

长沙市农村中学生暴力遭遇影响因素 多元多水平分析

周立波 罗家有 方俊群 孙振球

【摘要】 目的 分析长沙市农村中学生暴力遭遇发生影响因素。方法 以过去一年是否遭遇情感暴力、躯体暴力、性暴力为多应变量,采用多元多水平 logistic 回归模型分析三种暴力类型间的相关关系及暴力遭遇的影响因素。结果 在 3620 名被调查者中,情感暴力、躯体暴力和性暴力的发生率分别为 21.5%、24.3% 和 2.0%。采用多元多水平模型将班级水平的随机效应分离出来后得到情感暴力与躯体暴力、情感暴力与性暴力,躯体暴力与性暴力之间的相关系数分别为 0.337、0.133、0.131;躯体暴力发生率在班级内存在聚集性($\chi^2=4.286, P=0.038$);情感暴力与性暴力在班级内存在关联性($\chi^2=4.239, P=0.039$);躯体暴力与性暴力在班级内存在关联性($\chi^2=4.482, P=0.034$)。暴力发生率的影响因素有性别、吸烟、父母吵架、欺负他人、自尊水平。结论 多元多水平模型可将班级水平的随机效应分离出来,使结果的估计更加准确;在采取有关措施降低中学生暴力遭遇发生率时,除应注意一些个人因素和家庭因素的影响外,班集体小环境所起的作用也不容忽视。

【关键词】 情感暴力; 躯体暴力; 性暴力; 农村中学生; 多元多水平模型

Using multivariate multilevel model in studying the influential factors of violence among rural high school students in Changsha ZHOU Li-bo*, LUO Jia-you, FANG Jun-qun, SUN Zhen-qiu. *School of Public Health, Central South University, Changsha 410008, China

Corresponding author: LUO Jia-you, Email: jiayouluo@hotmail.com

【Abstract】 **Objective** To provide evidence for setting up violence intervention programs in rural middle schools, through studying the influential factors. **Methods** Taking variables including emotional, physical and sexual violence in the past year as the multi-dependent variables before multivariate multilevel model logistic regression model was adopted to analyze the correlations among the three kinds of violence and the influential factors. **Results** Among 3620 respondents, the incidence rates of emotional, physical and sexual violence weres 21.5%, 24.3% and 2.0% respectively. The correlation coefficients between emotional violence vs. physical violence, emotional violence vs. sexual violence, physical violence vs. sexual violence were 0.337, 0.133, 0.131 respectively when the random effect of class difference was separated by multivariate multilevel model. There was an internal aggregation of the incidence rate on physical violence in different grades ($\chi^2=4.286, P=0.038$) and an internal relevant between emotional violence vs. sexual violence ($\chi^2=4.239, P=0.039$), physical violence vs. sexual violence ($\chi^2=4.482, P=0.034$). The influential factors on the incidence rates of violence would include: sex, smoking status, family without harmony, tendency of bullying others and the level on self-esteem etc. **Conclusion** When the random effect of class difference was separated by multivariate multilevel model, the estimated results would be more precise. Other than paying more attention to both individual and family influential factors when taking measures to reduce the incidence rate of violence in high school students, the effect of environment in the class should not be ignored.

【Key words】 Emotional violence; Physical violence; Sexual violence; Rural middle school students; Multivariate multilevel model

暴力是指蓄意地运用躯体的力量或权利,对自

身、他人、群体或社会进行威胁或伤害,造成或极有可能造成损伤、死亡、精神伤害、发育障碍或权利剥夺的行为^[1]。青少年最容易受到暴力的影响^[2]。近年来,中学生暴力事件的发生率呈上升趋势,严重影响了学生的学习、生活以及身心健康。青少年除了容易成为受害人外,往往也是暴力行为的实施者。中学生中最常见的暴力方式是情感暴力、躯体暴力、

DOI: 10.3760/cma.j.issn.0254-6450.2009.04.006

作者单位: 410008 长沙,中南大学公共卫生学院(周立波、罗家有、孙振球);湖南省妇幼保健院(方俊群)

第一作者现工作单位: 511515 广东省清远县出入境检验检疫局卫生检疫科

通信作者: 罗家有, Email: jiayouluo@hotmail.com

性暴力,且三者具有一定的相关性^[3]。本研究旨在探讨农村中学生三种暴力的发生率及其影响因素,为中学生暴力干预提供科学依据。

对象与方法

1. 调查对象与抽样方法:研究对象为长沙市农村中学生,包括 7~12 年级初中和高中生。采用随机分层整群抽样方法,先从长沙市 4 个郊县中,随机抽取 1 个作为样本县;然后,根据县教育局提供的学校名册及相关数据,以学校为单位,按学校规模、地理位置进行分层,随机抽取 6 所学校 22 个班级,被抽取班级的所有学生列入调查对象范围。调查时间为 2006 年 9 月和 11 月。采用自填式问卷,每班由一名老师和一名经过培训的调查员负责,按统一步骤和指令要求学生回答调查表的每个问题;学生知情同意参与调查,诚实回答问题,匿名填写,以确保调查真实准确。

