

伤害流行病学研究的若干方法学问题

孙业桓 余德新 黄子惠

伤害作为一个重要的公共卫生问题,已引起广泛的关注。近年伤害流行病学研究在国内已得到广泛开展,但也应看到在伤害流行病学研究中存在诸多方法学问题。由于伤害(事故)具有可反复多次发生以及伤害类型、发生伤害的原因、发生地点、就诊场所等的多样性的特点,因而在进行伤害流行病学研究时应对伤害研究的复杂性有充分的认识,简单地将传统的流行病学研究方法应用在伤害研究领域会出现诸多问题并可能导致错误。本文仅就有关伤害流行病学研究中涉及的部分方法学问题作一讨论。

1. 关于伤害的定义:所谓伤害(injury)是指机械能、热能、化学能、电能和放射线等能量在传递到机体组织并与之进行能量交换的过程中,在机体组织无法耐受的水平上所造成的组织损伤,或是由于窒息引起缺氧称之为伤害。然而,从病理生理机制上所定义的伤害,无法满足流行病学研究的需要,如进行高血压的流行病学研究,第一步就是要定义血压达到什么水平视为异常,即首先要解决的问题是“病例”的诊断标准问题。就伤害研究而言,即是在何种情况下视其为发生了一次伤害,并被作为一个“统计病例”而纳入研究;在伤害研究领域一般称之为伤害的标准化定义(standardized definition),也有人称之为伤害的操作性定义(operational definition of injury)。

对伤害的界定,迄今国内外尚没有统一的标准。根据研究的目的,大致可从三个范畴定义伤害,分别以伤害发生后导致死亡、住院或出现其他情况(比如因伤害导致一定时间的活动受限、停工、休学等)作为判断依据。一般情况下,以因伤害致死亡或住院作为依据定义伤害,标准易于掌握,更易操作,但会造成低估伤害的问题,即大量的伤害并不需要住院治疗,致死性伤害更少。因此,根据不同的目的,不同的研究者可能会使用不同的操作性定义。由此,笔者认为应强调不同地点、不同时间、不同研究者的报告在进行比较时应注意可比性问题,翻阅有关伤害的文献,出现对使用不同定义的研究相互比较的报告并不少见。

近几年中国大陆开展的伤害研究基本上采用了由王声^[1]1998年提出的建议标准。凡具有下列三种情况中任何一项,即作为伤害的统计对象:①到医疗单位诊治,诊断为某一种损伤;②由家人、老师、同事或同伴对受伤者作紧急处置或看护;③因伤请假(停工、休学、休息)半日以上。而金会庆等^[2,3]对儿童非致死性意外伤害的流行病学研究采用的定义

为意外伤害发生后造成一定的身体损伤,父母或其他人曾带儿童看过医生,未看医生者不列入研究范畴。美国国立卫生统计中心1997年开展的全国性健康调查(NHIS)所拟定的伤害定义是“所谓伤害必须导致伤者寻求医疗帮助”^[4]。然而由于不同国家、不同地区有不同的医疗体制,因此,采用此定义对于发展中国家或欠发达地区,显然会低估伤害的发生情况。Junkins等^[5]对学校伤害的研究定义是,凡具有下列中一项,即作为伤害的统计对象:①至少导致缺课半天;②曾接受医务人员的治疗。Petridou等^[6]在一项干预性研究中采用的伤害定义为,凡具有下列中一项,即列为伤害的统计对象:①曾接受医务人员的治疗;②导致至少持续2h的不适。鉴于伤害研究的操作性定义存在较大差异,为此我们建议,有关的专业学会应发挥学科优势,根据研究目的拟定相应的具指导作用的操作性定义,既要考虑到中国的具体情况,同时应注意与国际上较为广泛采用的定义之间的有机整合,或可在今后的研究实践中以不同的人群为研究对象,探讨不同定义之一致程度,以便在相互比较时能以一合理的系数校正。

2. 关于伤害发生情况的频率测量:测量(measure)伤害发生的频率是进行伤害研究的基础,以死亡作为结局的伤害死亡率指标与一般的疾病流行病学研究并无区别,然而对于在流行病学最常用的一些发生率测量指标(包括其称谓问题)当应用到伤害研究时则有诸多值得商榷之处。

