

发展中国家的
艾滋病预防和关怀
项目评估
项目管理人员和决策者手册



新星出版社

发展中国家艾滋病预防和关怀项目评估

项目管理人员和决策者手册

家庭健康国际 (Family Health International)

美国国际发展署资助

主 审：	郝 阳	孙新华
编译人员：	王维真	文 毅
	孙江平	韩孟杰
	程 峰	朱成斌
	胡翼飞	袁山英
	马英鹏	周 郁
	吴 迪	

新星出版社

二〇〇六年十二月

图书在版编目 (CIP) 数据

发展中国家艾滋病预防和关怀项目评估项目管理人员和决策者手册 /

王维真等编译. —北京: 新星出版社, 2006.12

ISBN 7-80148-778-8

I.发... II.中... III.发展中国家-艾滋病-预防(卫生)-手册 IV.R512.910.1-62

中国版本图书馆 CIP 数据核字 (2006) 第 140469 号

发展中国家艾滋病预防和关怀项目评估项目管理人员和决策者手册

郝 阳 孙新华 / 编(著)

责任编辑: 温志宏

装帧设计: 王 燕

出版发行: 新星出版社

出 版 人: 谢 刚

社 址: 北京市东城区金宝街 67 号隆基大厦 100005

网 址: www.newstarpress.com

电 话: 010-65270477

传 真: 010-65270449

法律顾问: 北京建元律师事务所

经销电话: 010-65512133

邮购电话: 010-65276452

邮购地址: 北京市东四邮局 7 号信箱 100010

印 刷: 北京高迪印刷有限公司

开 本: 889 × 1194 1/16

印 张: 9.75

版 次: 2006 年 12 月第一版 2006 年 12 月第一次印刷

书 号: ISBN 7-80148-778-8

定 价: 28.00 元

序 言

人类为了遏制艾滋病的流行已付出了20年的努力。世界各国都在尝试、推广着自己和他人的努力和经验。近20多年来,国际社会对艾滋病的预防和控制投入了大量的资源,制定了许多政策和法规,采取了许多有效的干预措施,开展了众多的项目和活动。如何评估这些付出的努力和取得的效果,越来越受到政府和专家们的重视。

虽然我国艾滋病流行的态势越来越严峻,但随着我国控制艾滋病疫情的政策环境不断完善,和投入力度的不断加大,加之越来越多的多边、双边国际合作项目在我国各地开展起来,对这些国内外项目的监督评估要求也纳入了我国政府的工作日程。不同的项目有着不同的目标和要求,如何确保这些目标 and 要求的实现和符合评价他们的防治效果是摆在我们项目管理者面前一个很现实的问题。了解和掌握监督评估的技术要求,不仅有利于项目管理者有效的管理项目,同时对于国家项目和防治任务的高质量实施也有着极大的推动作用。

监督与评估是一套科学的管理体系,贯穿于项目的整个过程,甚至在项目设计之初就要有针对性的设计相应的监督与评估计划。好的监督评估计划可以有效的确保项目实施进度,同时还可以及时发现项目执行中的问题,进而进行调整和改进,提高项目质量,确保项目目标的完成。高质量的项目产出,不仅可以扩大项目影响,而且还可以吸引更多的资源投入,在更广泛的领域进行应用和推广。因此,监督评估工作质量的好坏对于项目成败的意义不言而喻。

本书为美国国际开发署(USAID)和家庭健康国际(FHI)共同实施的艾滋病预防与关怀实施项目(IMPACT)开发的《在发展中国家评估 HIV/AIDS 预防与关怀项目—决策者与项目管理者手册》(Evaluating Programs for HIV/AIDS Prevention and Care in Developing Countries 孝 A Handbook for Program Managers and Decision Makers)中文版。该书共分三个部分十七章,较全面的对监督与评估的概念、作用、评估策略、评估方法等方面进行了介绍。书中清晰的概念不仅提供了坚实的理论基础,丰富的实例也会帮助读者更加深刻地理解其内容。正像该书的书名一样,该书是项目管理者 and 决策者在进行项目管理和决策制定时非常有用的工具。

由于艾滋病防治项目的监督与评估还是一个比较新的领域,在翻译中难免有不足和需要探讨的地方,希望读者见谅并及时指出。

感谢家庭健康国际(FHI)中国办事处对此书的编译和出版给予的支持。

原 序

评估是项目实施过程中经常做的事情。本手册倡导将评估作为一个有成效的艾滋病预防和关怀项目最初计划时的关键部分。

读者将会发现作者已将项目分成不同阶段，并提供了综合和策略性很强的评估工具。手册所推荐的方法在预防、关怀和支持等方面的项目中，对实施者提供了有用而重要的信息。从人们理论上知道实施制定有效干预并不表明实际有效。但我们具有增加和保持全球政治承诺、将资源投入到艾滋病项目的能力，也表明我们的干预能明显地降低新发感染和改善艾滋病感染者/病人的生活质量。

