卫生政策研究进展

Progress in Health Policy Research

2009 年　第 10 期
（总第 12 期　医改监测与评价专刊）

上海市卫生局政策法规处
上海市卫生发展研究中心　　2009 年 10 月 22 日

编者按　随着国家和各地医改方案的陆续发布，国际组织对中国医改的关注程度日益提高。医药卫生体制改革是一个动态的过程，监测与评价是其中必不可少的重要一环，尤其是在此次范围广、规模大、内容复杂、投入大的医改背景下，更需要建立强有力的监测与评价体系，以监测改革进展、衡量结果、评价影响，并确保改革预期目标的实现。目前，卫生部统计信息中心正在筹备此项工作，并初步拟定上海、湖北两地为首轮试点地区。本期结合卫生部统计信息中心和世界卫生组织共同举办的“中国医药卫生体系绩效评价研讨会”以及国务院医改领导小组办公室和世界银行联合举办的“中国医改效果评价培训班”的会议精神，邀请胡善联教授撰稿“国际组织对我国医药卫生体制改革监测与评价框架的建议”，在已有资料的基础上整理编撰世界卫生组织“全球视角下中国医药卫生体制改革监测与评价综合框架”，同时撰文介绍医改效果评价的主要思路和方法，为上海开展医改监测与评价工作提供借鉴。谨供有关领导和同志参考。

目　录

专家解读

国际组织对我国医药卫生体制改革监测与评价框架的建议…… 1
1　医改监测与评价框架 …… 1
2　医改评价目的及WHO为中国设计的医改评价框架…… 2
3　欧盟医改评价指标体系 …… 3
4　卫生矩阵网络 …… 4
5　WHO驻华代表处的建议…… 5

它山之石

全球视角下中国医药卫生体制改革监测与评价综合框架…… 6
1　背景 …… 6
2　监测与评价的一般框架 …… 7
3　国家案例 …… 16
4　该框架在中国的应用 …… 22
5　结论 …… 24

医药卫生体制改革

医药卫生体制改革效果评价简介…… 26
1　什么是效果评价及其必要性 …… 26
2　效果评价的设计——因果推断技术工具箱 …… 27

关于上海市拟开展医药卫生体制改革监测与评价工作的思考…… 31
1　背景 …… 31
2　总体原则 …… 31
3　监测评价目标 …… 32
4　监测评价内容 …… 33
5　监测评价的主体与方法 …… 36
6　基本步骤 …… 37
7　工作保障 …… 39

国际组织对我国医药卫生体制改革监测与评价框架的建议

胡善联

《中共中央国务院关于深化医药卫生体制改革的意见》和《医药卫生体制改革近期重点实施方案（2009-2011）》的公布引起了很多国际组织的关注。世界卫生组织（WHO）应中国卫生部领导的邀请于2009年9月为中国提供了“全球视角下中国医药卫生体制改革监测与评价综合框架”。目前，卫生部统计信息中心正在积极组织和筹备开展此项工作。本文拟对国际组织提出的一些评价框架进行介绍。

1　医改监测与评价框架

当今世界对卫生系统改革的评价战略框架大体可以分成三种。第一种是2004年世界银行和哈佛大学所提出的“控制把手（Control Knobs）框架”，即通过筹资、支付方式、组织、规制、行为五个主要的控制卫生系统改革的阀门,影响卫生系统的中间目标(提高效率、质量、可及性)，实现最终目标（健康状况、财务风险的保护、消费者满意度的改善）。第二种是2006年WHO提出的“系统模块(Building Blocks）框架”，即通过加强卫生系统改善人群健康结果。卫生系统的七个模块包括服务提供、卫生人力、信息、医疗产品、疫苗和技术、筹资、领导和治理，通过实现中间指标（可及、覆盖、质量、安全），最终达到总体改革目标，包含改进健康（水平和公平）、反应性、社会和财务风险的保护、提高效率。第三种是2008年WHO提出的“初

胡善联：上海市卫生发展研究中心主任，复旦大学教授

本文系根据中华人民共和国卫生部和世界卫生组织于2009年8月在天津共同举办的“中国医药卫生体系绩效评价研讨会”主要精神综述而成

级卫生保健（PHC）评价框架”，包括四个领域的改革：（1）全民覆盖的改革（包括健康公平性、团结和社会包涵）；（2）卫生服务可及性的改革（包括有效、安全的卫生保健）；（3）公共政策的改革（包括保护和促进社区居民健康）；（4）领导改革（包括建立可以信赖的卫生部门）。尽管有上述内容不同且存在互相交叉的诸多评价框架可供选择，但这并不影响卫生改革的监测与评价工作（如指标的选择和测量需求，分析和资料的应用以及沟通等）。

2　医改评价目的及WHO为中国设计的医改评价框架

医改监测与评价的目的包括三个方面：第一，管理健康投资，监测改革的投入、过程和结果；第二，常规卫生系统绩效评价，这是国家决策的关键；第三，评价卫生改革投资的结果。

为此，WHO为中国当前医药卫生体制改革设计了一个监测和评价框架，包括不断递进的四个阶段：投入与过程—产出—结果—影响（表1）。

表1　中国卫生系统改革的监测与评价步骤

<table>
<tr><td></td><td>投入与过程</td><td>产出</td><td>结果</td><td>影响</td></tr>
<tr><td>指标群</td><td>治理、筹资、基础设施、卫生人力、供应链、信息</td><td>干预措施的可及性、服务意愿、干预的质量、安全性、效率</td><td>干预的覆盖率、疾患的风险、行为和危险因素</td><td>改善健康结果和公平性，社会与财务风险的保护</td></tr>
<tr><td rowspan="2">资料来源</td><td rowspan="2">管理资源（财务跟踪系统、国家卫生账户数据库、卫生人力、基础设施、药品、政策数据）</td><td>机构评价</td><td colspan="2">人群为基础的调查（覆盖率、健康状况、公平性、风险保护）</td></tr>
<tr><td colspan="3">临床报告系统　（服务意愿、质量、覆盖率、健康状况）</td></tr>
<tr><td>分析与综合</td><td colspan="4">资料质量评价、估计与预测、研究结果的应用、进展评价、卫生系统绩效</td></tr>
<tr><td>沟通与应用</td><td colspan="4">有目标性的综合报告、常规的国家回顾过程、全球的应用</td></tr>
</table>

卫生改革评价的短期目标是期望了解投入的增长情况，长期目标则是围绕卫生服务的可及性、质量、覆盖率、财务保护及患者满意度等方面，对健康结果的影响进行评价（如发病率和死亡率降低情况），其反映在医疗保障方面的措施有：（1）新的政策改变（如医疗保险、基本药物、卫生服务）；（2）医疗保险保费或筹资政策的改变；（3）新筹资政策的执行过程；（4）服务的可及性、价格的可承受性和服务利用情况，开放的机构数、初级卫生保健人员数；（5）结果的公平性（结果包括个人自付比例、灾难性支出比例及政策改变），干预措施的覆盖率；（6）病人满意度和反应性。WHO 为中国医改近期五项重点工作设计了总共 122 个评价指标，其中，医疗保险系统改革有 34 个指标，基本药物政策有 11 个指标，医疗服务体系和初级卫生保健改革有 31 个指标，基本公共卫生服务均等化有 27 个指标，公立医疗机构改革的评价有 19 个指标。

国际组织建议中国评价指标应该主要集中在短期投入和过程指标方面，从中筛选出一批核心指标。该组织将 1993-2008 年作为改革前的阶段，以 2007/2008 年的数据为基线资料。

3　欧盟医改评价指标体系

欧盟 27 个成员国的卫生体系各不相同，这与中国卫生体系的情况很相似。欧盟不同国家间健康状况、卫生服务质量和可及性的差异巨大，改善卫生体系绩效的工作显得尤为必要。当前，欧盟正在实施一项监测与评价的综合性长期计划，这将有利于提高优质卫生服务的可及性、减少健康的不公平性、提高患者的满意度、减少危