在人民卫生出版社出版的预防医学专业统编教材《流行病学》第4版的“伤害流行病学”一章,对“伤害发生率”的描述是“单位时间内(通常是年)伤害发生的人数与同期人口数之比,是进行伤害研究与监测的常用指标。但伤害发生率远不如发病率计算简单,而是会出现很多种情况的分母。……”按此描述分子为“受伤人数”,然而,在伤害流行病学研究中经常遇到的情况远非如此简单,一个人在一年中可能会发生2次或2次以上的伤害,如何计数?由谭红专主编的《现代流行病学》一书,除上述类似的描述外,同时给出用发生伤害的人次数作分子的计算公式,但仍然称之为“伤害发生率”。此相同的称谓势必会给读者、决策者和公众错误的信息。因此,我们建议以不同的分子计算的“伤害发生率”应给予不同的称谓,以人次数作分子的可称之为“伤害事件发生率(event-based rate of injury),从公共卫生的观点出发,伤害事件发生率更具实际意义。Menckel, Laflamme^[7]则将“伤害事件发生率”称之为伤害比(injury ratio)。

伤害事件发生率是一个有单位的发生率指标,单位由分母的性质决定。从概念上来讲,它是“比”。由于其分子是事

件计数资料,而且相对于每个人来讲可能会发生多次,因此其可信区间的计算以及组间比较不同于一般概念的“发病率”,有人建议应采用 Poisson 方法^[8]。虽然 Poisson 方法仍被用于此类重复发生事件的描述和比较,但其存在的问题和缺点早在 1920 年 Greenwood 和 Yule 就已进行了描述, Glynn Buring^[9]综述了相关的文献后认为,采用负二项分布的方法更能真实地描述和分析此类资料。

不同人群伤害事件发生率的比较,除应考虑前述的伤害“诊断”标准统一问题外,还应考虑到因混杂因子存在,直接比较会导致得出错误的结论。处理混杂因子存在时组间比较的常用策略——标准化法,同样可应用于伤害事件率的比较,并且已有此类事件计数资料事件发生率标准化率计算的报道^[10,11]。

3. 关于伤害发生影响因素的效应测量:效应测量指标比数比(odds ratio, OR)及率比(rate ratio)当应用于伤害流行病学研究时,应重新审视。Comming 等^[12]以老年跌落伤为例对其做了系统评述,每年 65~ 岁老人大约有 30% 发生跌伤,此较高的伤害发生率并不影响暴露组与非暴露组跌伤率率的计算,但如果将老年人群跌伤者与非跌伤者进行比较(病例对照研究)则高比例的跌伤者会对效应测量指标选择产生重要影响。因为跌伤者并非罕见,显然如此计算的比数比并不能作为 RR 的较好的估计值。Tinetti 等^[13]的报道,清楚地表明了这一点,即一个 336 人的老年人群队列与镇静剂使用有关的跌伤粗相对危险度(RR)为 3.1,而采用比数比这一指标计算时所得 OR 值为 31.1,其间差异是因为采用不同的效应测量指标的缘故。该研究报告两种粗效应测量指标(RR 与 OR)如此大的差异原因是暴露于镇静剂组跌伤者比例非常高(一年中至少发生一次跌落的比例达 93%,非暴露组也达到 30%)报告中当采用 logistic 回归分析在校正混杂因子后校正的比数比仍达 28.3。Comming 等^[12]提出了几种变通的解决办法,比如可以定义发生跌落 2 次或以上者作为“病例”,这样“频率”减少,再有,就是报告未校正的和校正的所有比数比的计算结果,但由此会使读者对结果评价产生理解和解释上的困惑。

