

中国儿童死亡率差异分析

郝虹生

中国人口的死亡率经过几十年的持续下降,就目前的社会经济发展水平而言,已经达到了相当低的水平。但中国是一个人口众多、地域辽阔的多民族国家,不同地区、不同社会经济群体及民族之间在死亡率方面仍存在较明显的差异。对中国人口死亡率的地区差异分析表明,影响地区之间预期寿命水平差异最大的是儿童,特别是婴幼儿死亡率^①。因此降低儿童死亡率是进一步降低我国人口总死亡率的潜力所在。与其他年龄人口的死亡率相比,婴儿和儿童死亡率对社会经济因素作用的反应通常是最为敏感的。

本文利用第三次全国人口普查原始数据,专门就包括婴幼儿在内的儿童死亡率差异进行分析。

一、资料与方法

(一)资料。本文使用的资料是从第三次全国人口普查的户记录和人记录原始数据中抽取出的一个随机样本。由于普查原始数据是按地区排列的,因此使用等距抽样方法抽选样本,这样可保证样本单位均匀分布于各省、市、自治区。样本是按下列原则抽选的:

(1)以居住在家庭户内的已婚妇女为抽样单位;

(2)样本中每个妇女年龄在15~44岁之间,且至少生育过一个活产子女;

(3)普查时该妇女有配偶,且能够通过“与户主关系”的编码识别。

剔除少数有明显逻辑错误的样本单位,

共抽选出5738个符合上述要求的妇女作为样本。

(二)因变量的构造。特拉塞尔(Trussell)和普雷斯顿(Preston)曾提出一种方法,利用妇女曾生育活产子女数和存活子女数来构造儿童死亡指数,称为儿童死亡比(Child Mortality Ratio)^②。利用该指标,可以用单个妇女的儿童死亡比作为分析所用的因变量,用妇女及其丈夫和所在家庭户的一些特征指标作为自变量,以妇女个人作为分析单位进行多元统计分析,从而观察比较不同类型妇女之间的子女死亡差异。

这里需要对本文的研究对象“儿童”做一点说明。儿童通常是指15岁以下的人口,而本文主要是探讨婴幼儿的死亡率差异。由于研究所用的是育龄妇女的子女存活资料,因此可能会有一部分妇女的子女年龄较大,甚至可能超出儿童年龄范围。不过,可以认为这个问题对本文研究的影响不大。因为从年龄结构来看,处于育龄后期的妇女在所有育龄妇女中所占比例不大,样本中绝大部分妇女的子女应属儿童范围之内;其次,从死亡年龄模式来看,从婴儿开始,死亡率在儿童段是迅速下降的,大多数儿童死亡发生在婴幼儿组,因此有把握认为,本文中主要反映的还是婴幼儿死亡率差异。

构造儿童死亡比的基本思想是在对妇女年龄的作用进行标准化处理之后,用每个妇女曾生子女中死亡的比例作为衡量儿童死

① 郝虹生等:《中国分省死亡率分析》,《人口研究》,1988年第4期。

② Trussell, J. and Preston, S.: "Estimating the Covariates of Childhood Mortality from Retrospective Reports of Mothers", HEALTH POLICY AND EDUCATION, Vol. 3, pp. 1-36.

亡的指标。我们所以对妇女年龄的作用进行标准化,是因为妇女年龄越大,她的子女年龄也相应越大,因而暴露于死亡风险的时间就越长,其子女死亡的比例也就越趋于提高。

儿童死亡比的定义是每个妇女所生子女的实际死亡比例与按该妇女年龄的“期望”死亡比例之比。此处的期望死亡比例即起到标准化作用。用公式表示第j个年龄组第i个妇女的儿童死亡比为

$$CMR_{ij} = \frac{PD_{ij}}{EPD_j}$$

式中 PD_{ij} 为第j个年龄组第i个妇女的子女死亡比例;

EPD_j 为第j个年龄组妇女的期望子女死亡比例, 妇女年龄组为15~19岁, 20~24岁, ……40~44岁。

期望死亡比例的计算是将布拉斯 (Brass) 儿童死亡率估计方法反转。布拉斯曾提出儿童死亡比例与生命表指标的如下关系:

$$q(a) = d(j) \cdot k(j)$$

式中 $q(a)$ 为a岁前死亡概率, $d(j)$ 为第j个年龄组妇女的子女死亡比例, $k(j)$ 为第j个年龄组的系数。(a, j)的对应关系为 (1, 15~19), (2, 20~24), (3, 25~29), (5, 30~34), (10, 35~39), (15, 40~44)。系数 $k(j)$ 是全体妇女生育模式的函数, 根据不同年龄组妇女的累计曾生子女数算出。