险因素，这些方面是改革的核心问题。目前欧盟通过征求各国意见已确定了一套核心指标体系，首先确定了一个包括500个指标的长列指标体系，然后结合各国的优先重点领域，缩短为82个短列指标体系。欧盟通过改进生命登记和卫生服务报告，建立了一个EUPHIX网站，提供框架的全方位信息，也包含了投入与过程、产出、结果、影响四个部分，同时对不同国家的指标进行对比，以显示各国卫生改革的短期和长期效果，其主要指标包括:（1）卫生人才（如每10万人口医生数）;（2）卫生支出与筹资;（3）卫生服务的质量和安全（如肿瘤的5年生存率）;（4）卫生服务的可及性（如因经济原因、等待时间、地理距离等延误就医）;（5）干预措施的覆盖率（如疫苗接种率）;（6）危险因素;（7）健康状况与公平性。实施结果证明需要中央政府来建立核心的评价指标集，以提供国家之间的可比性。

4　卫生矩阵网络

卫生矩阵网络（Health Metrics Network，简称HMN）是一个非政府组织，其目的是在卫生改革的范围内推进卫生信息系统建设，建立一个通用的卫生信息系统框架，加强国家卫生信息系统建设（www.healthmetricsnetwork.org）。卫生信息数据资料来源于人口普查资料、居民登记、人口调查、机构的资源记录、服务记录和个人记录。该组织建议中国建立一个数据库，在地区级以及地区级以下建立分类数据，地方具有信息自主权，供众多政策决策者使用，同时应建立个人电子病历的试验基地。卫生矩阵网络目前已覆盖了全球80多个国家，希望加强与中国的联系，创造南南合作的经验。

5 WHO驻华代表处的建议

WHO驻华代表处强调了数据收集的基本策略和收集工作，前者包括从行政管理机构收集的数据、入户调查、常规的临床报告系统、生命统计系统。数据收集途径包括机构调查、定期回顾检查、监测系统和其他专项调查（如患者满意度调查、病人离院调查、高危人群和目标人群调查等）。WHO建议中国确立一个通用框架，将政策目标与数据需求联系起来，建立协调机制以确保各部门与研究机构信息共享，改善数据的可得性和质量（包括信度和效度），并定期测量，建立与政策制定者及公众交流的方法。

全球视角下中国医药卫生体制改革监测与评价综合框架

1 背景

中国在卫生领域已取得了重大进展，期望寿命从1950年的46岁提高到了2005年的73岁就证明了这一点。但是，中国也存在一些公认的严峻挑战，如城乡人口分布、公平性、筹资风险保护和服务质量等方面。中国政府已经认识到了这些挑战，并声明在2020年前要实现“人人享有基本医疗卫生服务”的目标。

有关研究和报告指出，中国卫生系统绩效并没有向人们基于20世纪80年代前取得的良好绩效所期望的方向发展。这些结论是在对国家卫生服务调查和专题研究的数据进行分析的基础上得出的，如医疗费用的迅猛增长导致了医疗服务的低水平利用和医疗费用自付比例的提高，这一点对于贫困人群的影响尤为明显。同时，这些研究报告还持续关注了医疗服务质量问题。

2009年4月，国务院医药卫生体制改革领导小组公布了医药卫生体制改革方案，该方案绘制了中国医药卫生体制改革的发展蓝图，即2020年前实现“人人享有基本医疗卫生服务”的目标。该方案还提出，在未来3年内，中国重点进行 5个领域的重大改革：1) 加快推进基本医疗保障制度建设；2) 建立基本药物制度；3) 提高基层医疗卫生机构的公共服务能力；4) 缩小公共卫生服务覆盖的差距；5) 进行公立医院结构和筹资机制改革试点。在2009-2011的3年期

本文摘自世界卫生组织2009年9月发布的《全球视角下中国医药卫生体制改革监测与评价综合框架》。卫生部统计信息中心对其进行了翻译，本刊编辑人员在此基础上进行编译和整理

间，中央和地方政府将预计投资 8500 亿元人民币（相当于 1240 亿美元），以保证医药卫生体制改革的顺利推进。卫生部统计信息中心承担了医改监测与评价的重任。目前该中心已经拟定 2009-2011 年医改监测与评价框架草案和一套相关指标体系，当前的主要任务是努力设计一个更广泛的长期监测与评价框架。

本文根据卫生部要求，向其报告包括医改效果短期、中期和长期监测与评价的综合框架的进展情况，致力于借鉴世界各地的有效实践经验，并提出适用于中国的有效方案。本文第一部分概述卫生改革和加强卫生系统监测与评价的一般框架，然后据此分析中国在指标确定、数据收集和测量、分析和综合、数据交流等方面的现状，最后为中国医改监测与评价框架提供一系列意见或建议。

2 监测与评价的一般框架

医改监测与评价综合框架的开发需要满足不同使用者的需要和多重目标的需求，其中包括：

- 监测程序化的投入、过程和结果，这是管理卫生系统投资的要求。
- 卫生系统绩效评价，这能够为国家关键决策提供信息。
- 根据健康状况、社会保障和公平性相关的结果评价卫生改革投资效果，以及提出效果评价的最佳方法。

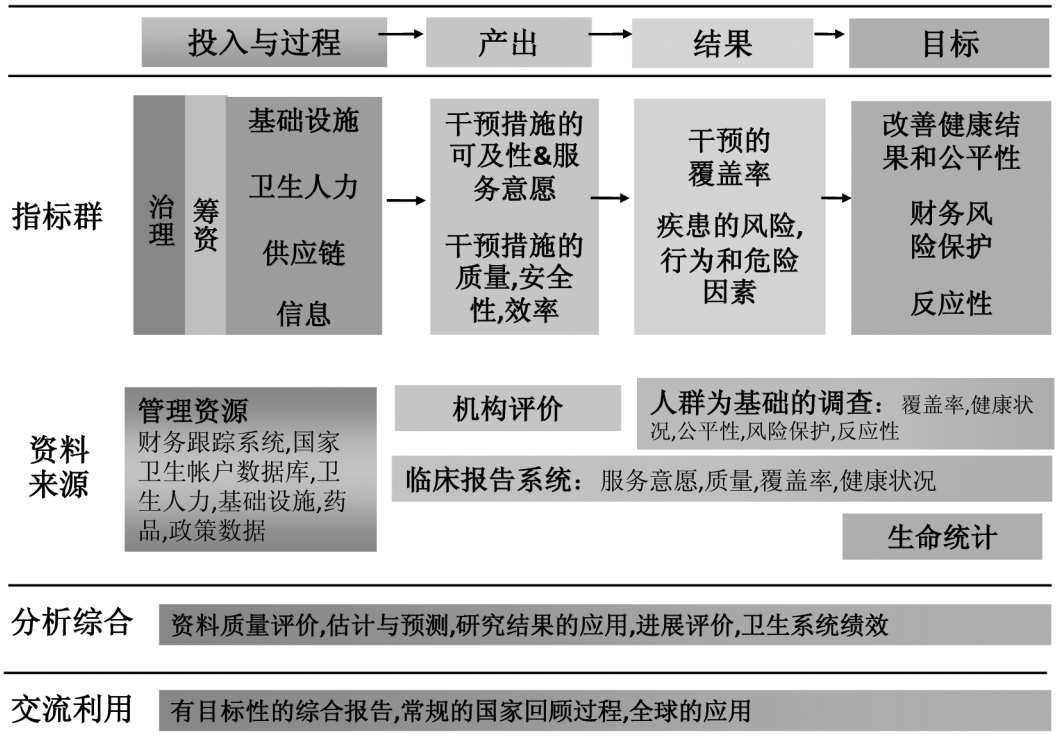
短期和长期数据需求之间的平衡是很必要的。以结果的投资活动为基础的短期数据需求往往关注结果链中投入与过程方面的指标，这需要数据支持，以确保改革政策能够在各地实施，并且帮助地方

管理人员关注改革政策的整体方向。

长期数据的收集需要产出和结果方面的数据——服务的可及性、质量、服务利用、患者负担以及患者满意度，其主要关注改革和卫生系统的整体目标，并且进一步找出差距，以便从长远考虑来加强卫生系统。建立以长期数据需求为视角处理归因问题也是很重要的，也就是说，财政和技术层面的改革有助于改善健康状况，降低发病率和死亡率，提高服务利用者满意度和改善公平性等。在这种评估中，国家间的比较和经验经常被用来衡量绩效状况。

框架不仅需要有利于识别结果或价值链中每个环节的核心指标，还需要将指标与数据集和数据收集方法联系起来，为来自多种数据集的数据分析提供工具指南，同时还要说明数据交流的方式，以便帮助不同层次决策者进行决策和交流。

