

第二十七章 监督与评估

第一节 监督与评估概述

近年来，监督与评估作为项目的重要组成部分，越来越受到关注。Michael Quinn Patton 在《定性评估研究方法（第二版）》中提到，“当一个人在进行成功和有效性检测或判断的时候，他就在从事评估工作。当这种有效性的检测是根据经验系统开展的，即通过细心地收集数据并对其充分分析，那这就是评估研究…… 评估是应用性研究，或者说是一种‘行动科学’，这一点将评估研究与基础的学术研究区分开来…… 应用性研究和评估的目的是让人们采取行动，提高决策力，用知识解决人类和社会问题……”

政府和项目投资者需要通过监督评估，了解数据报告的准确性、及时性和可比性，以确保项目能够达到预期目标，并保持资金的持续获得，更为重要的是因地制宜地利用信息，强化正在进行的项目。对于项目实施或管理者，需要通过监督评估获取所需的数据信息，及时对项目的开展作出调整，以保证项目按计划执行并做出科学的技术性决策。艾滋病项目的监督与评估是在特定的政治、经济和社会环境中，对监督与评估技术的应用。

本节主要讨论与艾滋病项目相关的监督与评估基本概念及其应用。

一、监督与评估概念

项目监督与评估就是通过科学的、可以接受的方法，系统地评价项目的操作和结果，并与某种标准进行比较，描述和解释项目的规划、执行过程和成效，并以此作为项目改善的依据。它是保证项目按计划、要求顺利进行的关键，也是一种重要的管理手段。

监督是指通过记录、定期报告、监测系统以及健康机构的观察和调查，对项目开展的关键要素（通常为投入和产出）进行常规跟踪，包括在项目执行中对执行活动、阶段性成果的评估，也包括项目执行活动中的信息搜集、整理、管理和应用。监督帮助项目管理者确定需要开展工作的领域和方向，并发现可改善的领域。在好的监督与评估系统中，监督十分有助于评估的开展。监督也可用于测量一定时间内的趋势，但

采用的方法应一致、严格，以确保能够进行比较。

评估是指对项目介入产生的目标结果的改变所进行的阶段性评价。换言之，评估是在一段时间后将特定的产出或结果直接与干预措施相联系。此外，评估也应将项目的产出与国家的方针政策和其他结果的趋势以及疾病影响相联系，即使项目仅作为影响疾病总体举措的一部分，对其评估也十分重要。

评估主要针对项目效果，监督针对项目执行的过程和环节。为了有效地遏制艾滋病的流行，提高艾滋病病人的生活质量，需要对项目实施有效的监督与评估，以确保项目的执行没有偏离项目的设计。监督与评估的目标和采用的方法不同，就所用方法的严格性来说，评估较困难、昂贵，尤其是结果和效果评估常常要求进行人群调查或其他严格的研究设计。评估还应考虑其他相关数据和调查资料，如人口健康调查、出生登记、死亡登记或哨点监测等数据。

二、监督与评估目的

艾滋病项目评估是一项目的性极强的研究活动，其首要的目的就是要“应用”，通过对项目的理论、规划、过程、产出、影响进行评价，为改善艾滋病项目的决策提供依据。艾滋病项目评估的问题与一般社会科学研究的问题也不相同，它不是来自研究者的头脑中，而是来自将要、正在或已经完成的项目中。艾滋病项目评估对结论的要求非常明确，即要对项目的各种状况进行判断，以有利于推动艾滋病防治工作。

不同的评估主办方进行评估的目的不同，如为了改善项目的设计、规划和管理，为了了解项目的进展状况等等。在理想的状况下，项目相关各方总是希望通过评估获得信息，以便于进行判断或者证实自己对项目的判断。但是，各方对信息的期待是不一样的。如资助方希望知道自己的钱是否花得值得，管理方如疾病控制部门希望检验自己的项目理论，执行方希望知道管理方是否满意。

当然，并不是所有希望进行评估的机构都可以开展项目评估。由于评估成本高昂，多数情况下，只有项目的资助方和管理方才有这样的能力来委托专门的评估机构进行评估。这样的机构通常包括国际合作领域的各种资助机构，如国际组织、基金会，还有政府机构，当然也有项目的管理机构等。

艾滋病防治项目评估一般的目的是通过对项目的考察来对项目进行决策：执行、继续、调整、选优、终止、扩展。与许多评估一样，对艾滋病防治项目进行评估可能是因为决策者或管理者希望知道项目到底进展得如何，以便于对项目进行调整；或者是政策的制订者希望知道政策的功效如何，以便于考虑政策的走向。所有这些都是为了决策的目的，所以，通常又被称为决策性的评估。在决策性评估中，最常见的是用于中期调整的评估。评估的一个重要目的就是要考察项目的进展，以便对项目进行调整。评估的任务除了了解项目的执行和进展以外，更为重要的就是发现偏离项目目标的状况或发现项目设计的缺陷，及时纠正或调整。如某个艾滋病关怀项目是对已经出

现艾滋病相关症状的病人展开关怀，让病人定期到项目点接受心理辅导等。但在项目实施一段时间以后，由于当地文化传统的压力，很少有人主动去项目点，由此有可能因项目设计的缺陷而使项目陷入困境。通过中期评估发现此问题后，在项目活动中增加了上门服务的内容。另一种最常见的决策性评估就是中期决断。仍以上述项目为例，如果评估发现，即使增设上门服务也难以接触到项目对象，整个项目无法继续，项目资助方就需要迅速决断，终止项目。或者，经过评估发现，项目活动执行的非常顺利并有意想不到的影响，譬如通过项目活动使得社会对艾滋病患者的态度大为改变，项目就需要继续下去，甚至可以尝试向其他地区推广或扩展项目活动的内容如增加更多的项目活动等等。

三、开展监督与评估需要考虑的问题

由于艾滋病防治项目活动的一些方面与现行的国家法律、社会习惯、文化传统之间的关系更加复杂，评估者在选择评估策略时，需要考虑以下因素：

1. 评估是否值得开展

由于艾滋病防治项目的目的在于通过项目干预艾滋病相关人群的行为，进而减少艾滋病的传播，降低艾滋病病人的身体和心灵伤痛，改善人类的社会环境，因此，评估者在进行抉择的时候，首先要考虑的是，是否值得去做？如果评估活动和评估结果的应用能够吸引更多人对问题的关注，并有利于问题的解决，那就值得去做；如果能够为更多的人带来利益，那也值得去做。

2. 评估为谁服务

正如上面说到的，评估结果的应用也是需要协调的问题。事实上，不仅是评估相关方，许多其他的人群也会用到评估的结果，特别是评估资助方、项目规划者、管理者、实施者、咨询机构、项目对象、项目潜在对象、媒体、学术机构等等。因此，评估者除了确定是否值得做以外，还必须非常明确地知道评估的结果到底为谁服务。同时，评估者还必须注意评估结果的应用是可以控制的，并不像一般学术研究那样，结果一旦发表，其应用就成为了不由作者控制的准公共物品。换句话说，在谁资助与为谁服务发生冲突的时候，评估者可以通过对结果应用的控制来协调。