如果将上述关系式反转, 则可以从某个标准生命表的死亡概率得到期望死亡比例:

$$EPD_j = q_s(a) / k(j)$$

其中标准死亡概率 $q_s(a)$ 来自某个与样本中儿童死亡率水平相同的模型生命表。根据特拉塞尔和普雷斯頓的比较试验, 选用不同的模型生命表时, 分析结果相当接近, 没有实质性影响。本文中使用的寇尔-德曼 (Coale-

Demeny) 西方模型生命表。选用生命表水平的标准, 是整个样本的实际儿童死亡比例和期望死亡比例相等, 即平均的儿童死亡比尽量接近1。通过多次试算, 发现使用西方模型的19.1水平, 可使样本的平均儿童死亡比达到1.0072, 因此 $q_s(a)$ 由该水平的模型生命表得出。式中 $k(j)$ 值是根据1982年普查全国机器汇总资料中的各年龄组妇女累计曾生活产子女数, 利用特拉塞尔的估计公式算出^①。

(三) 自变量。在国内外有关死亡率差异研究的文献中, 一般公认的死亡率影响因素包括: 医疗卫生条件、地理区域差异、城乡差异、民族差异以及社会、经济、文化差异 (包括职业、教育、收入等)。从第三次人口普查的户记录和人纪录中, 可以得到上述因素的一部分作为本文分析所用的自变量, 如母亲和父亲的受教育程度及职业, 母亲的民族, 户所在地理区域, 户的居住地特征 (市郊、镇、县) 等。另外, 样本中的妇女是由不同的出生同期群组成的, 如果死亡率不随时间变动, 由于妇女年龄已经在儿童死亡比的构造中被标准化, 因此妇女年龄将基本上不再是造成子女死亡率差异的一个因素。但实际情况是, 中国人口死亡率有明显的下降趋势, 时间越晚, 同一个年龄上的死亡率越低。这样, 妇女年龄仍会导致子女死亡的差异。考虑到这一点, 仍将妇女年龄包括在分析中, 作为控制变量使用。

由于分析所用的自变量大都是定类变量 (如民族、职业、地理区域等) 和定序变量 (如教育程度), 测度等级低, 不能直接用于多元统计分析, 因此必须将这类变量转换为虚拟变量组。每个初始变量由一组虚拟变量代表, 每个虚拟变量代表一个类别, 其中某一个类别取值为零作为参照类, 其余类别各构成一个虚拟变量。各虚拟变量所代表的类别在表1中给出定义。

① United Nations: "Indirect Techniques for Demographic Estimation", Manual X, 1983, p. 77.

表1 虚拟变量定义

变 量	类 别
母亲民族 NF0 NF1 NF2 NF3 NF4 NF5 NF6 NF7 NF8 NF9	汉 壮 回 维吾尔 彝 苗 藏 蒙 满 其它
母亲教育程度 EM0 EM1 EM2 EM3 EM4	不识字或识字不多 小学 初中 高中 大学
父亲教育程度 EF0 EF1 EF2 EF3 EF4	不识字或识字不多 小学 初中 高中 大学
母亲职业 OCM0 OCM1 OCM2 OCM3 OCM4 OCM5	不在业 农林牧渔劳动者 工人 商业与服务性工作人员 干部与办事人员 专业技术人员
父亲职业 OCF0 OCF1 OCF2 OCF3 OCF4 OCF5	不在业 农林牧渔劳动者 工人 商业与服务性工作人员 干部与办事人员 专业技术人员
居住地 P0 P1 P2 P3	县 镇 市郊 市区
地理区域 R0 R1 R2 R3 R4 R5	华北 东北 华东 中南 西南 西北