表 1　加强卫生系统 / 卫生改革的一般框架



2.1 核心指标的定义

一旦对框架和结果链进行界定之后，接下来要对医改监测与评价框架所有领域的核心指标进行定义。这些指标来源于现有表格资料，但是这些指标应该突出政策的重点并且覆盖卫生领域的各种问题，其中包括新出现的健康威胁和健康转型期问题；同样，指标还应覆盖社会和经济中健康决定因素（如教育和收入）。指标的定义应该与全球标准保持一致，并且还应包括所有必要的数据指标解释，以便对所测量的指标进行明确界定。监测与评价者应该根据国家级和省市级的测量能力界定指标的选择和属性（如测量的频率和分解的水平），那些无法统计的指标是没有利用价值的。同其它指标一样，卫生指标也应该具有有效性、可靠性、特异性、敏感性以及可行性／可负担性。同时，数据收集必须与决策制定相关或者适合更高层面的数据需要。

目前，很多国家可以利用经过充分测试的成熟表格，但是它们往往偏向于结果链的特定指标。在某些设置中，指标应该主要侧重于投入和过程，尤其是基于结果的筹资问题的报告；重点是面向结果和影响的指标，目前的挑战是确保全面的平衡；同时，还要关注数据的可用性和质量，尤其是指标的影响，在很多情况下，基线数据往往因其可能影响监测所产生的效果而不可用。

对于疾病方案和卫生体系来说，有很多的指标可用。世界卫生组织（WHO）选取了一套最低要求的核心指标体系（50项），为制定合理公共卫生决策提供依据。各领域指标的选取都考虑到了其科学

的合理性、相关性、为决策服务的实用性、对改变的反应性以及数据的有效性参考。但是，国家不应该只考虑后者，即使数据的可用性是有限的，也能促进国家对核心数据集的开发进行投入，准确跟踪医改绩效。对于每个指标来说，首选数据集应该是最佳选择。我们的目标是所有国家都具备一系列具有核心数据集的数据库。事实上，相对潜在指标来说，核心数据集是很少的。目前的挑战是要确保有一个合适的数据库，以保证数据集和核心指标的高质量和高效率。

最终，每个国家都需要维护指标和元数据的“注册”，这是国家卫生统计观察的需要，并需要对核心和补充指标予以明确和定义，同时要明确数据来源、分析方法和统计值。数据注册有助于数据收集的标准化，并为决策层提供高质量的数据以及方便数据间的比较和数据使用。

信息技术具有重要作用，它能高效地生成数据（如基于病人和医院记录或现场调查的数据收集），并通过不同机构、地区、区域和国家层面的可交互使用数据库达到数据的共享和交换。如果能提供及时和规范的数据，那么来自机构的数据经过处理和加工后，可以为公共卫生提供丰富的信息，这一点已经在某些疾病领域得到证明，尤其是在疾病爆发的报告方面（如 SARS），结核病方面亦然。然而，目前卫生信息化工作还很分散，并且常常出现信息需求与卫生信息技术开发无法衔接的情况。目前，很多国家都在制定卫生信息化开发战略和解决方案，这需要大量资金投入，WHO 正在考虑开发一个

协作中心平台，以便各国共享经验，并为所有国家提供工具指南。

2.2 检查数据来源

接下来要检查各指标的数据来源，以获得符合要求的数据。一般来说，每个指标都有一个首选的数据来源，然而在一些国家，某些重要数据来源（如死亡数据的民事登记）可能是不完整的、非功能性的或者需耗费巨资获得的。在这种情况下，可以利用替代性数据或者利用多渠道分析和重新整合数据，以获得最准确的数据估计值。数据的来源可分为两大类：一类是可以生成整体性的人口数据，另一类则是健康管理相关数据和业务活动的结果数据[1]。更具体来说，框架主要涵盖了五类数据来源：

（1）行政数据：提供投入和过程数据，其中包括国家卫生账户（财政数据）、人力资源数据库和记录、卫生基础设施、药品和商品、现有政策数据和可能实施的政策数据。

（2）机构评估：提供产出的数据，尤其是服务提供和服务质量。这些数据都是通过对机构的抽样调查或对所有机构的普查收集而来。通过访谈和详细记录，可以收集服务有效性和满意度方面的数据，更深入的评估（如离职访谈、陌生客户和医疗供方的观察）对于进一步提高评估质量和服务评估是很有必要的。

（3）临床报告系统：提供产出数据，以及健康结果和影响数据。包括疾病管理和临床咨询的常规报告数据、死因数据和服务提供等常规数据。

[1]Health Metrics Network. Framework and standards for country health information systems. World Health Organization. 2007

（4）人群调查数据：通过服务覆盖和利用情况调查、发病和死亡事件调查、以及生物和临床试验问卷调查提供结果和影响数据。大多数调查是基于家庭的随机抽样，有些是针对目标人群进行的。

（5）生命登记系统：提供影响数据，尤其是有因死亡和出生，这是最理想的高水平覆盖系统。

其他信息来源（如卫生研究、临床试验和纵向社区研究）同样可以反馈到卫生信息系统中。

数据集的最佳选取取决于一系列因素，包括流行病学、使用专门的测量仪器成本和能力问题以及方案的需求等。在很多情况下，资源组合可以促进优质信息的获得和效率。例如，免疫接种率根据年度机构报告来计算统计，常规的关于免疫接种数据的家庭调查将提供一个独立的次要的数据集，以用来验证并且有可能调整机构免疫接种率的估计数据。

一套共用原则可以使得任何数据集都能被用来生成数据包和指标。评价数据集应该检测核心指标体系的延伸程度，以保证数据收集质量，这包括指标的标准定义、合理的数据收集方法以及对指标数据收集和计算的描述方法。

一般来说，可取的方法是最大限度地利用现有数据，提高数据汇编，以便广大用户能通过一种简单的形式来使用并协调数据收集，以减少数据的重复和分裂。

2.3 数据合成与分析

下一个关键步骤就是数据的合成、比较和分析，尤其是审核不

一致且相矛盾的数据，对偏倚和误差进行说明和核算，并进行卫生状况和趋势的一致性评估。这种高层次的分析为疾病负担、风险行为、卫生服务覆盖、指标趋势以及卫生系统行为等方面都做了估计。有些国家使用了省市级的分析。统计分析的复杂程度因国家而异，有必要制定标准化工具，以便能够更广泛地使用这种分析方法。国家学术机构、公共卫生机构和研究机构的介入有助于更广泛地理解这种分析方法，从而提高卫生相关统计的质量。国家、省市级和地区建立共享的信息和数据库是很重要的，这样可以提高信息使用价值，并可能获得高质量的数据分析。

数据的质量评估是很重要的一方面。所有的数据都是有缺陷的，因此对数据进行系统评估是必要的，以此决定是否需要对数据进行调整。其方法包括记录性回顾、不同来源的数据比较（如机构和调查数据）、服务提供和覆盖情况的数据比较（如蚊帐和避孕套的使用情况）、一致性和趋势分析、外部确认（如与其他群体或研究机构的研究进行比较）以及建立模型。该程序应以透明的方式来执行，而加强指标的可信度也是至关重要的。

深入研究选定人群和进一步利用现有研究结果对分析来说是一项很重要的投入，这样的研究能够保证其他数据集产生数据的准确性，同时更深入地洞察某些事件所发生的原因。

数据的合成和分析环节，需要工具支持数据的质量评估、校正偏差以及缺失值的估算和预测，在此过程中信息技术同样扮演着重要的角色；需要持续开发和微调数据收集工具，以便进一步深入分

析其使用者，并在资源有限的情况下多层次、多方面地使用资源。

2.4 数据交流与使用

最后一步是在数据合成分析后，根据利益相关者的广泛需求转换为合适的数据材料，这就要求对数据进行整合、交流和传播，以不同的形式和语言传递给更高层次的政策制定者和决策者，促使信息被卫生体系内各层次部门所使用，如卫生服务管理、卫生系统管理、规划、宣传和政策开发服务部门等。这涉及到一系列使用者，他们来自于不同的技术领域和职业范围，掌握着一系列相关词汇和方法。数据的传播应该尽量满足每一个独立用户的需求，并为每个用户设计提供最有效的数据包和传播渠道。数据开发者应该认真对信息传播进行规划，以满足不同用户的计划周期和需求。一般来说，通信专家能够为不同的用户提供不同的信息包，信息技术则为不同用户提供了有效的数据交流平台。