3. 评估是一个有条件的社会实践

基本的条件包括时间、经费和知识。很多评估都有明确的时间要求，譬如在项目的某个阶段必须拿出某类评估结论。评估的费用也是一个重要的约束，希望评估能兼顾的目的越复杂，涉及的问题越多，所需要的费用就会越昂贵；对评估结果的精度要求越高，对知识的要求也就越复杂，评估所需要的费用也就越昂贵。

4. 评估目的是否具有可协调性

在既定的约束条件下，如果评估目的众多，评估者是否承担评估工作的抉择还取决于评估目的之间是否是可协调的。并且，通过沟通和解释，是否可以解决多目的之间的兼容性，使得协调以后的评估目的在给定的条件下可以完成。

5. 结果应用的可能性

对于评估者的抉择而言，评估结果的应用也是一个重要因素。如果评估结果不能获得应用，就是评估的最大失败。评估的着眼点就是通过评估，使得遏制艾滋病流行的干预活动变得更加有效，进而改善人类的社会环境。如果评估结果不能获得应用，评估本身就失去了意义。

对于艾滋病防治的决策而言，评估结论不是唯一的依据。评估结果要获得应用，就必须使得评估结果与决策之间具有：a. 意识形态（ideology）的一致性；b. 利益和兴趣（interests）取向的一致性；c. 信息（information）的充分性；d. 制度（institution）的保障，即评估结果应用的“4个I”。

第二节 监督与评估的基本要素

艾滋病防治项目的监督与评估与其他社会项目的监督与评估没有本质的区别，基本的原理来源于实验设计。但由于艾滋病防治项目对象的特殊性，严格运用实验原理会遇到前提条件的约束，为此，需要运用多种方式进行条件替代，以具体灵活地解决监督与评估中遇到的问题。

一、实施监督与评估的基本条件

对艾滋病防治项目实施进行评估是有条件的，否则是无法进行监督与评估的。实施评估的基本条件包括：a. 有正常规划和实施的项目；b. 有可以协调并达成妥协的项目目的；c. 有项目各方对评估的支持和帮助；d. 有足够的可支配的资源包括经费、时间和团队。

但是，即使所有上述条件都能够获得满足，评估者仍然需要实地考察项目的可评估性。项目的可评估性，又叫作评估性评价，就是系统地考察（无论是理论上的还是经验上的）是否可以对项目进行全面评估。

按 Joseph S. Wholey 的说法，对评估时机的把握可以通过以下的方式进行：a. 项目是否按照计划在执行；b. 项目运行是否稳定；c. 项目是否已经获得了一些正面的结果。通常，对这些问题的回答可以通过两种策略来进行：

第一，把项目希望达到的产出与实际获得的产出进行比较。这样做的目的是为了建立评估框架，即在项目活动与项目结果之间建立可以通过实证资料进行检验的关系。在评估框架的建立中，和获得项目目的一样，评估者也需要与项目的关键方面一起共同协商，建立项目目的、项目活动以及活动与目的之间的逻辑关系，分析理论要素之间的逻辑关系是否有可能导向预期的项目目的。

第二，搜集可能得到的资料，与项目的关键方面如项目资助方、管理方、执行方、项目对象、旁观者一起，讨论已有的资料是否支持从项目活动到项目目的的逻辑关系，

由此来检验项目理论、考察项目运行的正常性，甚至探讨项目成功的可能性。

还需要注意的是，评估性评价并不仅是看项目可否进行评估，多数情况下，评估性评价就是对项目的一个初步评估。如果初步评估的结论是，既不能形成评估框架，也不能看到项目的一些正面结果，评估者就有责任将初步的结论告知项目相关的关键方面，说明项目的基本状态，给项目一个改进的时机。如果初步评估的结论是项目是可评估的，那么，评估者除了了解项目的进展以外，还需要把评估中可能遇到的问题清理出来，以便评估设计和实施中能够有效地获得相关的信息。

二、监督与评估的指标与测量

在艾滋病防治项目的监督和评估中，指标就是展示项目进展和产出的“率”、“量”、“值”等。在一般意义上，指标是变量属性的表达工具，也是用来衡量事物发展变化的工具。测量则是获得变量属性特征值的研究活动。所谓变量属性特征值，就是与项目相关事物发展变化的程度。以“安全套推广”项目而言，基本的指标就是目标人群的“覆盖率”，覆盖人群中安全套的“正确使用率”。在使用安全套的人群中，由于使用安全套对艾滋病传播所带来的影响，如“感染率”或“阻断率”，这些“率”就是针对项目不同阶段的产出。

（一）评估指标

对艾滋病项目实施评估的指标从项目理论中来，所有项目问题都需要转化为指标才能够获得答案。而指标的多少取决于项目的性质和对监督与评估的要求。一个多目标的项目常常不可能用单个指标来测量；而单个目标的项目则可以用多个指标来测量。和搜集资料的理念一样，变量的测量也不是指标越多越好。监督与评估中，每个指标都需要有相应的实证资料支撑，多一个指标就多出一系列的工作，这就意味着需要花费更多的时间、费用和人力资源成本。

如果项目稍复杂一些，就会发现评估的某个产出变量常常需要多个指标来测量。例如，在艾滋病病人临终关怀项目的监督与评估中，就需要多个指标来测量：a. 对病情控制的效果；b. 病人生活质量的改善程度；c. 病人的生活满意度的改善程度等。采用其中的任何一项指标都不能很好地评估这类项目的产出。事实上，每个指标只能提供某个特定视角的信息，用于测量项目产出或项目过程的一个特征。如果某个变量具有多种属性，就必须采用多指标的综合测量。对艾滋病防治政策倡导的评估也是如此。

但是，这并不意味着每个评估中涉及的变量都需要采用多指标测量，有些时候，指标的概括性往往能够帮助评估者简化工作。如艾滋病防治工作的总体产出，虽然涉及防治工作的多层面、多维度，但我们可以用一个概括性较强的指标来测量，那就是艾滋病的流行状态；又如“安全套推广”项目，如果仅是希望了解项目活动是否能够（而不是在多大程度）干预性工作者的“安全性行为”，那么，任意一个与安全性行为有关的指标就足矣。再者，即使是复杂的项目如临终病人关怀，如果开展监督与评估

的目的仅是了解项目活动是否覆盖了足够的项目对象，那么评估者也可以用项目活动的覆盖率来对项目进行评估。在开展覆盖率评估时，许多情况下，很难确定用于评价的分母，因此，主要寻找分子，或从干预中寻找受影响或受益人群的子集。如有可能仍应包括分母（如果已有百分比，也应报告分子，以评价不同时间段和不同人群的覆盖率）。

（二）评估指标的测量

在项目的监督与评估中，针对具体项目活动的各种监督与评估问题，最终只有落实到指标测量上，才能够获得结论。

进行指标测量的办法有两个，一是运用已有的测量；二是建立新的测量。最有效的方法就是寻找类似的项目，从类似的评估中，甚至不同项目的类似问题中进行借鉴和移植；在无法借鉴已有测量的情况下，评估者需要建立新的测量。在监督与评估中，针对不同的指标，建立新测量有两种不同的策略，具体包括：

1. 建立临时性测量

某些变量的属性是变化的，很难建立稳定的测量。例如艾滋病临终关怀项目，在没有抗病毒药物的条件下，对临终关怀是一种测量。当抗病毒药物有效地延长了患者生存时间的时候，对临终关怀的测量就变化了，从生命的维持延展到生活质量和幸福感等。但是，针对艾滋病临终病人，并没有成熟的生活质量或者幸福感的测量指标，在这种情况下，就要建立临时性的、策略性的测量。

