

准试验研究

窦亚兰 戴江红 黄爱龙

【关键词】 准试验; 研究设计; 流行病学

Quasi-experiment study Dou Yalan¹, Dai Jianghong¹, Huang Ailong². 1 Department of Epidemiology and Health Statistics, School of Public Health, Xinjiang Medical University, Urumqi 830011, China; 2 Chongqing Medical University

Corresponding author: Dai Jianghong, Email: epi102@sina.com

This work was supported by a grant from the National Science and Technology Major Project of China (No. 2012ZX10001-007).

【Key words】 Quasi-experiment; Study design; Epidemiology

随机对照试验(randomized controlled trials)是在人群中进行的、前瞻性的、用于评估干预措施效果的对照研究。公认的第一个随机对照试验是1948年英国医学总会进行的链霉素治疗肺结核试验。

在实际情况下,因研究条件所限不能随机分组或不能设立平行对照,这种研究称为“类试验”或“准试验”(quasi-experiment)。准试验研究的一个特点是对研究对象采取非随机分配措施,当研究对象多、范围广,较难做到随机分配时,可选择准试验研究。准试验研究的另一个特点是可不另设对照组,而以自身为对照,即干预试验前后相比较^[1-2]。准试验研究设计已运用于多领域,如评估国家政策方针的影响、卫生干预措施的效果等^[3-5]。

伴随着准试验的广泛应用,产生了一些研究类型,包括中断时间序列设计(interrupted time-series design)、回归不连续设计(regression discontinuity design)、回归点位移设计(regression point displacement design)、前测后测设计(pre-test/post-test design)等。

1. 中断时间序列设计:

(1)概念:中断时间序列设计是一类用于探索干预措施的效应是否显著大于潜在时间趋势,评价干预措施效果(如国家指导方针等)的研究。该设计在干预措施实施前后的多个时间点即时间序列收集数据,估计干预前时间序列的潜在趋势后,根据潜在趋势线和干预后的时间序列趋势评估干预措施效果。该设计经常很难收集到足够的数据点,因此常用数据库或数据程序作为数据来源,分析中常使用时间序列模型^[6]。

(2)应用实例:2005年11月英格兰政府取消了酒精销售的时间限制,为了评估该法案对暴力事件发生的影响,有研究收集了2004年2月至2007年12月英国某市的暴力事件发

生率、抢劫率、总犯罪率,以法案实施时间2005年11月为中断点,分析法案实施前后暴力率的变化^[7]。

该研究以实施法案后犯罪率的整体趋势、深夜暴力发生率作为变量,通过差分自回归移动平均(ARIMA)模型检验研究假设,使用时间序列的自相关函数(ACF)、偏自相关函数(PACF)和ARIMA模型的残差验证时间序列和模型的适配性。为了分析评估该法案的影响,每个时间序列通过并入ARIMA模型的转换函数以验证政策实施的时间对暴力或其他犯罪发生的影响趋势。

研究发现,该法案的实施未影响总体暴力发生率,但使暴力事件转向周末03:00—06:00集中发生。法案实施后暴力事件发生频数略微增加;暴力事件发生比例显著增加。

中断时间序列设计的其他应用:评估英国对医生采取按工作绩效支付薪酬的政策对高血压患者治疗和转归的影响^[8];加纳北部为提高孕产妇和儿童健康质量设立的国家项目用以评估大型卫生干预措施的效果等^[9]。

2. 回归不连续设计:

(1)概念:回归不连续设计给予干预措施一个截点,比较干预前后的实际数据分布,建立统计学回归模型,通过比较干预前后回归趋势的差异评价效应,也可用于评估政策教育干预等,为了达到和传统试验设计相当的功效,需要大量研究者参与和精确的回归模型建立^[3,10]。

(2)应用实例:2001年日本政府推出了根据年龄制定不同医疗费用分担率的政策,将达到70岁人群的医疗保健费用分担从30%减至10%^[11]。为了评估政策对居民身心健康的影响,2007年的一项研究从全国代表性横断面调查中获取10 293名64~75岁成年人数据,以70岁年龄点作为截点,以身心健康测评得分为变量,比较研究对象70岁前后身心健康得分的变化趋势。研究根据既往数据总结出老年人身心健康得分随着年龄变化而平稳变化的回归趋势,再对70岁后的变化做出估计的回归趋势,通过对两组回归趋势进行比较以反映政策带来的效果,并建立线性模型以衡量政策的影响。结果显示,该政策的实施有益于日本老年人的身心健康。

