

多维网状结构 :流行病学病因研究的一种思路与方法

杨廷忠 Brand Brosbrook

1. 构想依据 近年来国内学者对恶性肿瘤和心血管病等慢性疾病病因的流行病学研究已经做了不少的探索和卓有成效的工作,对这些疾病的控制起到了积极的作用。对于疾病的病因已有多种认可的解释模型:三角模式、轮状模式、树状模式和网状模式等^[1]。对恶性肿瘤和心血管病等慢性疾病而言,其中涉及社会、行为和心理因素且具有举足轻重的作用,各种因素不但与疾病发生联系而且因素之间也相互关联和相互衔接,实际上疾病是多种因素相互作用而导致的结果。所以用网状模型描述这些疾病的发生与存在比较符合实际情况。我们认为病因因素与疾病的联系呈多维网状结构。这个结构是一个整体。在这个结构中,因素与因素、因素与疾病间相互关联、前后衔接;在这个结构中,既有单个因素的连接点又有多个因素的交汇点;在这个结构中,疾病可以不是终点。就病因因素而言,它们不但具有直接作用,而且还具有间接作用;不但具有主效应作用,而且还具有反馈作用,特别是可能还有支配若干因素的潜在因子存在。

图1显示遗传、个性、社会因素、心理压力、行为和冠心病间呈现出多维网状结构的关系。各种因素在多个维度上相互关联。值得强调的是模型中5个行为因素后面有一个潜在因子在支配着,这个潜在因子就是心理压力,它是这些行为因素的根,是十分重要的行为源病因因素。只要将此得以控制,这些行为因素就有可能在某些程度上得以控制,这对于疾病的预防具有非常重要的作用。从疾病的预防效果而言,干预措施应指向某些因素的系统根基。多维网状结构模型是一种整体思维方法,而目前流行病学危险因素分析的方法则将每个因素与疾病的关系孤立地看待,这样各种因素对致病作用的位点和性质就难以识别,对实施干预带来了很大困难。同时只重视末梢而忽略根基的思维方法决定了干预效果的低效性。

2. 实际操作:多维网状构架不但具有解释实际问题的合理性,而且具有可操作性。在病因研究中往往涉及到一些社会、行为、心理变量,这些变量的测量和测量误差估计是一个棘手的问题。因为它们大多是一些抽象的概念,往往变异性大,测量难以把握,甚至被认为这种研究缺乏可靠性。其中很多变量本身无法直接被观察,需要用其他的多重测量来表

达。目前流行病学研究对于这一问题的处理办法通常有两种,一是使用一些现成的测量工具,但现有的测量工具十分有限,难以适应对于各种疾病的研究,更不能表述研究者活跃的思维,使其只局限于分析若干个变量与某些疾病之间的关系;二是直接用某一测量项目表达某一概念,但这种做法的误差非常大。如饮酒行为应由饮酒的频率、量和种类多重属性来决定,显然任何单一的测量难以准确反映饮酒行为的概念。行为的测量应建立在概念的准确把握和严格的实施上。将多重变量测量的理念和操作纳入到多维网状结构模型中去可以较好地解决这一问题。不但应考虑如何应用若干个测量项目准确地表达某一变量的属性,而且应将这部分内容作为整个结构的有机部分进行整体思考和判断。至于最终测量效果,应用结构方程模型还可以做出估计。