2. 建立可积累的测量

如果某些变量的属性是相对稳定的，就可以建立积累性的测量。如针对艾滋病病人的临终关怀项目，医护人员的可及性就是一个稳定的测量。无论关怀项目的实施历时多长，能够方便和有效地获得治疗，就需要有便利地接触相关医护人员的途径。

需要注意的是，新的测量并不是一蹴而就的事情。一个有效的测量往往需要长时间的积累和发展，甚至很多人的智慧投入，甚至只有如此才能获得相应的效度和信度。并且，在评估中，测量问题几乎发生在项目的所有环节中，尤其是与结果相关的环节中，从投入到产出，从结果到影响。

三、成功的监督与评估系统具备的特点

各个国家有着不同的监督与评估需求，艾滋病疾病负担和国家健康结构可部分地说明这一点。正如许多国家成功的项目所证实的那样，成功的艾滋病监督与评估系统应具有如下一些共同的要素：

1. 专门的监督与评估部门

- 设在卫生部内，并有指定的技术和数据管理者。除其他工作外，该部门应结合国家更广泛的数据需求。
- 提供次国家级、区域和省份的监督与评估指导原则和指引。

- 提供监督与评估与其他部门，如教育、民政和公安等部门进行联系的指导原则。
- 监督与评估预算应介于各种来源的 HIV/AIDS 总预算的 5% ~ 10%，参考值为平均 7%。
- 国家监督与评估预算中相当一部分来自国内资金（而不应完全依赖外部基金资助）。
- 建立正规的监督与评估的工作联系，特别是与相关的部委、非政府组织、捐赠者以及致力于加强应用性研究的国家级研究机构。
- 一个由多部门参与的工作组，提供投入并在指标选择和监督与评估设计及实施的各个方面取得共识。
- 监督与评估部门或隶属于该部门机构的专业知识：流行病学、行为/社会科学以及资料处理和统计学、资料分发、资源跟踪（财政资源和物品资源）等专业知识。

2. 明确的目标

- 具有明确目的、目标、操作计划且清晰定义的国家规划或项目，国家监督与评估计划应每 3 ~ 5 年修订一次，操作计划应每年更新。
- 对国家规划或项目计划的实施进展根据目标进行定期检查/评估。
- 协调国家和捐赠者的监督与评估需求。

3. 评估指标

- 不同级别的监督与评估需要优先考虑的指标及附加指标。
- 不同时间内具有可比性且具有明确目标的指标。
- 与其他国家具有可比性的许多关键指标。

4. 数据收集和分析

- 国家级数据收集和分析的总体计划，包括数据的质量保证。
- 在各行政级别（包括地理区域）监督与评估收集数据和定期分析指标和相关资料的计划。
- 第二代监测，监测中行为数据与疾病监测数据相关联。

5. 数据的发布

- 国家级数据分发的总体计划，并及时提供自由、透明的基本数据集。透明度是真正负责性的基本要素。
- 广为分发、信息含量高的年度报告。
- 与决策者、制定计划者和实施者沟通并讨论监督与评估及研究结果的年会。
- 负责生成和传播研究结果的信息交流机构。
- 收集所有有关 HIV/AIDS 的中心数据库或资料库，包括正在进行的、透明度高且公众可用的研究信息。
- 国家和捐赠者监督与评估需求方面的协调。

6. 专项调查

- 选择优先成果/评估调查。
- 包括所需的定性调查。

- 包括操作性研究的调查。

第三节 评估的策略方法与实施

一、完整的国家项目评估框架

我们需要对艾滋病防治项目开展的不同时期对其进行评估，而且各阶段开展的评估应该统筹考虑，因为单一的数据收集方式不可能提供促进项目执行或影响政策改变的所有必要信息。多种评估和使用的多种方法（定性和定量）可以相互补充，而且满足于不同的评估需要。

1. 需求评估

需求评估始于艾滋病防治项目的计划阶段（或再计划阶段），用于在项目普遍实施之前确立或决定要干预和评估的问题。在此阶段，项目的伸缩性很强，而且项目实施方很容易对下一步如何进行做出决定。

需求评估可探究干预的需要，为确定项目的目的和目标提供必要的信息，有助于项目计划者尝试制定有效、可行的干预策略以及如何执行的决定。需求评估可在项目执行的同时作为探索性工具，为项目主管提供反馈意见，通过调整项目目标改变现状。需求评估研究能够发现不适宜的或无效的干预方法、设计和概念。

由于艾滋病形势的紧迫性，许多预防项目没有深入开展需求评估即实施干预措施。在减少 HIV 性传播的社区干预项目中，特别容易缺少这种评估。很多情况下，干预措施是基于非干预措施实施人群的一些主观想法设计而来的。在一些涉及行为改变干预措施的文献中，充满了一些理论上十分完美的想法，但在“现实世界”中根本无效，主要是由于这些想法不能被目标人群所接受或者用能够让他们可接受的方式进行宣传。充分地理解出现的问题可以更好地使项目策划者重新设计干预项目，从而使项目更吸引目标人群。现在，需求评估被越来越多地应用于预防项目的设计中。

需求评估使用多种研究方法，能够为项目设计者迅速地提供有关信息。这些方法包括回顾现有的信息、小组讨论、个人深入访谈、参与观察和使用问卷开展短期定量调查等。

需求评估最常见的方法学问题是缺乏外部有效性且不能外推。由于需求评估的结果来源于小规模快速评价过程和/或试点调查，我们不能将它外推到更大的人群。虽然有这些限制，需求评估研究通常依然能够发现不适宜的或无效的干预方法、设计和概念。但是，在项目设计阶段，即使开展了充分的需求评估，仍不能保证一个预防项目最终的有效性，也可能是因为项目没有很好地执行。

2. 过程评估

项目一旦开展，就需要检查是否正确、按期以及按照预算开展。过程评估用于回

答一些基本问题，如“计划开展的干预活动实际上实现的程度如何？”以及：“提供了什么服务？给谁？什么时候？频率如何？多久？多大范围？”等等。投入（有关人力、资金、材料和时间等的基本资源）和产出（直接的服务促进，如分发物品、培训人员、设立服务机构）是过程评估的两个重要的因素。回答这些问题常常用量化的指标。用定性的证据解释是否一个预防项目成功或失败在回答过程评估问题中同样重要。过程评估需要使用数据、熟悉项目的细节，不但要观察预期效果，还要观察不曾预见的结果。了解干预活动起作用的过程能够帮助我们解释干预的结果，但过程评估不能说明干预活动是否有效果。

过程评估在促进或调整干预活动中通常起到重要作用。在不断改变的流行状况中，它通过提供必要的信息来调整干预策略或项目目标。过程评估贯穿于项目的整个执行过程，而且通过不同的方法学途径评价提供的服务。这些方法既包括查阅服务记录和常规的报告系统、重要的调查访谈、服务接受者的推出访谈、“假扮顾客”直接观察（例如性病服务和自愿咨询检测服务）等定性方法，也包括采用人群调查评价项目覆盖率和利用障碍的定量方法。多种定性和定量研究设计相互补充，提供了较全面的信息。