为保证数据利用的广度，使用者应该充分运用各种方法，如利用目标、省市排名表、分组结果和基准性能等对数据交流进行评估。基于监测与评价结果的补充性工作和定性研究可能有助于定义趋势下的潜在因素，实现与大众的有效沟通，对于显示改革进展或加强影响或激励公众参与关键改革问题等都是很重要的。总之，表达技巧在改善数据交流和传播方面至关重要。

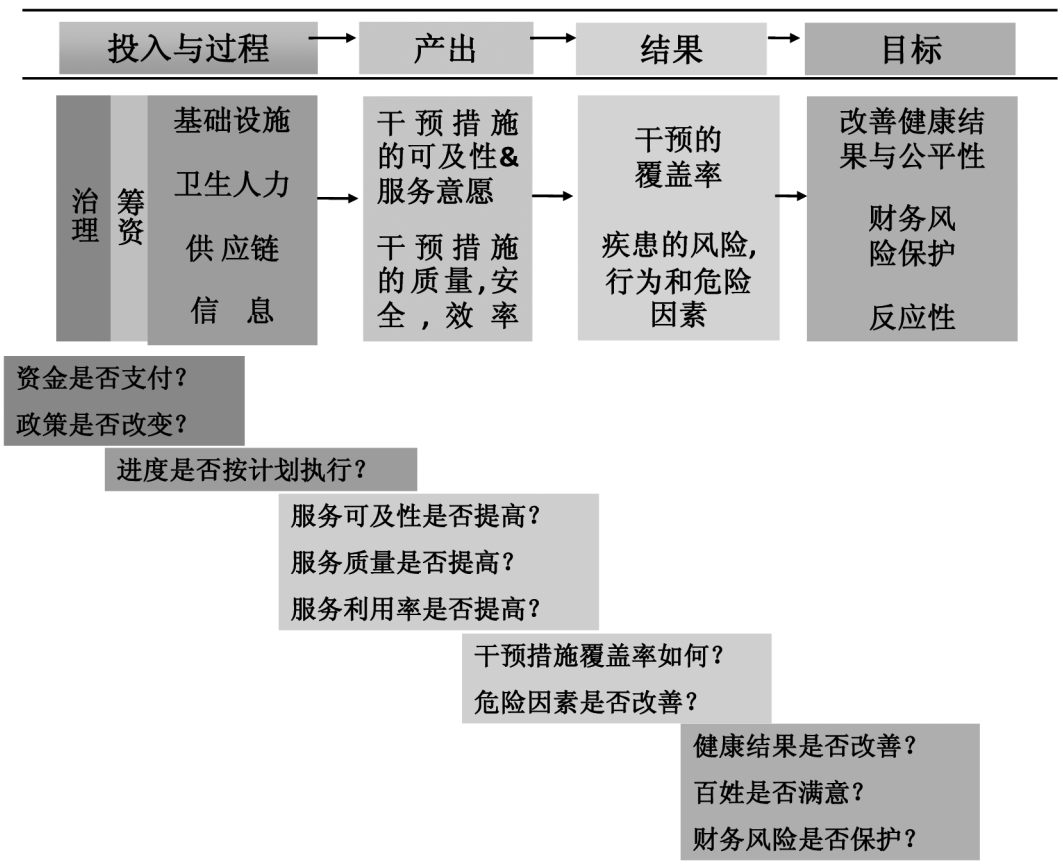
在信息需求和信息推广使用方面宣传信息文化，可以实现加强信息需求和供应以及质量间的动态链接，在实践期间，需要建立机构机制，并采取信息使用的激励措施。国家级的卫生数据分析能力

是至关重要的。有经验表明，最有效的机制是将数据 / 信息与实际资源配置（预算）联系起来，并且开发指标主导的规划。机制的关键之处是明确国家的数据处理过程，并在这些过程中加强数据的有用性、质量和使用性，而不是提出新的数据要求。

2.5 运用逐步法评估影响

卫生改革将会有覆盖全国的不同的优先干预措施，因此，运用干预和人口比较的准实验设计来评估改革的结果和衡量改革的具体方面是不太可能实现的。在这种情况下，随机抽样几乎不可能，大规模的评价研究设计是运用循序渐进的方式评估健康结果、覆盖率和风险行为，可及性和服务质量，以及干预费用。

表 2 递进法评价卫生改革



第一步，首先根据潜在逻辑性，跟踪卫生改革的资源；第二步，充分利用不断增长的资源、费用、人口和商品等，努力扩大服务；第三步，改善服务的可及性和质量；第四步，加大干预曝光率（即干预覆盖率和降低风险行为发生率）；第五步，如果资源充足，加大干预措施覆盖率对健康的影响效果将取决于干预效率（在理想的情况下的工作效率），以及效果转化为效益的相关因素（如干预措施的覆盖率和质量）和环境因素（如流行病学、经济变化、政治稳定）。因此，最终的疾病趋势变化需要利用这些因素加以解释，与此同时也必须考虑环境变化（如经济状况和政治稳定情况）。第六步，将发病率和死亡率水平与前四步联系起来进行评估。健康结果被定义为：在所有措施和方案的联合作用下，发病率和/或死亡率产生的可测量或估计的总体效果，还包括人们对卫生服务和财务风险保护的满意度。中国现有指标的使用可以显示医改所希望带来的影响。投入、过程和结果指标可以基于可获得的基线（趋势）数据和额外资源的水平来设置相对准确的目标。尽管上述非方案因素可能会影响目标的实现，但是目标对于结果和影响指标是非常重要的。

3　国家案例

本部分主要从机构设置、数据收集和分析等角度出发，选取部分国家作为典型案例，描述它们各自在卫生系统评估中的不同做法和实践。

欧洲共同体健康指标方案（ECHI）

欧共体约有5亿人口，各国的卫生体系和卫生状况大不相同。

因此，卫生状况和趋势以及改革情况的监测，需要一个通用的框架、指标体系和共同的报告机制。欧共体健康指标方案(ECHI)的目的是：根据对卫生和卫生决策点[2]的一个概念性的想法，为公共卫生领域开发指标清单。指标的选取使用了以下四个标准：

第一，**综合性：**包括公共卫生领域的所有方面。

第二，**满足用户的需求：**指标集需要覆盖委员会和成员国公共卫生政策的主要方面。

第三，**创新性：**数据集应不仅仅为数据收集，还必须明确其收集的必要性。

第四，**基于前期工作：**利用欧洲统计局和其他委员会等主要数据提供方的国际机构研究成果，同时尽可能多地采纳经济合作与发展组织和世界卫生组织欧洲区域所确定的指标和变量标准。

欧洲共同体根据涵盖欧洲公共卫生信息（EUPHIX）的数据交换网站制定网络纲要（ICHI，健康指标国际纲要）。短列表包括了40项指标，综合列表则包括了80多项指标，其中部分指标还需要进一步细化（如性别、经济状况，医学原因）。对所有的指标都有一个宽泛的元数据描述，指标关注点不应局限于指标比较，而应扩大到产出、可及性和质量、覆盖率、危险因素及其决定因素和健康状况等指标。

该方案对数据质量和各国之间的统计比较方面给予了很多关注，但是很多指标没有数据可比性。对于某些比较来说，运用不同的方法收集数据是有问题的。

[2] http://ec.europa.eu/health/ph_information/dissemination/echi/echi_en.htm

墨西哥

墨西哥卫生系统绩效评价的目的是衡量覆盖的有效性，以及根据卫生系统实际提供的干预措施所获得的潜在的健康收益范围，将所有的覆盖率数据整合为一项单一措施，并且计划在全国范围内实现卫生总费用核算的全覆盖，为国家提供卫生系统效率情况亮点的评估。本项目由墨西哥卫生部（卫生信息办公室）、国家公共卫生协会和哈佛大学[3]共同实施。

该国数据的可用性和质量主要受益于全国卫生服务调查资料(包括生物标志物)，其数据主要来自于2000年、2002年和2005-2006年的全国调查、行政登记和相关重要统计。

