%)	myc 表达 (%)	Rb 反向表达 (%)	FHIT*
细胞未分型	36~64			65	
鳞 瘤	41~54	48~70	11		40
小细胞癌	71	11	35		80
大细胞癌	53	48~70			40
腺 癌	27~34	48~70	53		40

* FHIT : Fragile histidine triad, 表现为 RNA 转录异常

有关代谢酶方面, 已知 CYPIAI 是肺癌的敏感性标记物, GSTMI 是机体抗氧化系统的重要标记物, 在 CYPIAI 高诱导性和 GSTMI 缺失者, 患肺癌的危险性大为增高。在人群中开展肺癌分子流行病学的研究, 目前的报道还不多, 这有待于今后的加强。

参 考 文 献

- 1 Alavanja MCR. 肺癌危险因素中的种族、家族和地理学的异质性. 肿瘤, 1996, 4: 450~456.
- 2 中华人民共和国卫生部. 全国卫生统计年报资料, 1982~1988.
- 3 Du YX, Cha Q, Chen XW, et al. An epidemiological study of risk factors for lung cancer in Guangzhou, China, Lung Cancer, 1996, 14, Suppl 1: S9~S37.
- 4 Chen BH, Hong CJ, He HZ. Indoor air pollution and its health effects in China —— A review.
- 5 IARC monograph on the evaluation of carcinogenic risk to human, 1987, Suppl 7: 17~74.
- 6 World Health Organization. Guidelines on studies in environmental epidemiology. WHO Environmental Health Criteria, 1983: 27.
- 7 Rylander R. 低危险度流行病学和好的流行病学研究方法. 肿瘤, 1996, 16: 475~478.
- 8 National Research Council Committee on Biological Marker. Biological markers in environmental health research. Environ Health Perspective, 74: 3~9.
- 9 Wynder EL. Guidelines to the epidemiology of weak association. Prev Med. 1987, 16: 211.
- 10 Shapiro S. Meta analysis/ Schmeta analysis. American J of Epidemiology, 1994, 140: 772~791.
- 11 Ahlbom A. Pooling epidemiologic studies. Epidemiology, 1993, 4: 283~284.

- 12 Monson RR. Occupational Epidemiology, CRC Press, Inc. 1980.
- 13 Kreyberg L. Histological lung cancer types, A morphological and biological correlation. Acta Pathol Microbiol Scand, 1962, Suppl 157:1.
- 14 Charloux A, Quoix E, Wolkove N, et al. The increasing incidence of lung adenocarcinoma: Reality or artefact? A review of the epidemiology of lung adenocarcinoma. International J of epidemiology, 1997, 26:14-23.
- 15 Eatough DJ, et al. The Chemical characterization of environmental tobacco smoke, Proceeding of Internatic al Symposium at McGill University, Canada, 1989, 3-39.
- 16 Trichopoulos D, et al. Active and passive smoking and pathological indicatr of lung cancer risk in an autopsy study. JAMA, 1992, 268:1697-1701.

(收稿:1997-08-17 修回:1997-08-25)

·读者福音·

本刊举办第五期“高级现代流行病学培训班”的通知

近年来,流行病学发展很快,现职卫生人员流行病学知识亟待更新与补充,本刊举办四期“高级现代流行病学培训班”后,许多读者要求再办此班,学员反映学习内容新颖,学到最新的理论和方法,深深获益,是全面、系统提高专业水平的极好机会;为了满足广大卫生人员现代流行病学知识水平与学科发展同步进行的迫切要求,本刊拟举办第五期,也是最后一期“高级现代流行病学培训班”。

一、时间:1998年4~7月,共4个月。

二、方法:该班以函授为主,聘请我国知名的流行病学专家章扬熙教授等专家任教。采取做习题、发考卷、给标准答案等方法开展教学活动。

三、学习内容:培训班采用专家新编教材,内容新颖,并以实例说明,学以致用。主要内容包括:1. 调查设计与实验设计;2. 常用的流行病学分析方法;3. 分析流行病学研究;4. 实验流行病学研究;5. 防治措施的效果评价;6. 暴发疫情的调查与分析;7. 血清流行病学;8. 寿命表法;9. 遗传度分析;10. 疾病预测方法(包括几十种预测方法,并有实例);11. logistic 模型;12. cox 模型;13. 价值服务工程;14. 目标规划方法;15. 综合评价方法;16. DALY 及其应用;17. 卫生决策方法;18. 现代临床流行病学(诊断、预后与疗效研究方法等)。培训班适合各级卫生防疫、妇幼保健、医院、卫生院、医学院校及科研单位的预防、临床、教学、科研人员和卫生行政官员及有关管理人员参加,非学员一律不售资料,请勿寄款。

四、考核及结业:采用发、收考卷的方式进行考试。及格者发结业证书,并通过本刊公布优秀学员名单,供有关单位和部门使用干部和晋升时参考。

五、报名及学费:报名日期为 1997 年 10 月 10 日 ~ 1998 年 3 月 31 日。请用楷体写明姓名、性别、年龄、职称、单位、详细通讯地址及邮政编码。报名同时邮寄学费(含资料费)180 元(开收据、报销)。一律寄至北京昌平流字五号《中华流行病学杂志》编辑部尹廉收(邮政编码 102206)。款到陆续寄资料,本期名额有限,按报名顺序录取,未被录取者一律退款。