3. 有效性评估（评价结果和影响）

评估艾滋病预防项目的有效性通常也需要定量测量方法。这些测量方法用于评价项目目标实现的程度。有效性评估用于回答如下一些问题：“观察到什么结果？”“结果说明什么？”“项目改变了什么？”

考虑艾滋病预防项目执行过程包含多个时期，而且新的年龄组不断成为性活跃年龄组，有效性评估最好分为两层，即短期和中期项目效果（项目结果）评估和长期项目效果（项目影响）评估。在减少 HIV 性传播的干预活动中，艾滋病相关态度的改变、危险行为减少和采取保护性行为以及性病发病率改变被认为是最适宜的测量短期或中期（也称为“近期”）结果的指标，长期效果包括对艾滋病流行趋势、关怀支持和促进社会反应的影响。

结果和影响评估与过程评估密切相连。过程信息能够帮助评估者了解干预活动如何以及为什么起到了效果，或许还可以了解是什么原因起了作用。如果不评价项目执行的过程，检查结果或影响指标将会导致关于干预效果的错误结论。

评估的目的和目标被定义为，选择适宜的结果和影响测量方法评价艾滋病预防项目的有效性。干预的结果和效果可使用具有定量值的指标测量，有效性评估通常基于这些指标。多种干预措施协同作用对于行为改变的促进最为有效，因此，调查不应设计为仅测量一种干预措施的效果。更确切地说，评估应测量暴露于多项联合干预措施的人群组的行为趋势。如果要对一项干预措施进行评估，通常要开展严格且昂贵的试验。

4. 绩效评估（成本效益分析）

成本效益分析也用于测量项目的有效性，但它通过增加测量单位效果的项目花费扩大了分析的内容。通过比较花费和多项干预措施的效果，成本分析和成本效益估计

有助于对优先设置、资源分配的决策和开展项目设计。

绩效评估是对项目进行成本 - 收益分析和成本 - 绩效分析。在艾滋病防治项目评估中，绩效评估有时会作为影响评估的一部分，但在一般情况下，会与过程评估和影响评估一样，作为整个项目评估的重要组成部分。在艾滋病防治项目的评估中，评估者不仅要考察项目的产出，也需要考察构成产出的成本。如20世纪80年代初期，中国实施了针对吸毒人员的强制戒毒政策。到20世纪90年代后期，不少公共卫生和药物滥用方面的力量开始对这项政策的成本提出质疑，认为强制戒毒的成本过于高昂，需要采用其他方式进行替代。

在所有的评估类型中，绩效评估的技术性比较强，对基础条件的要求也比较苛刻，如要求开展前有过程评估和影响评估。绩效评估的基本思路是对项目的成本与收益或绩效进行比较。显然，要获得成本的信息，就需要知道成本的构成，而要做到这一点，对项目过程的评估就成为必须，否则评估者无法计算成本。同样，要知道收益或者绩效，就需要知道项目的影响。在这个意义上，开展绩效评估的项目必须是有条件进行过程评估和影响评估的项目，即只有规范的、稳定的、成熟的项目或模式化的项目，才有条件进行绩效评估。

二、评估效果的区分

由于时间、可用资金数、培训的人员数量等多种因素的限制，项目管理者 and 评估的策划人在选择如何对一个特殊的评估或项目进行评估时，必须在理想、最好和可行、有用、相关及基本之间进行平衡。有效地区分评估效果的方法就是从3个层面进行区

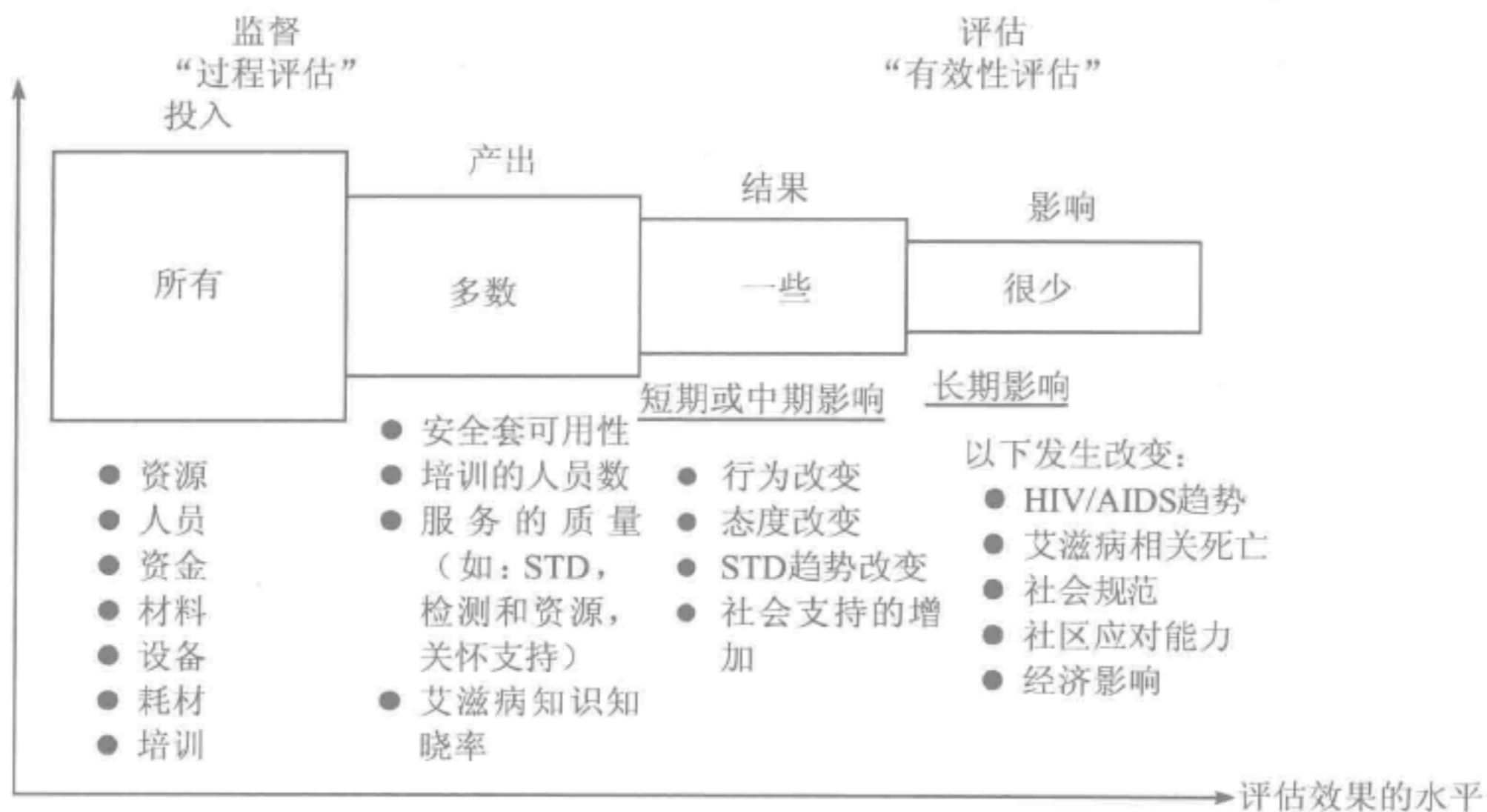


图 27-1 监督评估管道模式

分，即单独的项目层面、国家项目层面以及国际层面。使用多层的方法可以设定评估项目严格程度的优先顺序。

我们可以考虑单个项目作为干预措施实施的范围，多数情况下，不需要严格地研究设计，除非此项目用于新预防措施试点或回答未回答的研究问题。单独的实施标准干预策略的项目已经被证明在类似的条件下也是有效的，因此，应重点开展需求评估、过程评估和能力建设评价。监督和评估“管道图”说明，只有较少的项目才能保证开展有效性评估。见图 27-1。