墨西哥采用了一种测量有效覆盖率的综合方法，对每个州（共34个）实施了8项干预措施的有效覆盖性和18项干预措施的覆盖率的调查。数据的可用性和质量限制了理想方法的应用，而且每项指标的精确程度随数据集质量的变化而改变。视力障碍的有效覆盖被定义为：超过20岁的有视力障碍的成人戴眼镜的比例或解决他们的视力问题的比例。在2005-2006年的调查中，视力障碍矫正的需求是从自我报告有视力困难的数据估计而来的。服务的利用情况是通过眼镜使用情况的自我报告和使用效果估算得出的，服务质量则是通过对近视和远视者使用眼镜情况和解决视觉问题的困难性的两项数据估计得出的。在调查中如果调查对象认为两个问题的报告都没有问题，则被视为有效覆盖。对视力问题的评估的最佳客观方法

[3] Lozano R et al. Benchmarking performance of Mexican states using effective coverage. And Lancet Series. 2006

就是自我报告。

印度尼西亚

2001 年该国根据 WHO 卫生体系框架，进行了一次地区级的卫生体系评估，该次评估运用了一个由长期收入、女性受教育程度、每 10 万人口护士数、个人现金卫生支出和卫生机构可及性（通过紧急医疗服务利用测量）[4] 组成的投入指数，其重点是针对地区进行评估。评估面临的主要挑战是每个地区的测量手段都是不确定的，地区是一个相对较小的分析单位，人口越少的单位越要依赖于行政和卫生机构的数据，因此数据质量经常存在问题，而从省级水平进行的数据分析就极少会受到这些问题的影响。如果可能的话，可以运用模型来填补数据差距。产出指数由免疫覆盖率（已完成的）、专业接生、碘盐含量、灾难性支出和期望寿命等组成。

南非

2003 年，南非全国卫生行动要求全国各省和各地区在开发和执行年度计划时，必须衡量计划的年度实施情况并监测所取得的短期目标。卫生服务统计是各个地区的主要数据基础。非政府组织——卫生系统基金会开发了包括一个小指标集的“健康晴雨表”，其指标几乎完全来自于地区机构卫生信息系统（DHIS）[5] 报告，包括卫生服务投入、产出、结果和影响指标，以及卫生决策指标的选择。年度报告是按各地区的绩效情况排列的，并将数据整合在表中予以公

[4] Soeharsono S. Indonesia subnational health system performance assessment report. Presented at the meeting. Available at www.healthmetricsnetwork.org. Tandon A. Measurement issues regarding estimation of subnational health system efficiency. Presented at the meeting. Available at www.healthmetricsnetwork.org

[5] Monticelli F. South Africa’s District Health Barometer. Presented at the meeting. Available at www.healthmetricsnetwork.org

布。另外，该报告对选取的农村点和低收入城市人口进行了一次专门比较，并将所有的绩效数据与全国平均水平进行比较。

一般情况下，DHIS 生成的数据的质量似乎并不高，因此该国需要通过不断的努力去调整已知偏差，以提高数据质量。另外，该国需要追加投资，以提高基于人口数据的覆盖率和健康结果（如民事登记数据）等指标。

加拿大

加拿大在卫生信息系统的开发方面进行了大量的投入，以确保卫生系统的独立评估、数据的质量控制、数据的隐私保护、信息技术的使用和系统化的绩效分析。加拿大卫生信息研究所（CIHI）是一个独立的、非营利性的组织，它为加拿大卫生系统以及加拿大公民的健康提供主要数据，并分析相关数据。加拿大卫生信息研究所之所以在很多领域能够跟踪数据，这要归功于医院、地区卫生行政单位、医疗参与者和政府提供的信息，而其他数据来源则为 CIHI 更深入的数据分析奠定了基础。CIHI 的数据和报告主要关注卫生保健服务、卫生支出、卫生人力资源和人群健康问题。最近加拿大人口健康运动主要关注个人和社区的健康问题的影响因素，并致力于政策制定以减少不公平性。

CIHI 同样明确了国家卫生指标，推动了国家卫生指标的使用，比较卫生状况、卫生系统行为和特点。卫生指标的目的是：支持区域卫生管理部门监测人口健康和本地卫生系统的功能。每年印刷出版的卫生指标包括加拿大地区至少 5 万人口的卫生统计数据，涵盖

了大约 98% 的人口、省和地区。加拿大开发了卫生系统框架，目前已有 70 项指标，覆盖了基本卫生保健（14 项）、家庭保健（2 项）、其他方案和服务（13 项，大部分是保健质量）、灾难性药物管理和处方管理（1 项）、诊断和医疗设备（3 项）、人力资源（2 项），以及健康状况和风险因素指标（35 项）[6]。这些指标借鉴了医疗和人口等基线数据，同时对医疗质量方面的研究数据也有所采用。

泰国

有证据表明，泰国在运用循证决策和建立公共卫生运动方面取得了成功。泰国健康促进基金会（Thai Health）是一个独立的准政府机构。该机构有专门的预算资金（主要来自烟草和酒精的销售收入）确保其财政独立。Thai Health 的任务是通过与国内社团和公立或私立组织合作提倡和鼓励促进人群健康，并提高各个社会层次的循证决策能力，其所采取的战略是运用知识创造、社会动员和政治参与共同联合来解决困难，这被形象地形容为“三方联合能移山”。

任何情况下都会面临的挑战是克服方案管理人员夸大成功和淡化失败的心理倾向。泰国试图通过在每一个计划层次营造“信息文化”氛围，而省略服务规范方面的监测与评估。建立于 1992 年的卫生系统研究院（HSRI）不仅要评估卫生系统和政策，还要报告国家卫生目标的进展情况。虽然 HSRI 接受来自政府的资助，由独立委员会管理，但它是独立于政府的合法机构。因为 HSRI 拥有广受尊敬的资深学术成员，所以它在创造知识和将知识转化为政策等方面都做出了

[6] http://secure.cihi.ca/cihiweb/dispPage.jsp?cw_page=prtwg_2009_e

重大的贡献。

卫生信息系统开发办公室（HISO）获得 HSRI 和 Thai Health 的共同支持，得出重要卫生议题的分析结果（如地区间死亡率差异）。

4 该框架在中国的应用

专家组在“医药卫生体制改革监测与评价方案”（7 月 8 日）的草案中，制定了监测与评价框架，阐述了草案的主要内容和相关指标，关注了改革的 5 个优先领域。监测与评价的范围如下：

（1）监测 5 个重点领域，23 项政策措施和任务的进展情况。监测进度指标共计 130 项，包括加快建立基本医疗保障体系的 34 项进度指标，建立国家基本药物系统的 11 项指标，完善基层医疗卫生服务体系的 31 项指标，促进基本公共卫生服务均等化的 27 项指标，推动公立医院改革试点的 19 项指标，另有 8 项监测保障措施的实施情况的指标。

（2）通过设计专门的医改评估年度调查表，来评价社会公众对医改的满意度。

（3）根据结果评价医药卫生体制改革绩效。包括 3 个整体目标（改善居民健康状况、提高疾病风险分担能力和提高社会公众满意度）、4 个阶段性目标（提高服务的可及性、公平性、质量和效率）和卫生体系的 5 个关键部分（卫生筹资、科学支付、组织监管、医患关系等）。