监督与评估项目结果对公平分配资源、使项目最终产生可持续的影响非常关键。为达到此目的，本手册为国家项目管理者提供了综合框架以帮助评价项目的效果。本手册描述了必须回答的三个简单而又关键问题的方法，即：我们是否正在做正确的事情？我们是否正在正确地去执行？我们是否在足够大的范围去做并使之效果明显。回答这些问题后，项目管理者需要决定是否修改现有的项目和制定新的项目。

当全球进入艾滋病流行的第三个十年，我们必须展示我们所付出的努力和取得的效果。艾滋病流行的三十年，也是机遇和挑战共存三十年，也是表明我们可以并且已经扼制艾滋病流行的三十年。

Helene Gayle

美国疾病预防控制中心

艾滋病、性病、结核预防中心主任

致 谢

本手册所列的作者仅是部分直接对艾滋病预防和关怀项目监督与评估概念和方法的形成作出贡献的专家。

自从 30 年前艾滋病被发现，人们一直在通过干预活动以阻止其流行。项目管理人员、研究人员非常关注如何采取最佳的干预评估、监测调查来阻止艾滋病流行。

我们与众多国家进行了合作，包括美国国际发展署的 IMPACT 项目、家庭健康国际的全球预防和关怀项目、英国国际发展部、联合国艾滋病规划署、世界卫生组织。

另外，我们感谢美国国际发展署在其他国家资助项目的很多同事们，如 MEASURE、HORIZONS、AIDSMARK 和 SYNERGY 等，他们从不同角度探讨了评估和监测的问题，许多观点对本书做出贡献。

最后，所有艾滋病项目监督与评估的努力有赖于艾滋病感染者/病人或处于艾滋病危险人们的参与。无论是定量调查、深入访谈，还是小组访谈，我们都请求参与者交流和表达他们对性、注射吸毒、其他易感性等问题的看法，如果不是参与者在研究中的合作和参与，我们的项目不会有进展。我们非常感谢那些代表了处于艾滋病感染危险的人们，希望本书提供的可操作性的数据收集方法将为艾滋病预防和关怀项目提供更多的帮助。

简介

当艾滋病在全世界发展中国家流行和不断蔓延的时候,对减少传播的有效和适宜的干预设计比以往更重要,在与艾滋病进行斗争 15 年后,世界范围的防治信息不断汇集:一些国家在遏制新感染方面取得了成功,而一些国家感染率仍持续增高。

本手册试图在艾滋病防治领域内研究、探索和评估艾滋病预防及关怀项目的效果,目的是提高项目质量,即使在资源不足的机构本手册的方法也适用。近年来在开展有效干预方面取得了实质性进展,为评估项目提供了适宜的方法、策略和越来越具体而灵敏的评估指标。本手册将科学家、项目设计者、实施者、大学、当地社区组织、政府、私人机构等相互合作的经验结合起来。

然而,与其他公共卫生机构相比,艾滋病预防和关怀项目的评估开展相对较晚,关键的问题依然是如何最好地测量效果。关于评估和测量话题的争论将会继续下去。在描述关于评估话题的不同侧面时,以下的章节将面临这些挑战。

本手册的主要内容是与艾滋病的经性途径传播有关的项目评估,它是为特定的项目经理、项目决策者提供服务,通过实验和半实验方法评估干预的有效性。

第一部分描述概念、方法和框架的一致性(第一节),简述项目的综合性评估和监测计划的制定过程(第二节),为艾滋病项目评估打下基础。

第二部分描述了评估有效减少经性途径传播艾滋病所必需的核心项目措施(第三、四、五、六节),介绍了评估不同项目内容所用的不同方法。这一部分对如何评估艾滋病预防和关怀项目提供了长期的指导原则(第七节)。通过向资助者、决策者提供必要的反馈信息来决定所捐助的资源是否产生了预期的效果,这将有助于使他们增加承诺来资助关怀项目活动。

第三部分关注行为改变测量作为标准化预防实施的关键产出,包括了收集行为数据的最新方法学和调查工具(第八、九、十节)及评估它们的合理性和可信性(第十一节),运用定性信息的收集工具(第十二节),以及如何向关键人员宣传调查数据的有效措施(第十三节)。

最后,第四部分涉及了与项目影响有关的评估内容,强调了行为分析和血清流行病学数据的必要(第十四节),还描述了一个预测不同预防 HIV 传播策略影响的新工具(第十五节),为成本和成本-效益分析提供了可操作性指南(第十六、十七节)。