在国家项目层面，应该强调评估的几种分类，即干预结果、社会经济影响和社会规范的变化。当多项干预措施实施于重叠的目标人群时，指导原则是：一些评估不适宜在单独项目层面开展，而是应通过国家或地区项目协作开展。使用这样的评估策略，特别是在行为调查领域，不仅能节约资金，还能够起到单独项目无论如何也起不到的效果。国家项目层面的评估包括（但不限于）不同人群的行为趋势分析和与之相关的 HIV/STI 监测数据分析、安全套使用和针具交换等社会市场营销评估、性病病征管理、对国家项目总效果进行评分、社会经济影响评价以及国家艾滋病疫情的流行病学模型拟合。

国际层面的效果评估可以说明一些不确定的问题，如哪些预防干预措施最有效，以及在什么机构、对谁和在什么情况下最有效。但是，这种类型的评估需要大规模的社区基础的对照试验来完成。显然，这些试验超过了单独项目甚至国家项目的范围。由于直接测量 HIV 预防项目影响的大规模发病率研究难度较大且需要较多的资金，现在更加强调发展其他方法评价影响，比如通过模型方法。

三、制定评估计划的重要性的关键因素

（一）制定评估计划的重要性

在项目初期与投资者一起讨论一个完整的评估计划是十分有益的。从评估计划的过程中可以获得以下收益：

- ① 评估计划将为项目管理者 and 投资者提供评价评估需求、资源、能力和领域优先权的机会。
- ② 制定评估计划将向投资人说明项目计划如何对收到的资源担负责任。
- ③ 制定评估计划的过程中，常常可以发现现有的数据资源和/或过去的评估活动。利用这些现有数据和/或过去的评估成果能够帮助发展一个更有效的、低投入的评估计划。
- ④ 发展一个长期的评估计划能够说明有关评估优先顺序的未来决策。
- ⑤ 发展适当、全面的评估计划还可以有益地影响投资者的决策。

（二）评估计划的关键因素

在发展评估计划过程中，策划小组应考虑以下几点关键因素：

1. 评估范围

指定项目的目的和目标，建立基本概念框架（或逻辑模型），并确定现实期望结果。框架应整合投入、活动、产出、结果、影响，而期望结果是评估将要提供或说明的内容。

2. 评估方法

发展评估计划包括确定结果指标与测量方法以及数据来源和数据分析计划。

3. 计划的执行

描述开展的活动、参与者的角色和责任，以及确定活动的时间表，例如，何时分析数据，结果何时获得等。

4. 评估结果发布

确定谁将把评估结果转化为对于项目设计者、项目管理者 and 决策者有用的信息，结果将如何发布（如文章、口头演讲、项目材料等）以及为将来评估活动设定优先顺序。

四、建立全面整合的评估计划

评估计划一般包括几个常见步骤。这个过程中，主要涉及项目计划者、评估者、项目人员以及投资人。目标社区涉及的人员也应告知评估策划的过程。项目相关人在早期的涉入保证了项目终期评估结果的有用性。

（一）确定项目目的和目标

评估计划包含确立项目的目的和目标，并建立项目模型，可通过详细的书面描述确立项目的目的和目标（有时还有分目标），并描述项目如何运行来达到这些目标。此步完成后，项目的逻辑模型可以很容易地通过图表建立，并用于建立评估计划。

项目的每一个环节都应设立目标和分目标，而且每一个目标和分目标都应指定收集测量成功的指标，以确定项目的进展。

单个的项目不能开展影响评估，因为它的影响是与其他具有相同目标的项目在多个层面协同产生的。影响评估最适合在大的地理区域内开展，比如省或国家范围，它可以检测是否多个项目达到了期望影响的效果。这些影响可通过血清学监测系统测量，此类系统监测 HIV 和 STI 的流行趋势，也可通过重复的危险行为调查测量。地方机构通过外展、性病治疗和关怀等直接接触目标人群，他们应致力于项目执行的评估，而非结果或影响的评估。这意味着集中于投入的质量，如社区信息的培训和预实验，同时包括干预的覆盖面。

不同层面的项目进展时间不一，从过程完成水平的几个月（如人员培训）到结果和影响水平的几年。如果完成步骤一所必须的信息或数据不易获得，那么评估计划应先调整为收集它们。这一步完成时要以文字说明，同时，要在评估计划中阐明评估问题的优先次序。设定评估问题的先后顺序有时十分困难，因此，需要在后面的评估计

划制定过程中不断修订。

（二）检测现有数据和既往评估调查

包括确定现有的数据资源和其他的既往开展的、正在进行的和/或由其他资助方支持的评估活动。此步骤中，评估计划小组应评价是否其他小组正在计划开展类似的评估工作，如果是，应邀请他们共同制定评估计划。

（三）确定内部和外部评估资源和能力

确定评估资源不仅指确定评估的资金来源，还包括确定评估的其他类型的资源，比如有评估经验的人员可以帮助计划和开展评估活动。此步骤还包括确定管理和连接多个数据库及计算机系统的项目开展能力。

（四）确定评估问题、可行性以及适宜的设计和指标

评估专家和项目管理者需阐明优先评估的问题、合适的评估设计、结果测量或指标、数据需求和收集以及分析数据的方法。对于获得数据和维护数据系统的实用途径和方法应该加以讨论，而且数据系统应该具有可持续性并容易获取数据。一个全面评估的可操作性计划也应在此步骤中建立。因此，此步骤应有一个书面计划，简要描述评估的问题和评估设计、数据收集方法和分析计划以及整个评估的时间进度表。

（五）发布和使用评估结果的计划

此步不总是开展，但应该开展，因为它对于确保评估结果能充分地应用于项目促进和决策十分有益。此步包含如何使用评估结果的计划，如何转化评估结果为项目策略方针，以及如何将其发布给项目投资人和决策者。这一步也应包括建立与下一个评估的计划人的反馈循环和反馈机制，这样在新的评估中可以有效地吸取过去的教训。

将评估结果转化为对项目设计者有用的信息有助于项目的促进，这一点十分重要。此步最好与熟悉干预措施的个人和小组协作开展。评估结果如同所有信息一样，必须要市场化并针对其目标听众进行“包装”。

这一步在评估的计划和执行期没有考虑到，常常只在评估结束出现问题时，有人提出“评估计划是如何执行的？结果怎样用于促进 HIV 预防项目和政策？”的问题，评估结果的发布才成为一个问题。如果没有发布和使用评估结果的计划，这个问题常常由于评估者忘掉了细节或项目已经继续进行而无法回答。缺少发布计划将影响评估结果的作用和未来的评估活动。信息发布不充分也常常会导致评估工作变得多余，因为其他人并不知道以往评估活动的发现。另外，应改进一成不变的评估发布方式，否则无法促进项目的开展。因此，一个项目的总的评估计划中，应包括发布评估结果的计划。