（4）开展深入的定性研究，以深化对改革效果影响以及案例研究的剖析。

根据中国医药卫生体制改革的内容，专家组明确监测与评价框架的具体内容如下：

（1）包括从指标的选择、测量到分析、整合、交换和使用各方面指标。其中，中国应关注在数据收集和分析过程中不同部门行动者所扮演的角色。

（2）投入与过程、产出、结果和影响方面（包括短期和长期影响）。

（3）关注小范围的核心指标，并以其他指标补充作为补充。

（4）关注特殊疾病方案和卫生体系两个方面。

（5）从亚国家级（省市级）层面考虑。

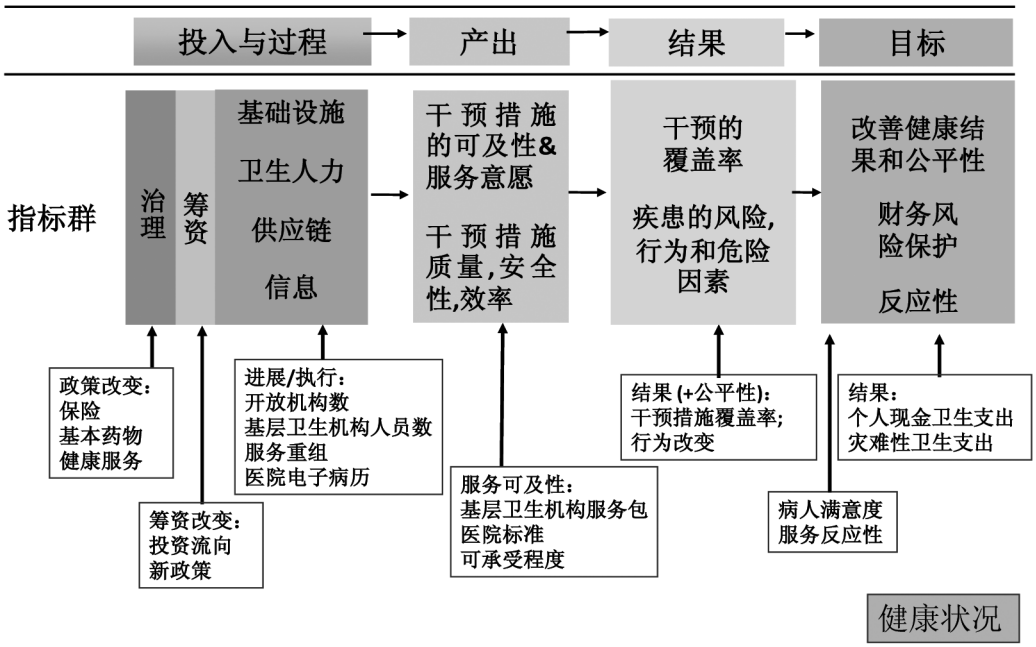
（6）明确分析工作，包括不同来源数据的评估、调整以及整合。

（7）与国际标准相衔接。

表 3 主要是卫生改革的 5 个优先领域，并展示了这 5 个优先领域和监测与评价框架的对应方式。这里同样有一个时间因素，显而易见，框架的右边部分需要很长时间才会产生很大改变，而框架的左边部分在短时间内就能产生结果。对于 5 个优先领域来说，框架左边和中间指标的短期改变是可以预见且可予以设定目标，如果加大改革执行力度并富有成效，甚至框架右边部分的患者满意度和灾难性卫生支出的风险防范等都是可以在中短期内实现的，而主要目标健康状况的改变将耗时最长。

接下来的部分主要描述框架每个领域的监测与评价状况，包括整体目标、改革的影响途径、数据集的当前状态、指标和数据收集的国际案例以及加强短期和长期数据收集的相关建议。

表 3　中国医药卫生改革优先领域的监测与评价框架



5　结论

- 准确评估卫生改革的短期影响是有可能的，但是监测和记录卫生改革长期影响的工作任重而道远。
- 建议使用一个融合卫生体系改革所有主要因素和指标的综合监测与评价框架开展监测与评价工作。
- 指标：平衡所有框架领域，包括短期和长期目标，优先使用现有统计指标，因为它们能够提供数据的基线水平和趋势；同时尽可能地使用国际指标，以加强国际间的可比性。
- 数据收集：有效的数据来源主要依赖调查、短期干预，甚至可扩及生物和临床手段，同时辅之以临床和管理数据库。
- 死亡率：为建立一个完整的死亡登记系统，需根据亚国家级

数据需求，完善妇幼保健监测系统，扩展生命登记系统，并扩展和加强人口监测站点系统。

（编辑：上海市卫生发展研究中心　荆丽梅　张　进　王月强　彭　颖）

医药卫生体制改革效果评价简介

1 什么是效果评价及其必要性

按照世界银行的定义，效果评价（Impact Evaluation）是指对社会政策给个人、家庭、社区和企业所带来的福利改变的评价。效果评价的核心问题是“得到政策干预与没有干预相比有什么不同。”由于很难直接观测受干预的群体与未受干预的群体，并将二者进行比较，所以效果评价的最大挑战是设立参照对象，即寻找在各个方面都与干预对象尽可能近似的群体。效果评价研究可以建立确定的因果关系，去除混杂因素的影响，将观测到的福利改变归因为政策干预。

效果评价的目的是为了提供反馈信息，以协助优化项目和设计政策。效果评价除了能够提高相关者的责任意识以外，同时也是一个保持动态学习的工具，决策者可以通过这个工具改善正在推进的项目，最终在项目推进过程中能更好的优化资金使用。其他项目评价方法还包括组织评论（Organizational Review）和过程监控（Process Monitoring），但这些方法都不能通过明确因果关系估算效应的大小。因果关系分析是非常关键的，能够帮助决策者理解不同政策在减贫等社会政策中的选择作用。

效果评价信息帮助决策者制定延续、修改、或终止的政策，并能够用来确定公共政策出台的优先顺序。效果评价是通过回答以下

本文系结合国务院医改领导小组办公室和世界银行于2009年7月在北京共同举办的“中国医改效果评价培训班”上的报告材料综述而成

问题提高政策的实际效果：项目是否取得了预期的目标？试点是否应该推广？某个大型项目是否应继续？社会效果是否可用项目的实施解释，或者源于其他同时发生的因素？项目的效果在目标受益群体（男性、女性、本地人）、地区和时间这几个维度上是否有变化？项目是否有正面或负面的非预期效应？项目的效果与其他选择相比如何？项目的效果与其耗费的成本而言是否物有所值？

2　效果评价的设计——因果推断技术工具箱

传统的描述性统计数据可以进行“前后对照”、“有无比较”等评价，但都有一定的局限性。“前后对照”设计是指接受政策干预的同一人群的前后横截面数据的对比，“有无比较”是指接受政策干预群体和未接受政策干预群体的比较。这两种方法都无法控制其他混杂因素的影响，前者无法控制随时间变化的其他效应，后者无法控制选择性偏倚。世界银行推荐的因果推断方法克服了上述方法的局限性，被广泛运用于全世界发展项目评估。

2.1　结果链

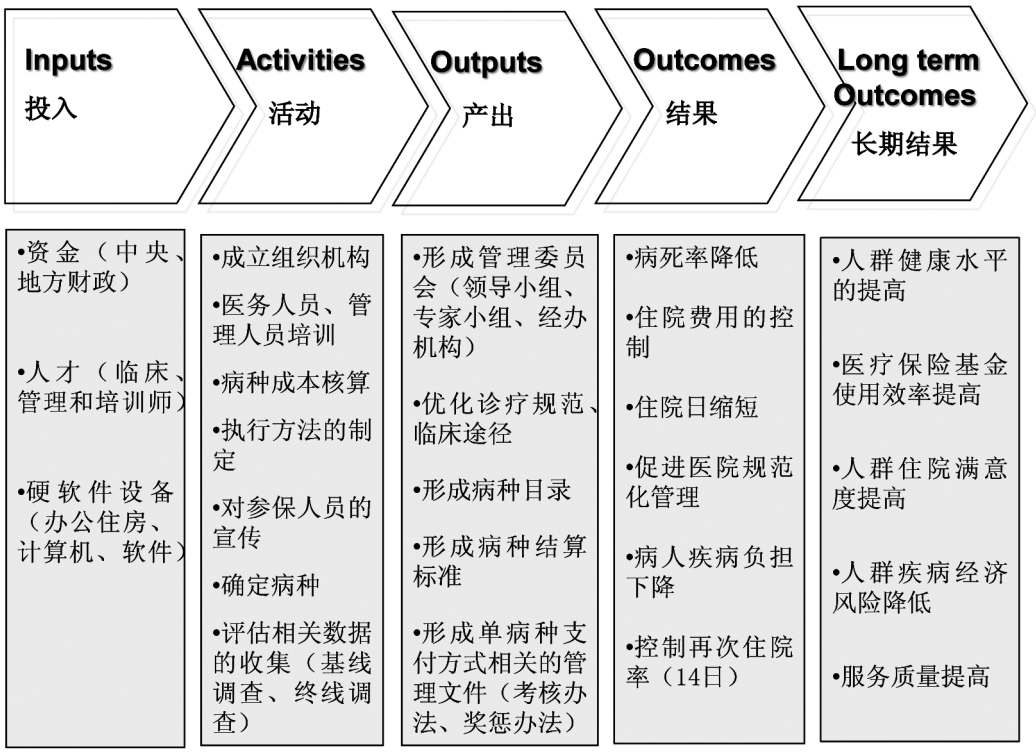
结果链（Results Chain）是描述政策干预在实施中逐渐发生作用并产生效果的过程的模型，包括投入、活动、产出、结果、长期结果几个部分。表 1 以医保住院支付改革为例，描述了政策发生效果的过程。