回归不连续设计的其他应用实例:评估空气质量预警的发布对南加州居民户外活动频繁度的影响^[12];评估美国医疗补助资格、私人医疗保险以及医疗利用率对儿童健康的因果影响^[13];探讨美国儿童进入幼儿园或小学与其BMI变化值、超重/肥胖是否存在因果关系^[14];通过评估中国高血压患者确诊后告知对其膳食结构的影响,探索告知诊断疾病信息是否会导致其产生更健康的生活方式等^[15]。

3. 回归点位移设计:

(1)概念:回归点位移设计中通常包括一个单一的干预组

DOI: 10.3760/cma.j.issn.0254-6450.2015.09.024

基金项目: 国家科技重大专项(2012ZX10001-007)

作者单位: 830011 乌鲁木齐, 新疆医科大学公共卫生学院流行病学与卫生统计学教研室(窦亚兰、戴江红); 重庆医科大学(黄爱龙)

通信作者: 戴江红, Email: epi102@sina.com

和多个对照组,并基于一个回归模型进行比较。该设计根据多个对照组的前测及后测值做出回归线,由干预组的前测后测值到该回归线的垂直位移评估干预效果。这类设计主要是一种分析策略,更适用于医疗保健方案的评价。它一般使用总体数据,因此更稳定、更精确,易于实施且成本低,在评估保健试点项目中显示出了巨大潜力^[16-17]。

(2)应用实例:这是一个以真实数据为背景的假设回归点位移设计研究,为了评估社区艾滋病教育计划“试点研究”的效果^[18]。研究中,对居民实施艾滋病相关知识教育干预措施前,1989年有10万新发艾滋病患者,干预后对1990年的新发患者人数进行统计。艾滋病患者数据来自美国96个地区1990年12月的艾滋病监测报告。其中95个地区已建立1989—1990年患者数的线性回归趋势,假设对剩余一个城市(迈阿密)的居民进行干预措施,模拟干预措施对艾滋病发病率的影响,迈阿密1989年的艾滋病发病率是0.532%,其1990年的艾滋病发病率已经从实际的0.584%被修改为0.484%。

4. 前测后测设计:

(1)概念:前测后测设计通常只安排1个被试组,在一段时间中按固定间隔对被试组进行一系列测试,让被试组接受干预后又按原来的间隔再进行一系列测试^[3]。特点是研究对象常与自身比较,分别在干预前后收集一系列数据,评估干预前后是否存在差异,解释干预效果。当然也有一部分前测后测设计安排干预组和对照组,在对结果进行分析时需要将对两组的后测结果进行比较,从而探讨干预效果。

(2)应用实例:前测后测设计在推进新的医学模式方面有所应用。如今许多国家的养老院在护理方面从传统方式转向了综合方式,对工作质量有着直接影响。有研究为了探讨综合护理方法在养老院的效果,在荷兰2家养老院采用前测后测设计实施干预(综合护理),一家养老院提供5个试验病房(采用综合护理方式),另一家养老院提供4个对照病房(采用传统护理方式),在干预前后对护理人员进行工作压力、满意度、工作质量等量表测量^[19]。对于研究数据首先分析干预前测量分值的差异,接下来使用协方差分析评价干预后的效果,最后对干预前测试分值进行校正,利用多元回归分析评估干预组和对照组之间的差异。结果显示,实施综合护理的病房护理人员更能根据患者需求提供护理和服务。

前测后测设计的其他应用:评估按摩疗法对急诊医护人员职业压力和焦虑的作用效果^[20];评估氯喹治疗疟疾的卫生策略对加纳的综合初级卫生保健服务的效果等^[21]。

上述四类准试验设计在目的与设计方面有相似之处,前三种设计在数据处理方面存在“可利用回归模型比较干预效果”等共同点,中断时间序列设计数据需以时间序列为基础,回归不连续设计需以截点值为界进行干预前后回归模型比较,回归点位移设计则是通过干预组到对照组回归线的位移估计干预效果。前测后测设计与回归不连续设计虽在设计方面有相似之处,但在数据处理方面存在区别;前测后测设计在单组设计的情况下,需根据干预前的多次测量结果求出回归方程,推算出在假设不接受干预的情况下以后各测试时间点上可能出现的结果,将其与实际干预后的测量结果进

行比较;前测后测设计在设立了对照组的情况下,则是以干预组和对照组的后测结果进行比较,在控制时间效应情况下评价干预措施的效果。

准试验研究存在局限性。在随机对照试验中,研究对象有相同的机会和概率被分配到干预组或对照组,因此干预组和对照组在基线上是相同的。而准试验设计中,研究人员会用非随机的方法来决定哪些研究对象进入干预组,哪些进入对照组,所选的标准也可以是未知的,因此干预组和对照组在基线上可能是不可比的,一些因素如成本、可行性等可能会影响到干预措施的实施,因此准试验受内在效度影响,不能确切证明在干预措施和数据间存在因果联系。