第四节 评估结果的发布和利用

如果评估的结果不能影响关键的项目相关人（包括决策者、项目管理者 and 项目实施的社区）采取减少 HIV 传播的行动，评估活动就变成了无意义的的数据收集活动。因此，采取能够促进对艾滋病有所反应的形式广泛地分发评估结果是任何数据收集系统的必要组成部分。一些人通过发表详细的完工技术报告发布评估结果，但如果发布结果是为了让项目相关者采取必要的行动，发布评估结果就变得更为复杂。实际上，有效地发布评估是一个持续的过程，它开始于评估项目的形成阶段，即项目设计和发展阶段，并在项目的整个实施过程中持续。

这个过程的主要事件包括：a. 在多个项目相关者中形成一致意见，包括关于开展评估活动的社区、需要收集和发布的信息以及发布信息的形式；b. 在项目设计阶段建立完善的信息发布策略，此策略应包括在数据收集完成后尽快发布一些重要发现，以维持大家对项目的关注并加速预防项目的执行；c. 在正式发布结果前准备目标听众的名单，这些人将理解评估结果的含义、限制和解释；d. 为不同的目标听众设计不同的分发材料，这些材料将通过通俗易懂的语言解释评估的结果；e. 主动地追踪目标听众，可通过回答问题、澄清结果的含义和解释、对适宜的策略、项目或行动提出建议。

一、结果的发布对象

结果的发布对象在不同的国家有所不同，取决于当地的情况、有影响的团体、社会和政治结构。在我国可能的对象可能包括：

- 决策者，包括国务院、卫生部或其他相关部委的政府官员
- 政府机构、非政府组织（NGOs）和社区组织（CBOs）的项目管理者
- 开展评估活动的社区的成员和领导
- 新闻媒体
- 项目投资方或投资机构
- 普通大众
- 大学和研究机构
- 私营老板
- 工会
- 宗教团体
- 军队
- 专业机构
- 省、市、县、乡、街道各级公务人员

通常建议发布的对象越广泛越好。上面列出的每个潜在的目标听众都拥有减少

HIV 传播的责任，而且能够通过了解评估结果采取相应的行动。但是，一般当地都会对公开讨论传播 HIV 的行为有些担心和敏感。这些敏感在一些重要的政府官员或高层的宗教领导中表现得更为明显，而他们的疏远往往严重地影响预防项目的效果。在设计发布评估结果的活动和材料时，对于这些担心，特别是对一些重要决策人的担心应多加关注。任何时候，信息的发布形式应考虑目标人群的担心，使他们参与到发布的过程中，而且避免使用他们反感的语言。一些间接的发布方法也可以尝试采用，包括向最接近决策者和宗教领导者的人解释重要的发现，他们可能把这些信息传递给他们的领导。由于评估社区的成员参与了数据收集的过程并能够根据结果改变一些行为，开展评估工作的实施者有责任确保社区的领导和成员了解整个评估的结果。他们在整个评估的设计、实施和结果分发过程中的主动参与将加强评估体系的建立，并有助于确保评估的可靠性和真实性。

二、可能的分发形式

开展评估的人员虽然完成了评估报告但不能就确定完成了评估工作，即使报告中包含了翔实的统计分析结果。虽然技术报告是必要的，但绝大多数群众不明白“P 值”和“可信区间”的含义，而这样的一些专业术语会使大众更加不明白所关心的问题。为了能够将适当的信息传递给可能的目标听众，通常可使用多种发布形式，包括：

1. 具有完成统计分析的翔实评估报告

这份报告作为其他发布材料的技术依据，而且适宜分发给在社区开展评估调查的项目主管和有关机构的技术人员。在报告中，一些重要的结果应该用清晰的、非技术性语言概括出来并突出显示。

2. 短讯或其他简短的大众媒体方式

这些材料应避免使用技术性语言。每个简报或短讯重点反映一到两个主要发现和含义即可，以免使读者糊涂。评估人员在媒体讨论评估发现时应尽可能提供文字的简评，这有助于减少错误的引用和错误的解释。评估者应该懂得新闻短讯是一种间接的向决策者和大众传递信息的方式，因此应计划撰写的内容，这一点十分重要。

3. 一至两页的政策摘要

由于决策者和开展评估社区的领导者处在一个对预防措施有重要影响的位置上，必须为他们专门制作一些材料。每个政策摘要应侧重一至两个主要的评估发现，讨论评估活动的意义，并针对他们很可能采取的影响预防活动的措施提供建议。向不同层面的决策者和社区领导人发放多种不同的政策摘要十分必要，这些材料将直接影响他们。

4. 小组发布会议或报告

大型会议或报告可以提供向社区成员、决策者、项目管理者或一般大众汇报评估结果的机会。发布的全过程包括公布数据之前、之中和之后的各种会议。公布数据之前的会议可以解释收集哪些数据，讨论数字解释的问题，以及筹划接受信息的人员名

单。数据公布时或之后的会议提供了开展以下活动的机会，如汇报评估发现、回答问题、澄清误解以及根据人们各负的责任和所能影响的活动对评估发现进行解释。

5. 单独会见

与有影响的决策者、评估社区的领导者或感兴趣的 NGO 和政府项目管理者进行个别的私人会议，为他们提供了获取并理解评估结果的机会。因为这样的会见需要充足的时间和充分的准备，评估人员需要注意选择那些能够对信息采取有效行动的人。

在为每个目标听众准备材料时，发布信息的根本目标是为听众提供信息，使他们能够鼓励别人或自己开展有意义的预防活动。

三、为特殊的听众特定结果的发布

当材料发布给大众传媒、决策者、大众，必须要采取合适的形式并被他们所理解。提高发布效果的一些建议如下：

① 避免使用学科专用的语言，比如一些统计术语（“卡方”或“统计学显著性”）或社会科学术语（“非固定性伴”或“肛交”）。这些词汇不容易被目标听众理解或容易被曲解。

② 为决策者和社区领导人制定的材料中，包括在他们的影响范围内可采取行动的建议。例如为教育部制定的材料要强调在初中阶段应加强校内性教育，因为相当数量的青年报告在较小的时候就发生了首次性行为。为最大程度地保证信息的可靠性，任何在政策简讯中的建议都应基于数据发现本身。

③ 在为其他听众制定的材料中，对评估结果含义的描述都应从听众的角度出发。例如有关工厂工人中的高危行为的报告应强调增加安全套的使用，或者有关男性高比例的婚前性行为的报告将会探讨婚姻可能会使年轻女性暴露于 HIV 的含义。

④ 将数据介绍给适宜年龄、性别和其他适宜特征的人群将帮助人们理解结果的含义。因为危险行为水平在不同人群之间、男性女性之间有差异，在同一个人的不同生命阶段也会有变化，因此报告中在解释评估结果的时候需要考虑这些因素。一些结果在报告时没有进行亚人群、年龄、性别或其他特征分类，这样很难让人们了解重要的事实，比如青少年的高危行为。