2.2　随机试验

随机试验（Randomized Trial）通过使用两个步骤建立模型。第一步是从确定的人群（总体）中随机抽取一个样本；第二步是将

这个样本中的个体随机分配到试验组和对照组。第一步确保了样本的外部有效性，即确保抽样误差在一定范围内，样本的结果能够代表总体；第二步确保了内部有效性，即确保因变量的观测结果是不受其他混杂因素影响的，仅由干预手段得出。

表 1　医保支付改革发生效果的过程



2.3　工具变量

工具变量（Instrumental Variable）是统计学、计量经济学、流行病学等学科中的研究工具，可用于不可能设计随机对照实验时的因果推断研究。即当随机解释变量与随机误差项相关时，寻找一个与随机解释变量高度相关，但与随机误差项不相关的变量，用该变量替代模型中的随机解释变量，进行模型参数估计，这一替代随机解释变量的变量为工具变量。例如，若评估吸烟对健康的影响，

除了吸烟以外，还有很多因素可以影响健康，甚至健康状况也可以影响是否吸烟。此时进行随机对照试验是很困难的，可通过观察烟草产品的税率来研究因果关系。因为在其他影响因素固定情况下，烟草税率对健康的影响必须通过影响吸烟行为完成，烟草税率和健康之间的相关关系就是吸烟影响健康的有力证据。因此，我们可通过烟草税和吸烟之间的关系估算吸烟和健康之间的关系。

2.4 界断点回归

界断点回归（Regression Discontinuity）是指将政策干预的目标人群根据是否纳入干预的标准进行排序，对前 50% 实行干预，对后 50% 进行对照。观测到的两组间的差值与基线相比就是政策的效果。该方法仅估计项目在切点周围产生的效果，因此不能总被外推广到一般人群。由于效果是在切点处进行估计的，所以一般而言，用这个方法评价所得的“效果”比相同样本量的随机试验所得到的观察值要低。

2.5 倍差法

倍差法（Difference-in-difference）通过监测和比较干预组与对照组的基线水平和结局的变化评价干预效果。干预的效果 = $(Yt_1-Yt_0)-(Yc_1-Yc_0)$，即干预组受干预后与基线的差值与对照组观测值与基线的差值的差，也叫双差法。倍差法的基本假设是干预组和对照组的趋势一致，需要至少三个时间点验证并估计干预效应，其中两个在干预前，一个在干预后。倍差法使用对照去除了其他随时间变化的效应的影响，相比“前后对照”更科学。

2.6 匹配法

匹配法（Matching）是指从调查中选取样本，使其观测到的影响因素与干预组一一匹配，而不存在选择偏倚，即要求找到基本特征与参与者完全一致的未参与者，比较两个群体在干预前后的差异，这种要求显然较高。Rosenbaum 和 Rubin 改进了这一模型，使用倾向计分匹配（Propensity Score Matching，PSM）的方法，测量干预组和对照组“是否参与”的概率，即参与倾向，再估算两个样本中参与的 Logit 或 Probit 值，限定样本以保证共同的子集，观察研究中重要的偏倚来源。为每个参与者找到一个倾向评分相同的未参与者，比较结果指标，其差别就是项目获益的估计值，然后根据每个个体获益的平均值计算出总体获益。

（编辑：上海市卫生发展研究中心 林海）

关于上海市拟开展医药卫生体制改革监测与评价工作的思考

1 背景

医药卫生体制改革作为一个连续动态的过程，应当遵循科学制定政策的逻辑顺序，即发现主要问题－寻找主要原因－设计若干方案－制定改革政策－实施改革政策－进行绩效评价，其中，监测与评价具有重要的理论研究与实践意义。政府卫生投入、改革措施落实、医务人员积极性等一系列涉及卫生改革的实施情况都需要一个循证、科学的监测体系进行量化，从而有利于决策部门及时调整和完善政策措施，确保医药卫生体制改革有计划、有步骤的顺利推进。此外，这个体系也将不断地为各级政府部门、相关管理机构和人员提供有效信息，实现评价结果的转化。

上海经济比较发达、医药卫生体系较为完善，这为上海医药卫生体制改革的平稳开展及其监测与评价指标体系的全面构建提供了有利条件。本市可以率先开展医改监测与评价的试点研究工作，为全国或其他地区提供可资借鉴的前瞻性经验。

2 总体原则

上海市在建立和实施医药卫生体制改革的监测与评价体系时，需要遵循以下原则：

（1）目的性原则：监测与评价体系设计必须紧紧围绕着上海市医药卫生体制改革的总体目标和阶段性目标，着眼于人人享有基本

本文系结合中华人民共和国卫生部和世界卫生组织于2009年8月在天津共同举办的“中国医药卫生体系绩效评价研讨会”报告的主要精神编撰而成

医疗卫生制度目标的实现，着眼于本市人民群众最关心、最直接、最现实的利益问题（如“看病难、看病贵”）的解决。

（2）针对性原则：监测与评价内容设计必须与本市医药卫生体制改革方案（尤其是当前五项重点改革）保持一致性，客观反映深化医药卫生体制各项改革任务和具体措施的落实情况。同时，该设计应满足本市各级政府、卫生行政部门、卫生服务机构在改革进程中的职责发挥和责任落实情况的评价需要。

（3）可比性原则：监测与评价指标设计必须从上海市社会经济文化背景和实际需要出发，既要保持历史数据、改革前后和改革过程监测与评价指标的可比性，又要注重总结和整合国际经验，应该采用国际上或全国通用的卫生体系与卫生改革监测与评价的一般框架设计，以确保关键性指标具有国内和国际可比性。

（4）可行性原则：监测和评价指标和方法设计要易懂、易掌握、易收集，尽可能用客观指标而非主观指标；要充分利用已经建立的卫生信息系统，与现有统计体系、调查制度相衔接。

3　监测评价目标

医改监测与评价工作要充分体现长远性、系统性和阶段性的要求，分别设定长远和总体目标、近期目标和具体目标。

监测与评价工作的长远和总体目标是：通过建立和完善医药卫生体制改革监测与评价体系，系统收集信息、客观反映进展、深入分析问题、科学绩效评价，全面评判医药卫生体制改革的总体目标和阶段性目标是否实现，覆盖城乡基本医疗卫生制度是否建立健全，

各项改革的任务和策略是否落实，人民群众是否得到安全、有效、方便、价廉的医疗卫生服务，城乡居民多层次、多元化的医疗卫生服务需求是否满足，居民健康状况是否改善。

监测与评价的近期目标是：根据国家医改监测与评价指标体系和《医药卫生体制改革近期重点实施方案（2009-2011）》（以下简称《实施方案》），设计上海市监测与评价框架与指标体系，客观反映国家医改工作和上海市重点改革进展和任务的落实情况，以及改革的政策措施对城乡居民“看病难、看病贵”问题解决的影响。其具体目标如下：

（1）系统了解和掌握上海市贯彻《中共中央国务院关于深化医药卫生体制改革的意见》（以下简称《意见》）与《实施方案》（尤其是近期五项重点改革的进展和具体任务）的落实情况。

（2）系统了解和掌握社会、城乡居民对医药卫生体制改革政策、措施落实的反应性（包括公众、医务人员和各级政府部门的满意度）。

（3）系统了解和掌握上海市各区县在落实《意见》与《实施方案》过程中结合本地实际而探索的有效实施途径，通过剖析典型案例，总结可以推广和借鉴的经验。

（4）按照《意见》与《实施方案》的要求，对五项重点改革的绩效进行系统评价，总结经验，找出差距，为进一步推进改革提供参考。

4　监测评价内容

根据监测与评价的目标要求，当前监测与评价内容应该主要围

绕《实施方案》提出的五项改革重点，包括可行性评价、进度评价和结果评价三个方面。

（1）可行性评价（条件评价和事前评价）

在上海市医改方案正式启动实施之前所进行的评价，其重点是评价医改方案与社会经济政策、社会经济文化发展水平的适应状况，拟解决的卫生问题与人群的客观需要的适应状况，医改实施可能产生的风险及其防范措施。