(四) 分析方法与模型。本文使用多变量分析, 以便在其它变量受控制条件下估计某个变量对儿童死亡差异的净作用。具体方法是以儿童死亡比作为因变量, 以虚拟回归形式进行协方差分析。虚拟回归方程如下:

$$CMR_{ij} = A + \sum B_k X_{ki} + E$$

① 有些少数民族样本人数仍嫌太少, 受随机误差影响较大, 结果不显著, 因此不一定反映实际情况。

其中 A=常数项,

B_k=变量K的虚拟回归系数,

X_{ki}=第i个妇女的k变量值,

E=余差项。

方程中各类别虚拟回归系数的含义, 是该类别与参照类相比的儿童死亡比增减量。

在虚拟回归分析中, 为了解决方差分布异质性 (heteroscedasticity) 问题, 对每个妇女的CMR_{ij}使用其曾生子女数做了加权处理。

对于多元分析结果, 我们关心的是:

(1) 各变量引起儿童死亡比差异的方向, 表现为虚拟回归系数的正负号, 正号的含义是儿童死亡比高于参照类, 负号为低于参照类。(2) 各变量对儿童死亡比差异的作用幅度, 由回归系数绝对值大小表示, 较大的正值意味着该系数所对应的这类妇女的子女死亡率明显高于参照类。因此, 虚拟回归系数在此处的实质意义, 是在控制其它因素的条件, 以参照类为基准, 比较各类型之间的子女死亡率差异。

二、分析结果

表2列出了各虚拟变量的回归系数, 从中可以观察比较在多元控制情况下各变量的净作用。

母亲民族: 本文以汉族作为参照类, 将少数民族与之比较。我国目前共有55个少数民族, 由于少数民族人口占总人口比例较小以及样本规模有限, 样本中少数民族妇女只有351人, 占8.12%, 因此不可能做很细的分组, 只能将人口最多的几个少数民族单独分组, 其余少数民族合并为一组。

分析结果表明, 代表各少数民族的虚拟变量的回归系数均为正, 这说明在控制其它变量条件下, 少数民族的儿童死亡率普遍高于汉族^①。由于历史的原因, 我国少数民族地区社会经济发展水平远远落后于汉族地区, 死亡率普遍很高。解放后这种情况已经

有了很大的改变,但从分析结果来看,少数民族儿童死亡率与汉族平均水平相比仍然偏高。目前没有证据表明,死亡率的民族差异是由种族或遗传等先天生理因素造成的,通常这种差异实质上反映的是民族间在社会经济、文化、教育、医疗卫生条件等方面的差异。在本文分析中,由于社会经济差异部分地得到了控制,因此,仍然存在的民族差异可能在很大程度上归因于文化、卫生因素,尤其是少数民族特有的生活习俗、育儿方式等。

分析结果还表明,少数民族之间也存在很大差异,其中维吾尔族的儿童死亡比不但远高于汉族,也明显高于其它少数民族。实际上,即使同是在新疆内部,维吾尔族的婴儿死亡率也明显地高于某些其它民族,如汉、哈萨克、回、蒙等民族^①。分地区死亡率比较表明,新疆的婴儿死亡率是28个省、市、自治区中最高的。由于维吾尔族占新疆人口的比例最大,达45.9%,因此新疆的婴儿死亡率高主要是维吾尔族决定的。

父母教育程度:母亲和父亲教育程度这两组虚拟变量都是以“不识字或识字不多”为参照类,所有回归系数都是负值,表明在控制其它因素之后,提高教育程度对降低儿童死亡率仍然有其独立的作用。从回归系数的绝对值来看,两组都呈递增趋势,反映出教育程度与儿童死亡率的反向变动关系。此外,在这种递增趋势中还可以看出,无论母亲还是父亲的教育程度,“小学”类对于降低儿童死亡率的边际作用都是最大的,即父母教育程度从文盲提高到小学水平所引起的子女死亡降低幅度最大。这说明,受过小学教育的父母,已经能相当有效地减少子女夭折。对于降低儿童死亡率,初等教育的投资最小,而收益最大。如果将母亲组与父亲组进行比较,可以发现所有教育程度上,母