⑤ 突出那些可以让目标人群产生行动的一些数据。因为评估常常寻找一些变化的趋势，一些人错误地认为缺少明显的趋势就没有意义，例如与非固定性伴具有较低但稳定的安全套使用率在推动安全套使用方面是一个“催化剂”。

⑥ 报告边缘群体的数据，如性服务人员或男男性行为人群，要细心、谨慎地关注该群体关心的问题。报告前要精心计划，并采取不增加对这些人群产生歧视的报告方式。

⑦ 报告应向听众说明为什么他们应该关注艾滋病。为财政部作的报告应说明如果不控制艾滋病的流行将很难达到预期的经济发展目标，为卫生专家作的报告应说明艾滋病将影响达到儿童生存和期望寿命的目标，为工会作的报告应说明工会成员和他们

的家庭将如何受到艾滋病的影响，将艾滋病与听众最关心的问题联系起来是一个报告有效的重要因素。

⑧ 告诉听众做一些事情可以改变现状。另外，报告的最后一部分应该描述现在可以采取的一些有效行动。这些行动包括听众自身可采取的行动以及他们可支持其他人完成的行动。

⑨ 使用预测方法举例说明过去的趋势可以改变，艾滋病将如何影响我们实现将来的目标。这包括使用流行病学方法预测 HIV 感染者人数、艾滋病病人数、艾滋病死亡人数，以及预测艾滋病的社会经济影响，如医疗卫生花费、孤儿数、经济增长、妇女生存以及郊区发展等。这些预测可用于说明艾滋病对家庭、社区和国家所有范围的影响。

四、结果发布的注意事项

我们不能低估发布评估结果的重要性。通过有效地发布，评估结果可以用于倡导采取一些有效的行为，并激发社会对艾滋病预防的更广泛反应。

① 发布结果是一个过程，它起始于项目评估的设计阶段并贯穿评估的整个过程。其目标不仅仅是广泛发布评估的结果，更重要的是要被人们理解并使人们行动起来。

② 通常，结果应尽可能地广泛发布，但是应该本着尊重目标听众和项目相关人，关注他们关心和敏感的问题。可能的目标听众包括决策者、评估社区的领导和成员、政府或非政府组织或社区组织的领导、大众媒体以及一般大众。

③ 发布形式有多种，包括技术报告、政策短讯、大众传媒的简评以及小组或个人会见。这些典型形式应该传递评估结果给目标听众，使他们行动起来。

④ 材料应该使用容易被目标听众理解的简单语言编写。不同的材料应针对不同的听众讨论评估结果的含义，并为他们提供建议采取适宜的行动。

⑤ 评估人员应该仔细了解向受歧视的群体发布信息的途径，避免使他们进一步受到歧视。

⑥ 有效地向重要的项目相关者发布评估结果的策略是扩大艾滋病预防的社会反应的重要因素。

第五节 中国艾滋病防治监督与评估指标体系

2005年7月，联合国艾滋病规划署发布了为实施《艾滋病承诺宣言》而建立的《核心指标构建指南》；2005年10月，世界卫生组织与联合国艾滋病规划署共同撰写了旨在为国家级和省级决策者提供指导艾滋病防治重要参考信息的《艾滋病防治战略信息框架》(Strategic Information Framework, SIF)；2006~2007年，卫生部制定了《中国艾滋病防治监督与评估框架(试行)》(简称《框架》)和《中国艾滋病防治监督与评估指标体系》

(简称《体系》), 用于指导全国艾滋病防治监督和评估工作。

一、监督与评估内容

《中国艾滋病防治监督与评估框架(试行)》从4个层次对艾滋病防治工作进行监督与评估, 即保障措施的落实和防治工作的开展、知识与行为的改变以及防治效果。

1. 保障措施的落实方面

经费投入、政府和部门防治规划和计划制定、多部门协作与配合、机构和人员能力建设、社会团体和民间组织参与等情况。

2. 防治工作的开展方面

大众宣传教育、咨询检测服务、高危人群干预(安全套使用、药物维持治疗等)、预防艾滋病母婴传播、血液安全、性病防治、艾滋病治疗(抗病毒治疗或中医治疗、机会性感染治疗)、为HIV感染者和艾滋病病人及其家庭提供关怀和救助等防治活动情况。

3. 知识与行为的改变方面

城市及农村居民、农民工、校内外青少年等大众人群艾滋病基本知识知晓率、暗娼和男男性行为者安全套使用率、注射吸毒者共针率等知识和行为的改变情况。

4. 防治效果方面

特定人群HIV和梅毒感染率、抗病毒治疗或中医治疗后病人存活率等, 结合知识与行为指标评价防治效果。

二、监督与评估指标

《中国艾滋病防治监督与评估指标体系》共有17个监督评估指标, 其中核心指标9个, 附加指标8个。核心指标为在监督评估过程中各地必须收集和填报的指标。附加指标为各地根据实际情况可以选择收集和填报的指标。各地附加指标的选择由国家级监督评估专家组确定。

(一) 核心指标

中国艾滋病防治监督与评估的核心指标见表27-1。每个指标的评估频次以及信息来源都给了明确的说明。

表 27-1 中国艾滋病防治监督与评估核心指标

指标名称	频度	信息来源
保障措施		
1. 省、地和县级政府本年度艾滋病防治专项经费数	每年一次	省、地和县级财政部门、防艾办

续表

指标名称	频度	信息来源
防治工作		
2. 在过去 12 个月中接受过艾滋病检测并知晓检测结果的高危人群比例	每年一次	综合监测或专题调查、VCT 门诊
3. 艾滋病预防措施覆盖的高危人群比例 (目标: 2007 年达到 70%, 2010 年达到 90%)	每年一次	综合监测或专题调查
4. 符合治疗标准的 HIV 感染者和艾滋病病人中接受抗病毒治疗或中医治疗的比例 (目标: 2007 年达到 50%, 2010 年达到 80%)	每年一次	卫生机构记录、病例报告统计
5. HIV 感染者、艾滋病病人及其家庭获得关怀、支持和服务的比例	每年一次	民政、卫生等部门的报表、专题调查
知识与行为		
6. 大众艾滋病基本知识知晓率 (目标: 城市 2007 年达到 75% 以上, 2010 年达到 85% 以上; 农村 2007 年达到 65% 以上, 2010 年达到 75% 以上; 校内青少年 2007 年达到 85% 以上, 2010 年达到 95% 以上; 校外青少年 2007 年达到 65% 以上, 2010 年达到 75% 以上)	两年一次	专题调查
7. 暗娼最近一次发生商业性行为时使用安全套的比例 (目标: 2007 年达到 70%, 2010 年达到 90%)	每年一次	综合监测或专题调查
8. 注射吸毒者在最近一个月未共用注射针具并使用了安全套的比例 (目标: 2007 年达到 70%, 2010 年达到 80%)	每年一次	综合监测或专题调查
效果		
9. 特定人群 HIV 感染率	每年一次	哨点监测、综合监测或专题调查

