

政策、项目干预效果评价 的不同设计及内部效度比较研究*

金承刚^①, 李 佳^①, 陈 丽^①, 梁小云^①

摘 要 目的: 比较政策、项目干预效果评价不同设计的内部效度。研究方法: 在理论上比较不同的评价设计及其效果分析模型, 利用深圳社康改革的门诊电子数据做实例, 对不同设计的内部效度进行比较。结果: 不同的评价设计得出不同的干预效果, 甚至是相反的结果。结论: 在政策及项目干预效果评价中, 建议采用双差法。

关键词 评价设计; 内部效度; 双差法

中图分类号 R1-9; R197 **文献标识码** A **文章编号** 1003-0743(2012)04-0079-04

Comparison of Internal Validity of Different Designs for Policy and Program Impact Evaluation/JIN Cheng-gang, LI Jia, CHEN Li, et al./Chinese Health Economics, 2012,31(4):79-82

Abstract Objective: To assess the internal validity of different program impact evaluation designs. **Methods:** Comparing the threats to internal validity of various impact evaluation designs and their impact identification models, and using electronic data from Shenzhen's Community Health Reform Pilot as an example to illustrate the internal validity of different designs. **Results:** Different evaluation designs come to a different result, or even to an opposite one. **Conclusion:** Double Difference for impact evaluation design is strongly proposed for its strong internal validity.

Key words impact evaluation; internal validity; double difference

First-author's address Social Development and Public policy Institute, Beijing Normal University, Beijing, 100875, China

1 背景

随着经济、社会的发展, 越来越多的政策、项目得以实施。如重大传染病综合防治试点、合作医疗制度、基本药物政策、社区卫生机构收支两条线、社区公共卫生服务项目等等。这些干预成为提高健康水平的重要途径。

这些政策、项目是否有效? 这需要系统的评价设计和数据分析才能回答。而科学评价成为循证政策的基础。本文介绍基本的评价概念和设计, 并模拟不同的评价设计分析其效度, 以提高评价结果的真实性。

2 效果评价的基本概念和设计

2.1 干预的效果

项目的效果是指通过项目干预后效果变量的值减去如果没有项目干预的效果变量值, 二者之差就是项目干预的效果。如开展了医疗保险项目干预, 目的之一是减少病人的自付费用。干预后, 可以收集数据得到第 i 个人干预后的自付费用 Y_{i1} 。如果这个人没有参加医疗保险, 自付的费用是 Y_{i0} 。那么, 干预效果则是这个人在干预后的效果变量减去如果不参加项目的效果变量, 即项目效果为: $\Delta Y_i = Y_{i1} - Y_{i0}$ 。

但是, 这在现实中是不可能做到的。因为同一个人不可能同时存在于两种状态中, 要么被项目干预, 要么没有被项目干预, 不可能同时获得数据 Y_{i1} 、 Y_{i0} 。因

此, 在评价中用术语“counterfactual”来描述这种问题。“counter”的意思是相反, “factual”是事实, 即和事实相反。其含义是如果不干预, 效果变量会是多少。

2.2 评价设计的“Counterfactual”处理

评价设计的一个重要任务是如何处理“counterfactual”。如果病人不治疗, 病人的健康状况会如何? 要想回答这个问题, 就要处理好“counterfactual”^[1]。处理方法: (1) 如果不治疗, 其健康状况和治疗前一样, 把治疗前作为“counterfactual”进行比较; (2) 如果不治疗, 病人的健康状况和对照组一样, 则把对照作为“counterfactual”进行比较。对“Counterfactual”的选择和处理, 直接影响到评价的结果, 甚至会得到相反的结论。

评价的金标准是临床随机区组设计, 由于随机分组, 两组在所有的变量上都是可比的, 用随机对照作为“Counterfactual”, 即潜含的意思是指如果不治疗, 那么效果变量的值和随机对照组一样。

2.3 评价内部效度和外部效度^[2]

内部效度 (internal validity) 是指针对给定的样本, 做出的统计推断是真实的。可以认为有四种原因来解释效果变量的变化: (1) 混杂; (2) 测量偏倚和选择偏倚; (3) 抽样误差; (4) 政策、项目干预^[3]。只有排除了前面3个因素后, 评价结果才具有较高的内部效度, 又称为净效果评价。而外部效度则是指把研究结果和项目经验应用到不同的地点、不同的人群、不同的时间, 是否能复制出同样的结果的能力。