表2 多变量分析结果

变 量	虚拟回归系数	T检验显著性水平
NF0	—	—
NF1	1.4487	0.2151
NF2	5.0762	0.0011
NF3	14.9930	0.0000
NF4	2.6085	0.1840
NF5	4.8832	0.0031
NF6	0.3896	0.9079
NF7	1.4196	0.5783
NF8	0.5821	0.7964
NF9	2.1133	0.0263
EM0	—	—
EM1	-1.2082	0.0001
EM2	-1.3938	0.0029
EM3	-1.6216	0.0337
EM4	-2.6388	0.1765
EF0	—	—
EF1	-1.0273	0.0039
EF2	-1.0880	0.0086
EF3	0.8323	0.1599
EF4	1.9929	0.4572
OCM0	—	—
OCM1	-0.3915	0.3372
OCM2	-1.1858	0.0584
OCM3	-2.4885	0.0045
OCM4	-0.7258	0.6511
OCM5	-2.4630	0.0083
OCF0	—	—
OCF1	-1.1082	0.2649
OCF2	-1.4744	0.1589
OCF3	-0.8116	0.5133
OCF4	-2.8852	0.0184
OCF5	-2.6366	0.0192
P0	—	—
P1	-1.4214	0.0210
P2	-2.0401	0.0015
P3	-1.3785	0.0212
R0	—	—
R1	-0.4775	0.3896
R2	0.0342	0.9397
R3	0.9211	0.0493
R4	3.1028	0.0300
R5	2.5933	0.0000
AGE	0.4100	0.0000
CONSTANT	-6.6752	

亲组的系数都高于父亲组,这意味着提高母亲的教育程度对于降低子女死亡的作用要大于父亲。

父母的职业:在父母职业各类别中,设“不在业”为参照类,其余各职业类别虚拟变量的系数都是负值,表明在业父母的儿童死亡率低于“不在业”的父母。父亲职业中儿童死亡最低的是“干部和办事人员”类

① 费梦潮:《新疆各民族婴儿死亡率的精确计算及人口的平均期望寿命》,《中国人口年鉴(1988)》,中国社会科学出版社,第187~191页。

别,母亲职业中最低的是“商业与服务性人员”类别;在父母职业中儿童死亡率与上述两组基本相同的是知识程度较高的“专业技术人员”类别。

如果说职业在一定程度上反映了个人和家庭的社会经济地位状况的话,那么,管理类和知识类职业的低儿童死亡率是与预期一致的。至于商业和服务行业的妇女为何有较低的子女死亡率,这一点尚不清楚。同一种职业状况,对于母亲和对于父亲的含义可能不完全一样,对儿童死亡比的作用也有所不同。父亲的职业也许在更大程度上反映了家庭的社会经济地位,而母亲的职业状况可能较多地与育儿方式相联系。

居住地:目前中国城乡之间在社会经济、文化卫生等方面存在很大差异,不少研究成果已指出这种差异会影响死亡率。这种差异无疑也会反映在儿童死亡率的差异上。在“县”被设为参照类的情况下,其余三个类别的系数均为较大的负值,表明市、市郊、镇这三个类别的儿童死亡率都明显低于县这一类。但在市、市郊、镇这三个类别中,儿童死亡比最低的不是市、而是市郊,其次是镇和市。在部分地控制了其它变量的作用后出现的这种情况,也许反映出影响儿童死亡最主要的不是居住地点本身,而是居住地点所代表的社会经济文化及其它环境条件。实际上每类居住地点的环境都会存在有利方面和不利方面。市区的社会经济条件通常好于市郊和镇,但在城市的某些优势因素作用被控制后,城市变量可能在更大程度上代表其不利因素的作用,例如交通事故造成的意外死亡较多,环境污染导致生态环境恶化等。

地理区域:与城乡差别类似,地区之间的社会经济差异也会影响儿童死亡率水平。在我国,东部沿海地区的社会经济发展水平最高,中部次之,西部最低。此处设“华北”为参照类,将其它地理大区与之比

较。各地域的系数反映出,华北、东北与华东的儿童死亡率差别不显著,中南地区略高,而西南和西北地区的儿童死亡率远远高于东部地区。儿童死亡水平的这种分布与社会经济发展水平的地理分布总趋势是一致的。在控制了民族、教育、职业、城乡等因素后,这种明显的东西差异仍然表现出来。这意味着东部和西部区域性差异中仍包含着许多影响儿童健康的其它因素,由于资料的限制而不能被模型所反映。

母亲年龄:在此模型中,母亲年龄是唯一的连续变量。该变量的回归系数表明,母亲年龄与儿童死亡比同方向变动。由于母亲年龄与儿童年龄显然是密切相关的,这种关系意味着较早出生的儿童经历了较高的死亡率。由于在构造儿童死亡比时,通过对母亲年龄的标准化对子女暴露于死亡风险的时间长度进行了控制,因此,回归系数在此处反映的主要是到普查为止的前一段时期内婴儿和儿童死亡率的下降趋势。