(二) 附加指标

中国艾滋病防治监督与评估附加指标见表 27-2。每个指标的评估频次以及信息来源都给了明确的说明。

表 27-2 中国艾滋病防治监督与评估附加指标

指标名称	频度	信息来源
保障措施		
1. 省级部门有艾滋病防治工作计划和相应经费支持, 并按计划开展工作	每年一次	省(自治区、直辖市)防艾办、省级部门工作计划、文件和总结
2. 社会团体和民间组织开展艾滋病防治活动的数量	每年一次	防艾办的文件及其他资料
防治工作		
3. 临床用血来自无偿献血的比例 (目标: 2007 年达到 90%, 2010 年达到 100%)	每年一次	卫生行政部门、血液中心(血站)的报告、专题调查
4. 阿片类毒品成瘾者接受社区药物维持治疗的比例 (目标: 2007 年达到 40%, 2010 年达到 70%)	每年一次	服务统计、药物治疗机构记录
5. HIV 感染孕产妇采取预防艾滋病母婴传播干预措施的比例 (目标: 2007 年达到 85%, 2010 年达到 90%)	每年一次	个案登记表、病历记录、随访记录、相关统计表
知识与行为		
6. 男性在最近一次与男性性伴发生性行为时使用了安全套的比例 (目标: 2007 年达到 70%, 2010 年达到 90%)	每年一次	综合监测或专题调查
效果		
7. 接受抗病毒治疗或中医治疗 12 个月后存活 的病人比例	每年一次	病人记录、报告
8. 高危人群梅毒抗体阳性检出率	每年一次	哨点监测

三、中国艾滋病防治监督与评估工作的组织与管理

(一) 国家级监督与评估

国务院防治艾滋病工作委员会办公室(以下简称国艾办)负责组织协调全国艾滋病防治监督评估工作的开展, 包括组织协调制定国家监督与评估框架和方案, 组织开展《规划》和《行动计划》的监督评估, 以及多部门联合监督。国务院防治艾滋病工作委员会各相关成员单位负责协调和配合涉及本系统的监督评估工作。中国疾病预防控制中心性病艾滋病预防控制中心负责监督评估的技术支持, 负责监督评估数据信息的收集、分析与报告, 并开展有关研究和培训等。

国艾办组建由艾滋病综合防治和评估领域专业技术人员组成的国家级监督评估专家组，为监督评估工作提供咨询，参加国家艾滋病防治监督评估方案的制订和修订，参与国家级艾滋病防治的监督评估工作，参与审阅年度监督评估报告，承担国家监督评估培训任务等。

（二）省级监督与评估

各省指定相关部门负责组织协调本省（自治区、直辖市）艾滋病防治监督评估工作，指定授权技术支持机构开展监督评估信息资料的收集、分析和报告，成立专家组协助开展省级监督评估工作。

四、中国艾滋病防治工作监督评估的方法

由国艾办根据防治工作需要，组织制定有关监督方案和监督评估指标，采取日常监督与评估和专项监督与评估等方式，每年组织不同形式的监督与评估活动。各地、各部门、各项目可根据需要，参照本框架，制定具体监督与评估方案，组织开展监督与评估活动。

为做好监督与评估工作，实现监督与评估的目的，国艾办组织建立监督评估信息系统，并根据需要组织开展专题调查。

1. 建立监督评估信息系统

根据监督与评估框架，建立国家艾滋病防治监督与评估信息系统，收集和整合各有关方面的信息（如艾滋病哨点监测、综合监测、传染病疫情报告、健康教育信息、项目信息等），将相关信息进行整理分析。

2. 开展专题调查

根据监督与评估框架，对现有信息收集系统收集不到的信息，开展专题调查，采用定性、定量的方法收集相关信息。

五、监督与评估的经费保障

国家级监督与评估工作经费由中央财政负担。全国各省（自治区、直辖市）需明确本省监督与评估的经费安排，保证本省监督与评估工作的顺利开展。

（邱泽奇 毛宇嵘）

参考文献

- [1] 艾尔·巴比. 社会研究方法基础 [M]. 北京: 华夏出版社, 2002.
- [2] 彼得·罗西. 项目评估: 方法与技术 [M]. 北京: 华夏出版社, 2002.
- [3] 国务院防治艾滋病工作委员会办公室. 中国艾滋病防治督导与评估框架 (试行) [M]. 北京: 人民卫生出版社, 2009.
- [4] 李强. 应用社会学 [M]. 北京: 中国人民大学出版社, 2004.
- [5] Campell D T. Reform as Experiments [J]. *American Psychologist*, 1969, 24 (4): 409 - 429.
- [6] Campell D T, Julian C. Stanley. *Experimental and Quasi - Experimental Designs for Research* [M]. Chicago: Rand McNally & Co, 1966.
- [7] Thomas D C, Campell D T. *Quasi - Experimentation: Design and Analysis Issues for Field Settings* [M]. Chicago: Rand McNally, 1979.
- [8] Cronbach L J, Ambron S R, Dornbusch S M, et al. *Toward Reform of Program Evaluation* [M]. San Francisco: Jossey - Bass, 1980.
- [9] Cullen M J. *The Statistical Movement in Early Victoria Britain: The Foundations of Empirical Social Research* [M]. New York: Harper and Row, 1975.
- [10] Emma E. Flook, Paul J. Sanazaro. *Health Services Research and R & D in Perspective* [M]. Ann Arbor: MI Health Administration Press, 1973.
- [11] McCord W, Joan McCord. *Origin of Crime: A New Evaluation of the Cambridge - Somerville Youth Study* [M]. New York: Columbia University Press, 1959.
- [12] Patton M Q. *Qualitative Evaluation and Research Methods* [M]. Newbury Park (CA): Sage Publications, 1990.
- [13] Power W, Helen W. *An Experiment in the Prevention of Delinquency: The Cambridge - Somerville Youth Study* [M]. New York: Columbia University Press, 1951.
- [14] Rehle T, Saidel T. *Evaluating Programs for HIV/AIDS Prevention and Care in Developing Countries* [EB/OL]. FHI, 2006. <http://www.fhi.orglen/hiv aids/pub/ardive/eva/chap/index.htm>.
- [15] Solomon R L. An extension of control group design [J]. *Psychological Bulletin*, 1949, 46: 137 - 150.
- [16] Samuel A, Stouffer. *The American Soldier: Combat and Its Aftermath (Plastic Comb)* [M]. Princeton: Princeton University Press, 1949.

- [17] Toulemonde, Jacques, Lise Rochaix. Rational Decision – Making through Project Appraisal; A Presentation of French Attempts [J]. International Review of Administrative Sciences, 1994, 60: 37 – 53.
- [18] Spencer Vibbert. What Works; How Outcomes Research Will Change Medical Practice [M]. New York: Ground Round Press, 1993.
- [19] Carol H, Weiss. Evaluation Research; Methods of Assessing Program Effectiveness [M]. Englewood Cliffs: Prentice – Hall, 1972.
- [20] Carol H, Weiss. Evaluation; Methods for Studying Programs and Policies [M]. Upper Saddle River: Prentice Hall, 1998.
- [21] WHO, The World Bank, UNICEF, et al. Monitoring and Evaluation Toolkit: HIV/AIDS, Tuberculosis and Malaria [M/OL]. The Global Fund, 2006. http://www.the-globalfund.org/documents/me/M_E_Toolkit_lores_en.pdf.
- [22] Joseph S, Wholey. Assessing the Feasibility and Likely Usefulness of Evaluation. in J. S. Wholey, H. P. Hatry, and K. E. Newcomer (Eds.), Handbook of Practical Program Evaluation [M]. San Francisco: Jossey – Bass, 1994.