本研究利用社康试点的数据, 来比较干预前后设计, 非随机对照设计及双差法设计的结果, 探讨评价设

* 基金项目: 部委基金项目 (2009ZX10004-903)。

① 北京师范大学社会发展与公共政策学院 北京 100875
作者简介: 金承刚 (1962-), 男, 博士学位, 研究方向: 项目效果评价和经济学评价; E-mail: cegin2005@126.com。

设计的内部效果，以提高政策、项目评价的水平。

3 常用的评价研究的类实验设计

理想的效果评价设计是随机区组，但是在政策和社区人群干预中，几乎是做不到随机区组。因此，在这种情况下，只能采用类实验研究设计。目前，以下是常见的评价设计^[4-5]：

3.1 干预前、干预后比较设计

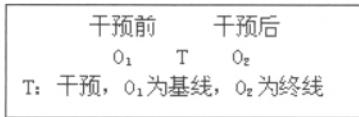


图1 干预前-后比较设计

这种设计分为两类，一类是前后的观察都是针对同一个人或机构，因此，形成一个追踪数据。另一类是干预前做一次随机抽样，干预后再做一次随机抽样，前后两次是截面数据(图1)。针对干预前测-后测两次截面数据设计的效果评价模型为：

$$Y = \beta_0 + \beta_T \cdot \text{Time} + \sum \beta_j X_j$$

Y: 效果变量; β_T : 干预效果; Time: 0为干预前, 1为干预后; X_j : 需要控制的混杂因素或控制变量

但如果政策干预和项目干预的时间较长，而且政策的实施是在一个快速变化的环境中，在基线和终线的两次数据收集之间会有很多事件发生，并作用于所要评价的效果变量，从而使项目评价的结果产生偏倚。

3.2 干预-对照组比较设计

首先，有一组机构或人群实施了某项政策或项目干预，由于无法随机选择对照，只能选择和干预组较为接近的作为对照(图2)。

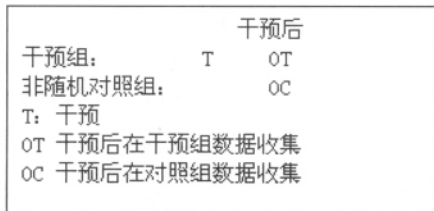


图2 干预-对照比较设计

针对干预前测-后测设计的效果评价模型为：

$$Y = \beta_0 + \beta_C \cdot \text{Group} + \sum \beta_j X_j$$

Y: 效果变量; β_C : 干预效果; Group: 0为对照组, 1为干预组; X_j : 需要控制的混杂因素或控制变量

这种设计的潜在的问题是选择性偏倚，导致评价结果的失真。选一个差的作为对照，那么效果往往为正，干预有效，如果选择一个好的作为对照，往往分析的效果是负的，则会认为干预使得效果差。

3.3 双差法

双差法(double difference)的设计是一组为干预组，非随机选择条件较接近的另一组为对照组。两组在干预前进行基线数据收集，在干预后，两组再进行一次数据收集(图3)。

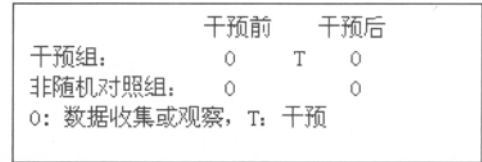


图3 双差法设计

由于是非随机选择对照，在干预前，两组的效果变量即存在差异 Δy_0 ，在项目干预后，两组又存在一个差异 Δy_1 ，而实际的干预效果是 $\Delta y_1 - \Delta y_0$ ，所以称为双差法(图4)。

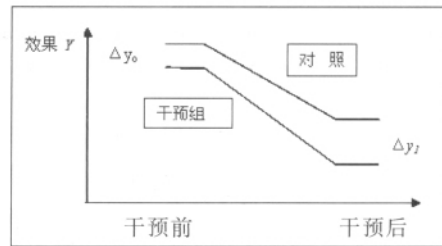


图4 双差法设计图示

$$Y = \beta_0 + \beta_C \cdot \text{Group} + \beta_T \cdot \text{Time} + \beta_{CT} \cdot \text{Group} \cdot \text{Time} + \sum \beta_j X_j$$

Y: 效果变量; Group: 0为对照组, 1为干预组; Time: 0为干预前, 1为干预后; Group·Time: Group和Time的乘积, 即交互项; X_j : 需要控制的混杂因素或控制变量; β_C : 俘获了干预前两组的差; β_T : 俘获了在干预期间项目干预以外的其它共同作用于两组的因素; β_{CT} : 则为控制了基线时的差异、干预期间其它因素的作用、以及其他混杂因素之后, 项目干预的效果。