（2）进度评价

在上海市医改方案实施后到总结评价之前对医改方案的实施进度与过程进行的监督与控制，即对医药卫生体制改革近期重点工作进展进行监测。评价的核心内容是评价贯彻《意见》和《实施方案》的进展和任务落实情况，同时对影响医改方案实施的有关因素进行分析，主要包括以下六个方面：

第一，加快推进基本医疗保障制度建设的进展，包括基本医疗保障的覆盖面、水平，基金管理和管理服务水平，以及城乡医疗救助制度的建设情况等。

第二，建立国家基本药物制度体系，包括基本药物目录遴选调整管理、建立基本药物供应保障体系和建立基本药物优先选择和合理使用等。

第三，健全基层医疗卫生服务体系，包括机构和人才队伍建设、补偿机制及运行机制的转变情况等。

第四，促进基本公共卫生服务逐步均等化，包括服务项目的覆盖、

重点公共卫生服务项目的开展及相关能力建设等。

第五，推进公立医院改革试点，包括公立医院管理体制、运行监管机制、补偿机制以及电子病历、常见病临床路径和多元办医格局形成等。

第六，保障措施，包括组织领导、财力保障、政策调整和舆论宣传等工作。

（3）结果评价

在医改方案实施一个阶段以后（2009-2011 年）对其是否达到目标和结果进行的评价，包括以下三个方面：

第一，社会公众对医药卫生体制改革满意度的社会学评估。该评估通过问卷设计、开展定期现场调查等方法，分别围绕人民群众对医药卫生体制改革总体感受、基本医疗保障制度、国家基本药物制度、医疗卫生服务体系和基本公共卫生服务均等化等五个方面进行评估。

第二，医药卫生体制改革阶段性（2009-2011 年）绩效评价。该评价的目的是了解和掌握医药卫生体制改革所取得的阶段性的成绩、效果及其影响，发现问题，分析问题，从而为决策部门及时调整和完善政策措施提供循证依据，以保障医药卫生体制改革有计划、有步骤的顺利推进。绩效评价的方法将借鉴国际上比较成熟的经验，结合上海市医改的实际需求，重点围绕三个总体目标（改善国民健康状况、提高疾病风险分担、提高社会公众满意度）、四个阶段性目标（增强卫生服务可及性、改善公平性、提高卫生服务质量和效率）、

五个卫生体系建设的关键环节（卫生筹资、科学支付、组织管理、监督规制、行为教育），设计医药卫生体制改革绩效评价的框架、内容和指标。

第三，医药卫生体制改革典型案例调查。该调查的目的是通过科学设计，进行定性调查，及时了解、系统掌握各地在贯彻落实《意见》和《实施方案》过程中的亮点，探索有效实现途径和做法。该调查通过典型案例调查与分析、抽象与总结，形成可以借鉴和推广的经验，以点带面、推动改革的深化和发展。

5 监测评价的主体与方法

根据政策方案的评价主体，评价大致分为内部评价和外部评价两类。内部评价是指参与医改方案实施各方的自我评价。这种评价的资料获取便利，评价结论可以直接反馈给制度执行者，但难以保证评价的公正客观。外部评价是指由医改方案制定和执行机构之外的评价者所完成的评价。外部评价组织一般能独立开展评价工作，有利于保证评价的客观公正性，但全面获取资料较难，且评价结论难以直接反馈给制度执行者。为了保证医改评价结论的客观公正和系统全面，兼顾参与医改方案实施过程的各方主体，比较理想的医改方案实施效果综合评价应采用内部评价和外部评价相结合的方式。

依据医改监测与评价的内容，本市应采用定量与定性相结合的评价方法。医改定量分析主要是通过分析指标体系的构建完成，而指标体系的设计应涵盖医药卫生体制改革近期五项重点工作内容，阶段性绩效评价内容和城乡居民满意度等内容。而定性评价主要通

过对政府组织管理方、医疗机构服务方和公众代表进行访谈，到相关机构进行现场观察和体验式观察，以及对有探索性、开拓性和创新性，同时具有借鉴和推广价值的典型案例进行调查与分析，以抽象与总结的方法完成评价。

6 基本步骤

6.1 准备阶段

（1）确定开展医改监测与评价工作的主体，分析监测与评价工作的可行性

该阶段应成立包括领导小组、专家技术支持小组和监测与评价实施小组在内的评价组织，结合相应背景分析开展医改监测与评价工作的可行性。

（2）制定医改监测与评价计划

第一，确定医改监测与评价的目的。通常评价的目的主要是判断医改预期目标的实现程度。（关于医改监测与评价的总体目标和具体目标在前文中已阐述）

第二，确定监测与评价对象。根据医改评价的目标及内容，评价对象应包括城乡居民、医改方案执行与监督机构及其工作人员，以及政府、医药企业、医疗机构等医改的利益相关者。评价对象的确定需要明确抽样方法，保证样本具有较好的代表性，能反映总体特征。

第三，确定监测与评价的指标体系。确定监测与评价的指标体系应当满足指标选择的有效性、可行性、灵敏性、特异性和可靠性

等特征，坚持全面性与系统性、科学性、客观性、可操作性和实用性的指标体系确定原则。通常就医改监测与评价而言，指标越多，反映问题相对越全面。但是并非指标越多越好，指标过多意味着收集资料难度增大，误差也随之增加。因此，在众多指标间需要进行筛选，原则上应当以最少的指标反映最多的问题。

第四，确定评估方法。首先要确定资料收集方法，根据各个指标的特性确定资料收集方法。通常情况下，除已有统计资料、文献等，医改监测与评价指标测量值的确定多依赖于社会调查实现。其次要确定分析方法，医改监测与评价的基本方法包括前后对比分析法、专家评定法和自我评价法等。

第五，挑选和培训评估人员。挑选合适的评估人员，并对其进行严格的一致性培训，这是保证资料可靠性的重要质量控制措施。一致性培训的主要内容应包括监测与评价的目的、意义、调查方法和评价方案等。我们建议制定一份完整的调查指南。

6.2 监测评价实施阶段

监测与评价的实施是既根据既定计划现场进行资料收集，又在收集资料的过程中发现问题，并对信息去粗取精、去伪存真的过程。资料收集后要根据评价目的进行相应分析。收集资料过程一定要注意质量控制，保证资料的准确性、可靠性。

6.3 评价总结阶段

此阶段主要是在综合分析工作的基础上，明确评价结论，完成评价报告。根据分析结果得出评价结论是一个推理演绎的过程，不

同的人对同样的分析结果可以得出不同的结论。因此，有必要对决策者、研究者各方的思维及行为特征加以分析，确保评价结论科学合理、客观公正。依据完成医改监测与评价报告的目标，本市应重点关注以下两项工作：一是报告是否客观公正地完成了既定计划目标、指标体系和评价内容；二是报告是否规范，是否符合基本的格式要求。

7　工作保障

（1）本市各级政府要加强对医药卫生体制改革监测与评价工作的领导。上海市医改领导小组要将医改监测与评价工作列入重要的议事日程，给予人、财、物的支持和保障，同时加强舆论宣传工作，推动该项工作的有效实施。

（2）本市各级卫生行政部门要做好上海市医改监测与评价试点研究的具体实施工作。积极与市发展与改革、财政、劳动保障、民政和药品监管等有关部门协调沟通，做好数据采集单位的层级督查，确保数据的准确可靠。

（3）本市医药卫生机构在信息收集中要实事求是、责任到人，采取网络直报上报数据，并应及时准确上报，不得迟报、误报、拒报、瞒报、假报，否则，一经发现，将严肃追究有关单位领导和相关人员的责任。

（4）对于上海市医改监测与评价试点研究执行中所发现的问题及时上报相关主管部门，并确保得到强有力的技术保障和有效的业务支持。

（编辑：上海市卫生发展研究中心 何江江）

欢迎批评指正，不吝赐稿

来稿来信请寄:

上海市北京西路 1477 号 801 室，邮编：200040

电子邮件：limei_jing@163.com

网址：www.shdrc.org（上海卫生发展研究网）

联系人：荆丽梅

电话：021-22121872

传真：021-22121879

内部资料，免费交流