2. 测量方法:暴力调查问卷主要包括两大部分:即暴力遭遇和危险因素。暴力遭遇采用 WHO 健康与生活经历问卷中的暴力遭遇部分^[4]。该问卷的制订考虑了不同国家之间的文化差异,可以进行跨文化应用(Cronbach $\alpha=0.84$),已越来越多地被世界各国所采用。本次调查中,中学生遭遇的暴力可以来源于任何施暴者(父母、老师、同学、邻居、陌生人等);暴力发生率为过去一年是否遭遇暴力。暴力类型包括情感暴力、躯体暴力和性暴力,其中情感暴力包括以下情形:在他人面前受到言语侮辱、贬低或羞辱,受到他人有目的的恐吓、被威胁要伤害自己所关心的人、被虐待或错误对待等;躯体暴力包括以下情形:被打耳光、被推搡、朝你扔东西、被人在地上拖、被敲打、被踢、被抽打、被扼、被卡、被烧和受到枪、刀或钝器攻击等;性暴力包括以下情形:当事人不愿意或感到害怕时被强迫发生性关系;当事人不情愿或不高兴时被迫进行性接触。

中学生遭遇暴力相关的影响因素主要包括人口学因素、行为因素、社会心理因素。各因素的定义与赋值见表 1。自尊水平的评价采用 Rosenberg 自评量表,10 个条目,旨在评价受评者自信和自我价值的自我感受,得分范围为 10~40 分。依据 $\bar{x} \pm s$,将自尊水平分为三个等级,即高($>\bar{x}+s$)、中($\bar{x} \pm s$)、低($<\bar{x}-s$);学习成绩依据调查前学校登记的学生学期主要学习科目学习成绩总分,分为三个等级,即好(排名前 25%)、中(排名 25%~75%之间)、差(排名 75%以后)。

表 1 变量赋值表

变 量	定义及赋值
应变量	
暴力遭遇(过去一年) 否=0, 是=1	
自变量	
年龄(岁)	11~ =0, 13~ =1, 15~ =2, 17~18=3
性别	女=0, 男=1
父亲文化	小学及以下=0, 中学=1, 大学及以上=2
父亲职业	农民=0, 经商=1, 工人=2, 干部=3, 其他=4
母亲文化	小学及以下=0, 中学=1, 大学及以上=2
母亲职业	农民=0, 经商=1, 工人=2, 干部=3, 其他=4
泡网吧	不或极少=0, 经常=1
玩电脑游戏	不或极少=0, 经常=1
吸烟	不或极少=0, 经常=1
饮酒	不或极少=0, 经常=1
欺负他人	不或极少=0, 经常=1
学习成绩	差=0, 中=1, 好=2
父母婚姻	稳定=0, 再婚=1, 分居=2, 离婚=3, 单亲=4
父母吵架	不或极少=0, 经常=1
零用钱(元)	$\leq 50=0, 51 \sim 100=1, 101 \sim 150=2, 151 \sim 200=3, \geq 200=4$
居住场所	家=0, 学校=1, 老师家=2, 租房=3
自尊水平	低=0, 中=1, 高=2

注:无序多分类变量和有序多分类变量均以哑元形式纳入

3. 统计学分析:多元多水平 logistic 模型是同时对多个结果变量进行 logistic 回归分析,将几个结果变量视为 1 水平上的观察单位,用一个虚拟水平表示^[5]。本研究存在三个水平,分别是班级、学生个体和一个虚拟水平(情感暴力、躯体暴力、性暴力)。分析水平之间呈三层嵌套关系,即虚拟水平 1 嵌套于学生个体水平 2, 学生个体嵌套在班级水平 3。以情感暴力(y_1)、躯体暴力(y_2)和性暴力(y_3)为反应变量拟合三水平 logistic 回归模型如下:假设 $y_{1jk} \sim \text{Binomial}(1, \pi_{1jk}), y_{2jk} \sim \text{Binomial}(1, \pi_{2jk}), y_{3jk} \sim \text{Binomial}(1, \pi_{3jk})$ 。将单因素分析有意义的自变量纳入得到含三元三水平 logistic 模型:

$$\begin{cases} \text{logit}\pi_{1jk} = \beta_{(1)0k} + \beta_{2j}x_{(1)jk} + \dots + \beta_m x_{(1)jk}, \beta_{(1)0k} = \beta_{(1)0} + \nu_{(1)0k} \\ \text{logit}\pi_{2jk} = \beta_{(2)0k} + \beta_{2j}x_{(2)jk} + \dots + \beta_m x_{(2)jk}, \beta_{(2)0k} = \beta_{(2)0} + \nu_{(2)0k} \\ \text{logit}\pi_{3jk} = \beta_{(3)0k} + \beta_{2j}x_{(3)jk} + \dots + \beta_m x_{(3)jk}, \beta_{(3)0k} = \beta_{(3)0} + \nu_{(3)0k} \end{cases} \quad (1)$$

在个体水平,设计如下随机系数矩阵:

$$\begin{bmatrix} \nu_{0k} \\ \nu_{1k} \\ \nu_{2k} \end{bmatrix} \sim N(0, \Omega_k); \Omega_k = \begin{bmatrix} \sigma_{\nu_{0k}}^2 & & \\ \sigma_{\nu_{0k}\nu_{1k}} & \sigma_{\nu_{1k}}^2 & \\ \sigma_{\nu_{0k}\nu_{2k}} & \sigma_{\nu_{1k}\nu_{2k}} & \sigma_{\nu_{2k}}^2 \end{bmatrix} \quad (2)$$

将以上随机系数带入式(3)可以去除班级差异和自变量影响,进一步求出 $y_1、y_2、y_3$ 之间的相关系数。

$$\text{cov} \begin{bmatrix} y_{1jk} \\ y_{2jk} \\ y_{3jk} \end{bmatrix} \begin{bmatrix} \pi_{1jk} \\ \pi_{2jk} \\ \pi_{3jk} \end{bmatrix} = \begin{bmatrix} g(\pi_{1jk}) \\ \rho[g(\pi_{1jk})g(\pi_{2jk})]^{0.5} g(\pi_{2jk}) \\ \rho[g(\pi_{1jk})g(\pi_{3jk})]^{0.5} \rho[g(\pi_{2jk})g(\pi_{3jk})]^{0.5} g(\pi_{3jk}) \end{bmatrix} \quad (3)$$

模型中 $\beta_{(1)}, \beta_{(2)}, \beta_{(3)}$ 为反应变量 y_1, y_2, y_3 的系数, $x_{(1)}, x_{(2)}, x_{(3)}$ 分别为反应变量 y_1, y_2, y_3 的自变量, 下标“ k ”为班级标号, “ j ”为学生标号, “ jk ”代表第 k 个班级第 j 名学生, $g(\pi) = \delta\pi(1-\pi)/n$ 为二项分布的方差, $\delta=1$ 表示限定二项分布, ρ 为两个结果变量之间的相关系数。

结 果

1. 一般情况: 共发出调查问卷 3960 份, 两个班级 120 名学生因为课程安排冲突未接受调查, 175 名学生因缺席也未参加调查, 在收回的调查问卷中, 45 份因质量不合格被剔除。共计收回有效调查问卷 3620 份, 有效应答率 91.4%。

在接受调查 3620 名农村中学生中, 男生 1755 人 (48.5%), 女生 1865 人 (51.5%), 年龄 11~18 岁, 男女生平均年龄基本相当 ($P>0.05$); 11~岁、13~岁、15~岁和 17~18 岁组人数分别占 9.8%、37.5%、24.3% 和 28.4%; 初中和高中生人数基本接近 (50.4%, 49.6%), 其中 71.4% 及 77.5% 来自于农民家庭 (依据为父母职业), 其父母文化程度在中学及以下者占 90% 以上。

2. 暴力发生率: 3620 名农村中学生中, 情感暴力、躯体暴力、性暴力的发生率分别为 21.5%、24.3%、2.0%; 除性暴力外, 情感暴力、躯体暴力差异有统计学意义 ($\chi^2=131.5, P=0.000; \chi^2=84.2, P=0.000$), 男生高于女生 (表 2)。

表 2 3620 名不同性别农村中学生暴力发生情况

性别	人数	构成比 (%)	年龄 (岁, $\bar{x} \pm s$)	暴力类型		
				躯体暴力 人数率 (%)	情感暴力 人数率 (%)	性暴力 人数率 (%)
男	1755	48.5	14.99 ± 1.89	574 32.7*	491 28.0*	29 1.7
女	1865	51.5	14.88 ± 1.86	305 16.4*	288 15.4*	44 2.4
合计	3620	100.0	14.93 ± 1.88	879 24.3	779 21.5	73 2.0