正像我们已讨论的那样,对伤害研究,研究者更关心的是伤害事件、伤害事件率,然而可应用于反复发作事件(recurrent events)的统计分析方法^[9,14],在伤害研究领域并未得以较多合理的应用,而在中国的伤害研究报道迄今未见应用此类方法。此类方法涉及多水平模型(multi-level model)广义估计方程(generalized estimating equation)及 Poisson 回归、负二项回归等指数分布族资料的统计学方法等。流行病学研究较常应用的一般 logistic 回归分析方法对于单位时间或单位空间内某事件发生数(count data)资料的影响因素分析,并不是优选方法,且在伤害(或各类事故)研究中,常无法满足 logistic 回归分析所要求的各观察对象间相互独立的应用条件,此在伤害研究领域应引起足够注意。近年随着多水平统计分析模型研究的发展,人们已认识到对具

有层次结构的资料,如采用传统的统计分析方法所建模型预测精度不高,且会得出错误的标准误、可信区间估计及不正确的统计检验结果^[15],在伤害(或事故)领域的研究常会忽视伤害研究资料具有多水平的特点,此也应引起注意。

4. 关于伤害信息收集过程的回忆偏倚:《中华流行病学杂志》曾有专文介绍了伤害信息收集中的几个问题^[16]。就非致死性的伤害流行病学研究而言,有关伤害的调查常常建立在回顾性自我报告的基础之上,此类调查信息无论在发达国家还是在发展中国家均是对常规生命统计资料的有效补充,对于常规资料不健全的发展中国家更有实际意义。但在应用此类回顾过去发生伤害事件的回忆信息资料时,应充分认识到其存在的偏倚问题。Harel 等^[17]对 0~17 岁儿童的研究显示,随回忆时间的延长,报告的伤害率会有显著下降,从 1 个月回忆期估算的 24.4% 的年伤害率下降到 12 个月回忆期的 14.7% 的年伤害率,其中尤以 0~4 岁年龄组和微伤的下降最为显著,而对于较严重伤害比如导致缺课、卧床休息、外科包扎或手术、住院的伤害则回忆偏倚较小。Mock 等^[18]在加纳对一般人群的研究得到类似的结果,报告的伤害率从 1 个月回忆期的 27.6% 的年伤害率下降到 12 个月回忆期的 7.6% 的年伤害发生率,因此建议欲采用回顾调查方式估算总的非致死性伤害发生率的回忆期不能跨度太长,以 1~3 个月为宜。在中国有关回忆偏倚对调查结果影响大小的研究未见报道,对于具有不同文化背景的国人来说回忆偏倚的影响大小有待研究。一年回忆期调查所获结果会低估伤害的危害,且在描述伤害发生月分布时,也会因回忆偏倚而致难以真实描述一年四季的分布特征。

5. 关于伤害流行病学研究设计类型:有关伤害流行病学专著或一般流行病学著作中“伤害流行病学”篇章多是套用一般流行病学研究设计分类体系,但当应用于伤害研究时,或许会发现某些情形下的研究无法套用(或套用不合理)某种分类类型。比如,在调查时测量的是已存在的暴露,而伤害结局则是按定义回顾过去一年是否发生和发生了几次,这显然不是真正意义的横断面研究,因为伤害与否是通过回顾得知的,也不能算真正意义上的病例对照研究,因为抽样并不是建立在结局存在与否的基础上,因此,Comming 等^[12]建议与其试图把某一研究归类,倒不如在研究报告中真实描述其研究设计和过程。

6. 结语:开展伤害流行病学研究应清醒的认识到其特殊性。应明确定义结局变量——伤害事件,注意控制和估计偏倚,正确使用统计学分析方法,审慎的解说研究结果。伤害流行病学工作者应进一步探讨传统流行病学方法应用的合理性,不断发展伤害流行病学理论,重视多学科的合作研究,充分借鉴国外的研究成果,开展有一定深度和规模的具中国国情特点的伤害流行病学研究,以满足我国 21 世纪伤害预防和控制工作所提出的更高层次的要求。

参 考 文 献

1998 4:299.