参考文献

- [1] Government Social Research Unit. The magenta book: guidance notes for policy evaluation and analysis [M]. London: HM Treasury, 2007: 11-13.
- [2] Liu DW. Experimental epidemiology [M] // Zhan SY. Epidemiology. 7th ed. Beijing: People's Medical Publishing House, 2012: 119. (in Chinese)
- [3] 刘殿武. 实验流行病学 [M] // 詹思延. 流行病学. 7版. 北京: 人民卫生出版社, 2012: 119.
- [4] Zuckerman IH, Lee E, Wutoh AK, et al. Application of regression-discontinuity analysis in pharmaceutical health services research [J]. Health Serv Res, 2006, 41(2): 550-563.
- [5] Murray DM, Pennell M, Rhoda D, et al. Designing studies that would address the multilayered nature of health care [J]. J Natl Cancer Inst Monogr, 2010, 2010(40): 90-96.
- [6] Kessler RC. Quasiexperimental designs in AIDS psychosocial research [M] // Ostrow DG, Kessler RC eds. Methodological Issues in AIDS Behavioral Research. US: Springer, 1993: 75-92.
- [7] Grimshaw J, Campbell M, Eccles M, et al. Experimental and quasi-experimental designs for evaluating guideline implementation strategies [J]. Fam Pract, 2000, 17 Suppl 1: S11-16.
- [8] Humphreys DK, Eisner MP, Wiebe DJ. Evaluating the impact of flexible alcohol trading hours on violence: an interrupted time series analysis [J]. PLoS One, 2013, 8(2): e55581.
- [9] Serunaga B, Ross-Degnan D, Avery AJ, et al. Effect of pay for performance on the management and outcomes of hypertension in the United Kingdom: interrupted time series study [J]. BMJ, 2011, 342: d108.
- [10] Singh K, Speizer I, Handa S, et al. Impact evaluation of a quality improvement intervention on maternal and child health outcomes in Northern Ghana: early assessment of a national scale-up project [J]. Int J Qual Health Care, 2013, 25(5): 477-487.
- [11] Linden A, Adams JL, Roberts N. Evaluating disease management programme effectiveness: an introduction to the regression discontinuity design [J]. J Eval Clin Pract, 2006, 12(2): 124-131.
- [12] Nishi A, McWilliams JM, Noguchi H, et al. Health benefits of reduced patient cost sharing in Japan [J]. Bull World Health Organ, 2012, 90(6): 426-435.
- [13] Neidell M. Air quality warnings and outdoor activities: evidence from Southern California using a regression discontinuity design [J]. J Epidemiol Commun Health, 2010, 64(10): 921-926.
- [14] De La Mata D. The effect of Medicaid eligibility on coverage, utilization, and children's health [J]. Health Econ, 2012, 21(9): 1061-1079.
- [15] Anderson PM, Butcher KF, Cascio EU, et al. Is being in school better? The impact of school on children's BMI when starting age is endogenous [J]. J Health Econ, 2011, 30(5): 977-986.
- [16] Zhao M, Konishi Y, Glewwe P. Does information on health status lead to a healthier lifestyle? Evidence from China on the effect of hypertension diagnosis on food consumption [J]. J Health Econ, 2013, 32(2): 367-385.
- [17] Linden A, Trochim WM, Adams JL. Evaluating program effectiveness using the regression point displacement design [J]. Eval Health Prof, 2006, 29(4): 407-423.
- [18] Conklin A, Nolte E. Disease management evaluation: a comprehensive review of current state of the art [R]. RAND technology report, RAND Corporation, 2011.
- [19] Campbell DT, Krauss BJ. Speculations on quasi-experimental design in HIV/AIDS prevention research [J]. J Meth Measur Soc Sci, 2012, 3(1): 52-84.
- [20] Boumans NP, Berkhout AJMB, Vijgen SMC, et al. The effects of integrated care on quality of work in nursing homes: a quasi-experiment [J]. Int J Nurs Stud, 2008, 45(8): 1122-1136.
- [21] Cooke M, Holzhauser K, Jones M, et al. The effect of aromatherapy massage with music on the stress and anxiety levels of emergency nurses: comparison between summer and winter [J]. J Clin Nurs, 2007, 16(9): 1695-1703.
- [22] Agyepong IA, Ansah E, Gyapong M, et al. Strategies to improve adherence to recommended chloroquine treatment regimes: a quasi-experiment in the context of integrated primary health care delivery in Ghana [J]. Soc Sci Med, 2002, 55(12): 2215-2226.

(收稿日期: 2015-03-20)

(本文编辑: 万玉立)