注: * $P<0.01$

3. 多元多水平 logistic 回归分析: 由于本次研究资料具有层次结构的多反应变量特征, 故本次分析以过去一年是否遭遇情感暴力、躯体暴力、性暴力为多应变量, 采用多元多水平 logistic 回归模型分析暴力遭遇发生的影响因素, 并与一般的多因素 logistic 回归分析的结果进行比较, 分析工具采用 SPSS 15.0 和 MLwiN (version 2.02) 软件。对班级 (水平 2) 的随机参数做假设检验得到: 情感暴力发生率在班级

内不存在聚集性 ($\chi^2=3.259, P=0.071$); 躯体暴力发生率在班级内存在聚集性 ($\chi^2=4.286, P=0.038$); 性暴力发生率在班级内不存在聚集性 ($\chi^2=3.512, P=0.061$); 情感暴力与躯体暴力在班级内不存在关联性 ($\chi^2=3.160, P=0.075$); 情感暴力与性暴力在班级内存在关联性 ($\chi^2=4.239, P=0.039$); 躯体暴力与性暴力在班级内存在关联性 ($\chi^2=4.482, P=0.034$)。

$$\begin{bmatrix} \nu_{0k} \\ \nu_{1k} \\ \nu_{2k} \end{bmatrix} \sim N(0, \Omega_k); \Omega_k = \begin{bmatrix} 0.065(0.036) & & & & & & \\ & 0.080(0.045) & & & & & \\ & & 0.147(0.071) & & & & \\ & & & 0.210(0.102) & & & \\ & & & & 0.307(0.145) & & \\ & & & & & 0.624(0.321) & \\ & & & & & & \end{bmatrix}$$

采用多应变量多水平模型将班级水平的随机效应分离出来后得到情感暴力、躯体暴力、性暴力间的相关系数与控制协变量后计算得到的偏相关系数相比均略有差异, 更为真实地反映三者之间的相关关系 (表 3)。

表 3 暴力类型间的相关系数矩阵

变 量	偏相关系数矩阵			多水平相关系数矩阵		
	y_1	y_2	y_3	y_1	y_2	y_3
情感暴力 (y_1)	1			1		
躯体暴力 (y_2)	0.377*	1		0.369*	1	
性暴力 (y_3)	0.133*	0.131*	1	0.108*	0.110*	1

注: * $P<0.01$

多应变量多水平模型理论考虑了数据的非独立性及分布于不同水平上的误差, 模型拟合采用 logistic 模型和 PQL 方法给予样本更为准确的参数化的估计^[6]。本研究考虑了班级效应和残差之间的相关性, 计算估计系数与一元 logistic 回归不同, 且所估计系数的标准误均小于相应的一元 logistic 回归系数的标准误。在单应变量模型中 17~18 岁年龄组、玩电脑游戏对躯体暴力的影响有统计学意义, 但在多应变量多水平模型中无统计学意义, 低自尊组对躯体暴力的影响程度有所降低 (表 4)。

讨 论

由于本次研究资料具有层次结构的多反应变量特征, 如果分别以情感暴力、躯体暴力和性暴力为应变量, 将所有个人变量拟合 logistic 模型, 所得结果将有以下问题^[6]: 其一, 传统回归模型的估计方法是建立在个体因素相互独立的假设上, 当这个假设不成立时, 回归模型中的各参数估计值的有效性和统计特性均会受到影响, 从而最终的统计推断结论将可能偏差; 其二, 如果从学生个体水平来分析, 则未将班级作为总体的一个随机样本, 不能说明不同班