这种设计可以很好地处理两组基线时的差异，并排除同时作用于两组的其他因素。

4 三种评价设计的效果比较

以下利用社区卫生改革的数据，采取上述3种评价设计，比较分析3种设计的评价结果和内部效度。

4.1 干预和数据来源

深圳市宝安区选择了观澜和西乡街道，将院办院管的社区卫生模式改革为相对独立的社区卫生管理模式。本研究选择观澜作为干预组，与其相接近的松岗作为非随机对照组。收集试点组和对照组在改革前一年和后一年的电子门诊数据。由于费用呈严重的偏态，因此采用中位数回归。

4.2 变量的选择及定义

一般地，门诊费用主要受到疾病种类和疾病的严重

性、人口、社会、经济状况、医保的影响^[6-7]。根据上述的有关变量从社康信息系统中提取数据,但数据中没有家庭人均收入。本研究因变量、自变量及定义见表1。

4.3 样本描述与分布

试点前对照组和试点组的基线水平有差异。具体反映在年龄、教育程度、医保类型的差异,特别是两组的门诊费用(中位数)就存在着明显的差异。同样在试点后,两组存在较大的差异,见表2。

4.4 三种评价设计的效果比较

表1 变量的选择及定义

变量	定义
因变量	
门诊费	门诊费用包括挂号费、西药费、中药费、治疗费、观察费等其它费用
自变量	
组别	0:对照组;1:试点组。
试点前后	0:试点前;1:试点后。
效应	组别 * 试点前后(两个虚拟变量的乘积)
性别	0:男;1:女
年龄	患者年龄连续变量
受教育程度	1:初中及以下;2:高中高职及以上;3:其他
医保类型	1:自费;2:劳工医保;3:其他医保

表3中的模型(1)是前后比较设计,在模型中用Time这个虚拟变量测量干预的效果,即在控制混杂因素后,干预前后所发生的变化。如果采取这种评价设计,那么干预的效果则为0.691元($P=0.000$),即改革导致了社区卫生门诊费用上涨了0.691元;模型(2)是非随机

表2 对照组、试点组样本特征

特征变量	对照组		试点组	
	%或 Mean±s		%或 Mean±s	
	试点前 (n=45 614)	试点后 (n=82 314)	试点前 (n=114 445)	试点后 (n=139 152)
性别(男性)	49.36%	49.12%	48.62%	47.94%
年龄(岁)	27±16.32	25±16.58	29±14.75	28±14.64
受教育程度				
初中及以下(含少儿)	61.53%	49.45%	62.94%	62.65%
高中、高职及以上	14.60%	24.15%	18.11%	20.20%
其他	23.87%	26.40%	16.95%	17.15%
医保类型				
劳工医保	23.92%	21.14%	22.26%	15.50%
其他医保	6.54%	9.94%	2.76%	2.46%
自费	69.54%	68.92%	74.98%	82.03
门诊费用(元)				
平均费用	48.13±35.79	49.19±35.66	42.09±38.49	44.64±43.56
中位数	39.00	43.00	33.73	34.50

对照设计,变量Group的系数则为干预效果,即在控制了混杂因素后,和对照组相比效果变量的变化;如果采用了这种设计,那么试点干预的效果则为-8.959元($P=0.000$),即改革效果非常明显,费用下降-8.959元;模

型(3)是双差法设计,Group和Time的交互项的系数则为干预效果,干预效果为-3.208元($P=0.000$)。

5 讨论与建议

根据结果可知,不同的评价设计得出不同的干预效果,甚至是相反的结果。因此,评价研究的设计对评价的内部效度、评价的真实性起到关键的作用。

前后比较设计的结果是改革导致处方费用提高0.691元。采用了前后设计,由于两次测量之间的时间长,而在这一段时间内会有很多的政策、新的做法产生,这些都会作用于效果变量,这种设计对干预效果的估计则是该改革干预和其它因素共同作用的结果,因此,是一个混杂了其它因素的偏倚的结果。而采用非随机对照的结果是改革导致费用下降8.959元。如果采用了非随机对照的设计,那么就可能带来选择性偏倚:如果选择了差的组作为对照组,则项目效果的估计是往往是正的,是有效的;而如果选择了好的组作为对照组,则往往导致对项目干预效果的估计为负的。