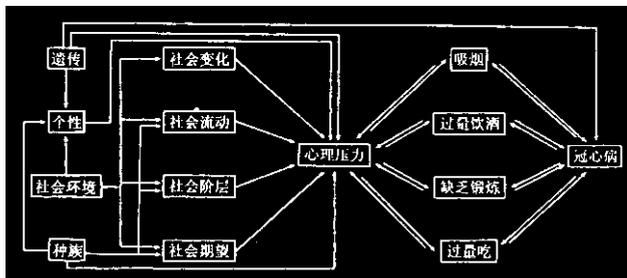


图1 遗传、个性、社会因素、心理压力和行为与冠心病间的多维网状结构关系

分析和检验多维网状结构可使用结构方程模型统计方法。该方法由两部分构成,第一部分为测量模型,说明潜变量和测量变量间的关系。用两个方程来表示: $y(p \times 1) = \Lambda_y(P \times m) * \eta(m \times 1) + \epsilon(p \times 1)$, $x(q \times 1) = \Lambda_x(q \times n) * \xi(n \times 1) + \delta(q \times 1)$ 。式中 y 和 x 分别为内生测量变量和外生测量变量所构成的向量($p \times 1$)和($q \times 1$); Λ_y 和 Λ_x 分别为 y 对 η 和 x 对 ξ 的因子负荷矩阵($P \times m$)和($q \times n$); η 为内生潜变量所构成的向量($m \times 1$); ξ 为外生潜变量所构成的向量($n \times 1$); ϵ 和 δ 分别为 y 和 x 的测量误差所构成的向量($p \times 1$)和($q \times 1$),它们被假定与 ξ 和 ζ 不相关并且它们之间也互不相关。第二部分为潜变量模型,表示外生潜变量和内生潜变量之间的因果关系。其形式为: $\eta(m \times 1) = B(m \times m) * \eta(m \times 1) + \Gamma(m \times n) * \xi(n \times 1) + \zeta(m \times 1)$ 。式中 B 为内生潜变量对内生潜变量 $\{\eta\}$ 效应系数的 $m \times m$ 矩阵, Γ 为外生潜变量(ξ)对内生潜变量 $\{\eta\}$ 效应系数的 $m \times n$ 矩阵,最后一

作者单位 310031 杭州,浙江大学医学院社会医学卫生事业管理教研室(杨廷忠);Cancer Center, University of California, San Diego, USA (Brand Brosbrook)

项为残差项构成的向量 ($m \times 1$), 它被假定接近于 0 并与 ξ 不相关。模型的拟合优度检验使用如下指标: GFI 为拟合优度指数, 表示由模型能够解释的变异和共变异, 其值在 0 和 1 之间, 在 0.9 以上时表示方程拟合良好。AGFI 为考虑了自由度后的 GFI, 其值通常 > 0.8 时表明拟合良好。RMR 为均方根残差, 表示平均拟合度的残差测量, 其值 < 0.1 时表明拟合良好。似然比 χ^2 检验无显著性表明观察和预测的共变异矩阵差别不明显, 方程拟合良好; 值得注意的是, 在大样本条件下即使其他指标拟合情况很好, 很小的差别会使 χ^2 值变得显著, 故对该指标的要求不必很严格。参数评价使用 t 检验法, 群内比较使用标准参数, 群间比较使用非标准参数^[2,3]。结构方程模型运算可以在 SPSS 统计软件 LISREL 程序上进行或在 SAS 软件 CALIS 模块的 EQS 程序上进行。EQS 编程首先指明 CALIS 模块, 再在 Ram 语句中列出各变量(标号)矩阵, 结果变量在前, 原因变量在后; 第一层次运算负荷值和路径参数, 第二层次运算误差参数^[4]。

3. 应用实例: 应用对江苏省铜山县 4 所中学 732 名高中男生吸烟和饮酒行为为期 4 个月的干预资料^[5]。在此我们根据行为学的理论假设了心理控制源、吸烟、饮酒和口的活动与情绪失调关系的网状结构模型。人口学特征、家庭背景变量作为混杂变量来设计。吸烟以每周吸的支数来报告, 按 < 1 支、1 支、2 支和 > 2 支编码。饮酒行为以每月饮的有效次数来报告, 按 < 1 、1~2、3~4 和 > 4 个有效次编码; 口的活动用吃零食和咬异物两项测量来表达; 心理控制源变量纳入分析的为内源性(I)、有势力的他人(P)和机遇(C)量表得分, 情绪状态纳入分析的为情绪失调评分(TMD)。以上变量和项目的测量见文献[5]。吸烟和饮酒状态在干预前后测量, 其他变量在干预后测量。纳入分析的变量均为干预未测得值。统计分析使用 SAS 软件上 CALIS 模块上 EQS 程序进行^[4]。分两步进行, 第一步分析吃零食和咬异物测量对口的活动的表达。模型拟合满意(GFI: 0.983 6, AGFI: 0.973 4, $\chi^2 = 7.76, P > 0.05$), 两因子负荷值检验均在 0.01 水平上显著。第二步分析外源变量和内源变量之间的关系。方程拟合满意(GFI 0.991 1; AGFI 0.975 6, $\chi^2 = 30.69, P > 0.05$), 结果见图 2。各变量平均数和标准差分别为: I: 28.80, 6.45; P:

22.86, 7.08; C: 19.30, 8.82; TMD: 49.68, 31.03; 吸烟: 1.21, 0.64; 饮酒: 1.52, 0.84; 吃零食: 0.37, 0.48; 咬异物: 0.27, 0.44。

图 2 只显示了主效应模型, 各种关系的相反作用加入模型后, 方程的拟合不理想, 有待于进一步探索。在这里结构方程模型表述了若干个变量与情绪失调之间较复杂的关系, 显示了各路径间的主效应方向。心理控制源直接或通过吸烟和饮酒对情绪失调产生作用, 吸烟和饮酒引起情绪失调, 而后者又可引起口的活动。由此可见, 在这几种行为中, 吸烟和饮酒是情绪失调的因, 而咬异物和吃零食是情绪失调的果, 可以理解为机体对情绪失调的行为调节方式。如将观察期间吸烟和饮酒启动或戒除变量指标代替图 2 中的吸烟和饮酒变量指标(两种状态的分类见文献[5]), 这种关系也成立。如果使用目前在病因研究中普遍应用的流行病学危险因素研究的方法探讨以上关系, 只会获得心理控制源、吸烟、饮酒、咬异物、吃零食等因素与情绪失调关联的信息, 而联系的内在机制则无法得知。

4. 评述: 多维网状病因研究模型应该建立在符合实际的理论构思和严格的变量测量之上。全面了解所要研究疾病的研究现状, 深刻领会相关的理论, 深思熟虑的思考和论证, 使研究更为严密并富有理论与实际意义。如果无任何理论依据和实际工作的基础就直接建构模型, 这种模型除了提供统计学相关的意义外, 别无任何帮助。同样没有准确的数据为基础, 就不会产生较理想的结构模型。使用结构模型的优势在于对于潜在变量的认识和评价, 这对于慢性病的病因研究颇为重要。显然, 得到支配与疾病有关联的若干因素的根要比单纯寻求各因素与疾病的关联更为重要。对于病因潜在变量的统计分析首先要看若干病因因素之间有无关联, 然后再分析这种联系能否被一个共同因子所解释及这一个因子是否导致疾病的发生。由于这一过程是一种验证的方法, 而不是探索的手段, 故这一方法的使用应该在其他流行病学病因研究方法的基础之上进行。既然称这种病因研究策略是多维网状结构, 结构中所涉及的要素间不但呈上下左右空间关系且呈前后的时间关系, 故这一方法所要求的应为纵向观察研究, 作为适用这一病因研究策略分析的结构方程模型

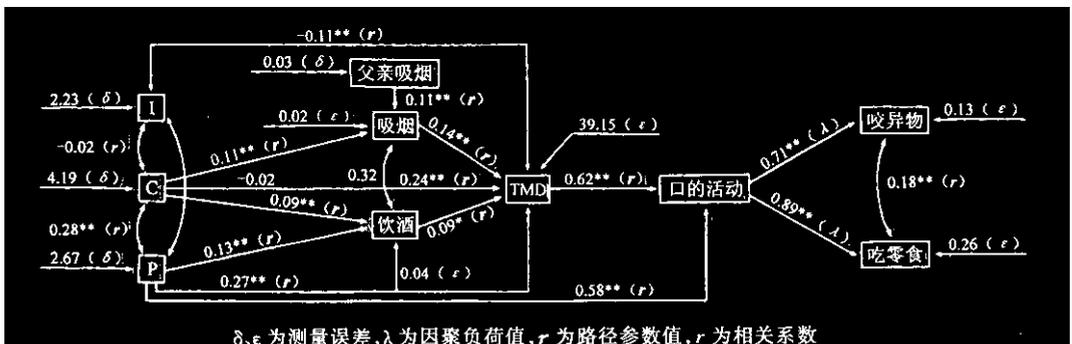


图2 根据实例模拟的心理控制源、吸烟、饮酒和口的活动与情绪失调关系的网状结构模型

初始的设计也是针对这类资料的。当然有些横向资料在逻辑上具有纵向资料的性质, 如, 父母吸烟与子女吸烟的关系,

故没有必要刻板地看待这一问题。通过横断面资料依据有关变量逻辑上的因果关系检验建立的模型,也许会为病因研究提供某些线索和为进一步的纵向研究打下基础。多维网状结构是一种因果关系模型。在这个模型中,有些变量只影响其他变量,而不受另一些变量的影响,常见的是一些人口学和社会变量,而有些变量则处于以上两种情况,常见的是一些心理、行为变量,过去我们常将疾病作为只受其他变量影响的变量来处理,实际上很多情况下疾病对其他变量有反向作用,正如图1所示;明确这一点对于认识疾病发生的机制大有裨益。