表 4 两种模型回归系数及标准误比较

变 量	一般单因变量模型			多元多水平模型		
	γ_1	γ_2	γ_3	γ_1	γ_2	γ_3
β_0 (常数项)	-0.526(0.164)*	-0.591(0.167)*	-4.181(0.145)*	-0.562(0.176)*	-0.564(0.198)*	-3.864(0.270)*
β_1 (年龄,岁)						
11~	1	1	-	1	1	-
13~	-0.232(0.145)	0.131(0.149)	-	-0.169(0.141)	0.143(0.143)	-
15~	-0.522(0.156)*	-0.095(0.157)	-	-0.390(0.149)*	-0.030(0.151)	-
17~18	-0.521(0.153)*	-0.318(0.156) [†]	-	-0.351(0.145)*	-0.172(0.150)	-
β_2 (性别)	-0.779(0.088)*	-0.868(0.086) [†]	-	-0.747(0.083)*	-0.842(0.082)*	-
β_3 (玩电脑游戏)	-	0.300(0.151) [†]	-	-	0.248(0.136)	-
β_4 (吸烟)	-	0.981(0.256)*	-	-	0.847(0.237)*	-
β_5 (父母吵架)	0.466(0.146)*	0.401(0.144)*	0.708(0.345) [†]	0.515(0.140)*	0.460(0.139)*	0.922(0.278)*
β_6 (欺负他人)	1.320(0.166)*	0.926(0.171)*	1.445(0.336)*	1.327(0.164)*	0.913(0.168)*	1.317(0.296)*
β_7 (自尊)						
低	1	1	-	1	1	-
中	-0.182(0.110)	-0.291(0.107)*	-	-0.173(0.106)	-0.258(0.103) [†]	-
高	-0.434(0.149)*	-0.357(0.143) [†]	-	-0.382(0.143)*	-0.275(0.135) [†]	-

注: $\alpha_{.1} = 0.05, \alpha_{.05} = 0.10$; * $P < 0.01$, [†] $P < 0.05$; 括号外数据为回归系数, 括号内数据为标准误

级之间是否存在差异,还可能增加犯第一类错误的机会;而多元多水平模型将班级的差异分离出来后,剩下的只是评价结果在学生间的变异度,如本次研究表明在单应变量模型中年龄、玩电脑游戏对躯体暴力的影响有统计学意义,但在多应变量多水平模型中无统计学意义,低自尊组对躯体暴力的影响程度有所降低。

在研究具有层次结构资料时,层次结构本身也是一种混杂影响因素,传统最小二乘法不能排除其结构效应,因而计算得到的部分相关系数偏大,而多水平模型则可通过高一级水平的设置将其随机效应分离出来,使结果更加准确^[7]:①要分析情感暴力、躯体暴力和性暴力之间的相关性,班级水平是混杂因素,借助多元多水平模型可将班级水平的随机效应分离出来,能够更好地估计三种暴力之间的相关系数;②多元多水平 logistic 回归分析既可以分别对各种暴力类型的影响因素进行评价,又可以比较某因素对各种暴力类型的影响是否相同,比单个结果变量的 logistic 模型提供了更多的信息;③利用虚拟 1 水平的数据结构可以有效地处理当某个体有个别观察结果变量缺失的情况,用一元 logistic 回归需要剔除这些个体,而借助多水平模型来估计就可以充分利用所有信息,只要缺失是随机的其估计将是有效且无偏的^[5]。

社会环境的复杂多变,政治、经济和文化各个领域都发生很多变化,青少年在成长的过程中,无时无刻不受到随着社会变革而来的各种不良因素的影

响。市场经济带来的思想意识形态的变化,使他们对物质的追求和非法占有的欲望越来越强,在这种社会大环境的影响下,目前在校园或周边地区,侵害学生人身、财产安全的事件时有发生。由于社会环境的复杂性和难以测量,本研究未能深入。本研究缺乏有关测量班级环境的变量,如师生关系、同学关系、秩序和纪律、竞争和学习负担;故不能进一步分析班级间差异的可解释因素,有待改进。

综上所述,在采取有关措施降低大学生暴力遭遇发生率时,除应注意一些个人因素和家庭因素的影响外,班集体小环境所起的作用也不容忽视,多元多水平分析结果说明有些班级小环境较之于其他班级更易于使其成员遭遇暴力。

参 考 文 献

- [1] 唐晓立. 世界暴力与卫生报告. 北京: 人民卫生出版社, 2002: 4.
- [2] 王声湧. 伤害流行病学. 北京: 人民卫生出版社, 2003: 402-405.
- [3] 荆春霞, 王声湧, 陈青山, 等. 广州市中学校园暴力发生情况及原因分析. 中国学校卫生, 2005, 26(1): 22-23.
- [4] WHO Multicountry Study on Women's Health and Life Experiences Questionnaire (Version9). Geneva, Switzerland: World Health Organization, 2000.
- [5] 丁浩, 丁红, 赵杨, 等. 临床试验中多个终点变量同时评价的多元 logistic 模型. 中国卫生统计, 2007, 24(3): 251-253.
- [6] 杨珉, 李晓松. 医学和公共卫生研究常用多水平统计模型. 北京: 北京大学医学出版社, 2007: 2.
- [7] 李智文, 刘钢, 陶育纯. 双变量多水平模型在多重相关中的应用. 中国卫生统计, 2001, 18(1): 9-10.

(收稿日期: 2008-07-19)

(本文编辑: 尹廉)