在多元分析中,初始变量被转换为一组虚拟变量,对单个虚拟变量的检查不能完全看出初始变量是否显著,为此,在多元分析中还对各组虚拟变量做了分组检验。检验结果表明,所有的初始变量均具有统计显著性。这些变量对儿童死亡率差异的解释作用相对大小依次为:母亲年龄、母亲民族、户所在地理区域、母亲教育程度、户居住地(市、郊、镇、县)、父亲职业、母亲职业、父亲教育程度。

三、总结与讨论

(一)从上述分析结果中可得出以下几个主要结论:

1.本文分析所用的自变量可以分为两类:父母个人特征和外环境因素。对外在环境因素的两个变量分析表明,东部和西部地区之间,城镇和农村地区之间在儿童死亡率上存在着明显的差异,这在很大程度上反映了户所在地区社会经济发展水平的作

用。

2.就整体而言,汉族妇女的儿童死亡率明显低于少数民族;同时少数民族之间也存在相当大的差异,某些少数民族的儿童死亡率很高,值得进一步研究。

3.父母的教育程度对于儿童死亡率有重要的影响。特别应该指出的是以下几点:

(1)初等教育对降低儿童死亡率的边际作用最大;(2)父亲和母亲的教育程度对子女死亡的作用大小是不同的,母亲教育程度的作用更大;(3)在本文分析中,母亲教育程度是所有可人为改变的因素中作用最大的变量。上述结论具有明显的政策含义,即普及妇女的初等教育是降低儿童死亡率的有效手段。

4.不同职业状况的父母之间儿童死亡率存在差异,部分地反映出家庭社会经济状况对儿童健康的影响,相对来说父亲的职业作用更大一些。

(二)在得出这些结论的同时,应该指出这项分析在几个方面的局限性对分析结果可能具有的潜在影响:

1.本分析中涉及一些父亲特征,此处“父亲”指作为分析单位的妇女在普查时点上的“当前”丈夫。妇女有可能离婚和丧偶后再婚。所以在分析中的“父亲”不一定是死亡子女的亲生父亲。

2.父母的某些标志会随时间而改变,如职业变动、教育程度提高、居住地点迁移等,因此子女死亡前和子女死亡时的父母某些特征与普查时点上的特征未必完全一致,而真正影响子女死亡的是前者。

3.分析的数据来自普查登记,某些可能影响儿童死亡率的个人特征和户的特征,如经济收入、子女营养状况、住房与卫生条件等没有包括在普查登记中。

4.限于计算机的存储容量,分析所用的样本量不足以进行更细的分组,某些变量如职业、民族,因为合并了一些类别而使原始数据信息有所损失,如有条件使用更大的样本做分析,将会取得更令人满意的结果。

(本文责任编辑:徐莉)

(作者工作单位:中国人民大学人口研究所)

(上接第59页)

表10

生育末胎各年龄组妇女生育量比较

年龄组	19及以下	20~24	25~29	30~34	34~39	40~44	45~49	50及以上
妇女数	20	151	318	416	507	484	136	6
生育活婴量	21	299	1013	1768	1762	3181	998	47
平均活婴量	1.05	1.98	3.18	4.25	6.45	6.57	7.33	7.83

从表10看到,末胎生育年龄越高,平均终身生育的活婴数越多,相反末胎生育年龄越低,平均生育子女数越少。而且每隔5个年龄基本上多生一个活婴。这是自然生育状态下不可改变的规律。

(三)婴儿死亡状况

这里所提的婴儿死亡状况不是指某一年内婴儿的死亡状况,而是指所调查妇女生育的婴儿在一定时期总的死亡状况。

在所调查妇女生育的24113个活婴中,一年内死亡的有4376个,死亡率为181.47%。其中,男性死亡率为196.47%,女性死亡率为162.47%。根

据1985年全国第一期深入的生育力调查资料,河北省1965年前婴儿死亡率为80.1%,其中男性为91.4%,女性为68.1%;1970~1974年婴儿死亡率为38.1%,其中男性为41.1%,女性为34.8%;1980~1983年为34%,都比这次所调查的婴儿死亡率低得多。它反映社会主义的新中国比封建社会的旧中国在医疗卫生事业、妇幼保健事业等方面都有了巨大发展,人民生活水平有了很大改善,人口的健康水平提高了。(本文责任编辑:王跃生)

(作者工作单位:河北师范大学人口研究所)