结构方程模型最初的设计主要针对连续变量的分析,特别是结构中的出口变量应为连续变量。但在流行病学病因研究中常常为分类变量。现已发展出适合于分类资料的结构分析模型,其基本思路是对分类变量Y进行标准化处理^[6]。同时可以采取变通的办法,如果出口变量为分类变量,可以在此变量上连接某个或某些连续变量,使其变为中间变量。也可采取分步分析策略,但这样做的缺陷在于可能使整个模型的拟合优度有所下降。所要指出的是结构方程模型分析能够用若干个指标对模型的拟合优度做出估计,这种估计是表示模型对数据总拟合度测量的统计学意义,而不

能用来判断模型的质量和模型的理论及实际意义。模型能对测定误差做出估计当然很好,但关键还是实际测量中的控制。最后要强调的是应用结构方程模型得出的结论只能用来排除不能用以肯定。

参 考 文 献

- 1 刘瑞璋. 病因概念及病因推断. 见: 耿贯一, 主编. 流行病学. 第2版. 北京: 人民卫生出版社, 1995. 46-52.
- 2 Bollen KA. Structural equations with latent variables. Toronto: John Wiley & Sons, 1989. 318-403.
- 3 Fuzhong L. Approaches to testing interaction effects using structural equation modeling methodology. Multivariate Behavioral Research, 1998, 33: 1-39.
- 4 Cary NC. SAS/STAT users guide. SAS Institute Inc, 1998. 370-432.
- 5 杨廷忠, 李鲁, 朱书宏. 高中学生吸烟、饮酒及吸烟饮酒并存行为与情绪失调关系的研究. 中华流行病学杂志, 2001, 22: 397-398.
- 6 Muthen BO. A general structural equation model with dichotomous, ordered categorical, and continuous latent variable indicators. Psychometrika, 1984, 49: 115-132.

(收稿日期 2001-12-11)

(本文编辑: 张林东)

· 短篇报道 ·

黄金海岸沙门菌引起食物中毒的调查

孙凤琪 计国欣

1998年7月,从广州市海珠区某制衣厂食物中毒病人粪便中分离出14株黄金海岸沙门菌,此菌型在国内首次检出。

流行病学调查:7月14日下午6时,该厂有35人在厂内擅自开设的食堂就餐,当晚10时就餐者陆续出现呕吐、腹胀,次日病情加重,继出现腹泻、高热、肌肉疼痛等症状。先后有30例中毒病人住院治疗,占就餐人数的85.7%。病人中除有白细胞、中性粒细胞升高外,还出现脱水、电解质失衡。经抗感染及对症治疗1~3d全部治愈出院。

细菌分离鉴定:按国家标准1985年《食品卫生检验方法》(微生物学部分)进行,同时挑取菌落在营养琼脂平板上

作沙门菌O-1噬菌体裂解试验,经37℃培养4~6h,被O-1沙门菌属噬菌体裂解。该菌株为革兰阴性杆菌,有动力,在SS琼脂平板上菌落呈圆形,外观光滑、湿润、无色半透明,菌落中心呈黑色,中等大小。生化反应符合沙门菌。

血清学鉴定:164种沙门菌属诊断血清由卫生部兰州生物制品研究所生产(有效期内使用)。采用玻片凝集试验,沙门菌属A~F群“O”多价诊断血清凝集(+++),生理盐水对照阴性。“O”因子血清6 δ 凝集(++);“H”因子血清第一相 γ 凝集(++),第二相I,V血清凝集(++),V阴性,W血清凝集(+++)。其抗原式为6 δ : γ :1,W。该菌株经中国药品生物制品检定所复核证实为黄金海岸沙门菌,检定菌号3526。

(收稿日期 2002-04-12)

(本文编辑: 段江娟)

作者单位:510288 广州市海珠区卫生防疫站微生物检验科(孙凤琪);中国药品生物制品检定所(计国欣)