- 2 金会庆,张树林,陈朴,等. 儿童非致死性意外伤害的流行病学研究. 人类工效学,1999 5:1-5.
- 3 金会庆,孙业桓,张树林,等. 事故倾向性儿童相关生理心理特征的 logistic 回归分析. 中国学校卫生,1999 20:53-54.
- 4 Warner M, Barnes PM, Fingerhut LA. Injury and poisoning episodes and conditions: National Health Interview Survey, 1997. Vital & Health Statistics-Series 10 :Data from the National Health Survey, 2000. 1-38.
- 5 Junkins EP Jr, Knight S, Lightfoot AC, et al. Epidemiology of school injuries in Utah: a population-based study. J School Health, 1999 69:409-412.
- 6 Petridou E, Tolma E, Dessypris N, et al. A controlled evaluation of a community injury prevention project in two Greek islands. Inter J Epidemiol, 1997 26:173-179.
- 7 Menckel E, Laflamme L. Injuries to boys and girls in Swedish schools: different activities, different results? Scand J Publ Health, 2000 28:132-136.
- 8 Rivara FP, Cummings P, Koepsell TD, et al. Injury control: a guide to research and program evaluation. Cambridge University Press, UK. 2001.
- 9 Glynn RJ, Buring JE. Ways of measuring rates of recurrent events. British Med J, 1996 312:364-367.
- 10 Stukel TA, Glynn RJ, Fisher ES, et al. Standardized rates of recurrent outcomes. Stati Med, 1994 13:1781-1791.
- 11 Glynn RJ, Stukel TA, Sharp SM, et al. Estimating the variance of standardized rates of recurrent events, with application to hospitalizations among the elderly in New England. Am J Epidemiol, 1993 137:776-786.
- 12 Comming RG, Kelsey JL, Nevitt MC. Methodologic issues in the study of frequent and recurrent health problems. Falls in the elderly. Ann Epidemiol, 1990 1:49-56.
- 13 Tinetti ME, Speechley M, Ginter SF. Risk factors for falls among elderly persons living in the community. N Engl J Med, 1988 319:1701-1707.
- 14 Gardner W, Mulvey EP, Shaw EC. Regression analyses of counts and rates: poisson, overdispersed poisson, and negative binomial models. Psychol Bull, 1995 118:392-404.
- 15 Skinner CJ, Holt D, Smith TMF. Analysis of complex survey. Wiley, Chichester UK. 1989.
- 16 汪宁. 伤害信息收集中的几个问题. 中华流行病学杂志, 2000, 21:372-374.
- 17 Harel Y, Overpeck MD, Jones DH, et al. The effects of recall on estimating annual nonfatal injury rates for children and adolescents. Am J Publ Health, 1994 84:599-605.
- 18 Mock C, Acheampong F, Adjei S, et al. The effect of recall on estimation of incidence rates for injury in Ghana. Inter J Epidemiol, 1999 28:750-755.

(收稿日期:2003-08-21)
(本文编辑:尹廉)

· 读者·作者·编者 ·

关于文稿申请“快速通道”发表的规定和要求

为了使反映我国流行病学领域中,有重大研究成果内容的论文尽快、及时在我刊发表,本刊自 2003 年起,对作者的来稿从审查到正式发表的程序,开辟了“快速通道”。为了使广大作者、读者了解文稿申请“快速通道”的规定和要求,特将申请“快速通道”的规定和要求公布如下(1)凡内容涉及有重大创新和为国内首创,达到或超过国内或国际先进水平的论文,均可申请进入“快速通道”,以使论文快速发表(2)作者本人提出进入“快速通道”申请(3)作者提供国内外数据库的查新报告(4)作者提供两位同行知名专家(作者所在单位的专家和作者的导师应回避)的推荐信,推荐信内容应包括本研究为“最新”、“首创”,及申请快速发表论文的理由(5)作者提供申请快速发表论文的作者署名无争议、发明权(即首创权)无争议的证明(6)作者提供论文一式 3 份(包括软盘)(7)作者提供由作者单位科研部门开具的介绍信。

凡符合上述规定和要求,获准进入“快速通道”的论文,将由本刊编委会总编辑、相关专业编委共同审议决定是否刊登(每篇论文需交纳快速审稿费 200 元,并请通过邮局汇款)如编委会审查后同意论文发表,本刊郑重承诺,该论文于收稿后 4 个月内正式刊出。

本刊编辑部