采用双差法设计的评价结果是改革导致费用下降3.208元。双差法不仅有效地处理了两组在干预前的差异,同时能有效地处理在干预期间项目外的同时作用于

两组的其它因素或政策,因此,双差法具有较高的真实性和内部效度,可以产生可信的循证证据。

建议在可能的条件下,特别是外部环境快速变化的情况下,尽量采用双差法设计。但如果外部环境变化慢

表3 三种评价设计的社康改革对门诊费用的作用:中位数回归

	模型(1) 前后比较设计	模型(2) 非随机对照设计	模型(3) 双差法设计
Group		-8.959*** (0.000)	-5.955*** (0.000)
Time	0.691*** (0.000)		3.997*** (0.000)
Groupo Time			-3.208*** (0.000)
性别	3.101*** (0.000)	1.900*** (0.000)	2.027*** (0.000)
年龄一次方	1.128*** (0.000)	1.392*** (0.000)	1.371*** (0.000)
年龄二次方	-0.029 5*** (0.000)	-0.035 9*** (0.000)	-0.033 9*** (0.000)
年龄三次方	0.000 236*** (0.000)	0.000278*** (0.000)	0.000 253*** (0.000)
教育(参照:初中以下)			(0.000)
高中高职	-1.760*** (0.000)	-3.668*** (0.000)	-3.189*** (0.000)
其他	5.716*** (0.000)	3.246*** (0.000)	3.602*** (0.000)
医保类型(参照:自费)			
劳工医保	-4.229*** (0.000)	-2.189*** (0.000)	-2.902*** (0.000)
其他医保	2.020*** (0.000)	7.562*** (0.000)	6.247*** (0.000)

*** 示: $P < 0.001$ 。

或小组,即可以认为其它因素处于不变的情况下,可以采用前后比较设计。而非随机对照设计,选择偏倚是一个问题,会导致过高或过低估计,但如果能采用参与概率匹配的方法(propensity score matching)可以提高内部效度,但只解决可测变量(selection on observables)所导致的选择偏倚,而且,这种方法对数据要求高,要求干预组和对照组在预测的参与概率上有一定的共同重叠(common support)^[8]。

参 考 文 献

[1] Menzies P. Counterfactual Theories of Causation[M]. The Stanford Encyclopedia of Philosophy, 2009.
 [2] Briggs D C. Comments on Slavin: synthesizing causal inferences [J]. Educational Researcher, 2008(37):15-22.
 [3] Hennekens C H, Buring J E. Epidemiology in Medicine [M]. Little, Brown and Co, 1987.
 [4] Cook T D. Quasi-experimentation[M]. Chicago: Rand-McNally, 1979.
 [5] Trochim W M K . Advances in Quasi-experimental Design and Analysis[M]. San Francisco: Jossey-Bass, 1986.
 [6] 陈延, 王梅. 我国医院门诊费用研究综述[J]. 中国卫生经济, 2004,23(4):41-43.
 [7] 王爱荣. 门诊急诊医疗费用影响因素分析[J]. 中国卫生事业管理, 2004,20(1):45-46.
 [8] Ravallion M. Evaluating Anti-Poverty Programs[M]. T. Paul Schultz & John A. Strauss (ed.), Elsevier, 2008: 3787-3846.

[收稿日期: 2012-01-21] (编辑: 高非, 滕百军)

欢迎订阅 杂志

邮发代号: 14-97
 定 价: 12.00 元/册, 144.00 元/年, 全国各地邮局均可订阅
 发行代理: 京盛龙康(北京)文化传媒有限公司
 邮局汇款 地 址: 北京市海淀区紫竹院路1号人济山庄D栋804室
 邮 编: 100044
 银行汇款 户 名: 京盛龙康(北京)文化传媒有限公司
 开户银行: 中国工商银行北京海淀西区支行四道口分理处转紫竹桥支行
 账 号: 0200235509201000303

《中国卫生经济》杂志订阅通知单

订阅单位名称	收件人:
	邮政编码:
收件人详细地址	_____省_____市_____县(区)_____街道_____路_____号
订购清单	_____年_____期至_____期各_____册(注:定价12.00元/册)
汇款金额	人民币:_____万_____千_____百_____拾_____元整(¥_____ . 00)
订阅单位联系人	办公电话: _____ 手机: _____
意见反馈	

注: 汇款同时请务必详细填写以下《中国卫生经济》杂志订阅通知单并邮寄或传真至京盛龙康(北京)文化传媒有限公司, 以便及时、准确邮寄杂志和发票。京盛龙康(北京)文化传媒有限公司发行部传真: (010) 88556120/88556121。