本手册的每一节试图按照项目评估周期的逻辑关系叙述,并作为独立的章节书写,而且各章节互相之间是有联系的。书中每一节与其他节互相之间的索引能帮助读者理解每个评估部分是如何与整体内容相联系的。

本手册的作者和编辑希望能给项目管理者、决策者以清晰的、可操作性的指导。

目 录

第一章 艾滋病项目评估的作用

- 第一节 监督与评估的基本方法和框架 1
- 第二节 综合性监督评估计划的制定 11

第二章 评估措施和实施方式

- 第三节 评估行为改变交流的干预措施 18
- 第四节 评估性传播感染的控制项目 26
- 第五节 安全套项目的评估 37
- 第六节 艾滋病自愿咨询检测项目的评估 46
- 第七节 对艾滋病感染者 / 病人关怀项目的评估 53

第三章 行为趋势测量的方法

- 第八节 在项目评估中行为数据的应用 61
- 第九节 监测艾滋病危险行为的抽样策略 73
- 第十节 行为调查的指标和问卷 88
- 第十一节 自报行为资料的有效性和可靠性评估 97
- 第十二节 定性资料在评估艾滋病项目中的作用 101
- 第十三节 资料收集结果的有效公布 115

第四章 项目效果评估

- 第十四节 对艾滋病疫情和相应控制措施的理解：将行为改变方面的数据与
艾滋病疫情的下降联系起来 119
- 第十五节 将调查数据转化为项目影响：AVERT 模型 127
- 第十六节 艾滋病预防和关怀项目的经济与财务成本评估指南 136
- 第十七节 艾滋病预防和关怀项目成本—效果分析指南 147

列表目录

1-1 综合评估的框架	2
1-2 潜在项目产出 / 影响测量	4
1-3 研究设计: 证据的等级层次 (按证据说服率递减)	5
1-4 项目评估水平的指标举例	7
2-1 某国家项目 A 的逻辑模型	15
4-1 STI 控制项目的监督评估指标和数据来源	28
4-2 评估 STI 病历管理质量的常用方法	30
4-3 STI 调查的不同类型	32
4-4 用做被动调查指标的 STI 症状	33
5-1 UNAIDS 安全套项目指标	39
5-2 典型东非城市干预策略的相对成本效益	44
6-1 自愿咨询和检测项目的说明性指标	47
6-2 VCT 服务各方面的可能数据来源	50
7-1 监督持续性关怀的标准	57
8-1 建议根据流行阶段选择目标人群来重复进行横断面行为调查	66
8-2 不同国家选择做行为监测系统反映小组的范例	67
10-1 HIV/AIDS 项目中的关键行为指标	89
15-1 用于 AVERT 模型的每次性暴露造成的 HIV-1 传播概率	133
15-2 喀麦隆女性性工作者队列估计和观察到的 HIV 血清阳转数	134
15-3 模拟南非一次干预的影响	134
15-4 估计的不同于干预部分的影响	135
16-1 资产预期使用寿命估计指南	142
17-1 效果的分析和评估	153

列图目录

1-1 监督与评估	9
4-1 卫生系统模型发现性病患者中有效治疗的障碍	26
4-2 STI 动态传播模型	27
4-3 STI 和 HIV 之间的相互作用	35
5-1 墨西哥的人们对安全套作为艾滋病预防方法的知晓情况	41
6-1 进行艾滋病自愿咨询和检测的人们报告安全套使用和禁欲情况	52

7-1HIV/AIDS 持续性关怀	55
8-1 过去一年中报告的与非固定性伴性交的男性百分比	69
8-2 坚持使用安全套的女性性工作者的百分比：柬埔寨 1996—1999	69
14-1 发病率、感染率和死亡率的关系	120
14-2 Nsambya 和乌干达 HIV 的流行率	122
14-3 人口中 15-24 岁人群有性行为的百分比（乌干达 1989 年和 1995 年	123
14-4 泰国青年 21 岁入伍时艾滋病感染率的趋势	124
14-5 泰国政府诊所报告性传播感染的数量	124
14-6 泰国北部 21 岁男性性行为改变和艾滋病 / 性病下降	125
17-1 两种干预项目的一个简单决策树	154

方框图表目录

2-1 评估组织内部和外部观点的价值	14
2-2 在每个项目中进行 KABP 调查：有用还是无用	15
3-1 问题举例	22
6-1 HIV 自愿咨询的组成与特征	48
6-2 实例：AIDSCAP/UNAEDS/WHO 多中心随机试验	51
7-1 乌干达基本 HIV 关怀的推荐标准	57
7-2 综合关怀指标举例	58
8-1 识别和描述亚人群的分布	64
9-1 有关选择亚人群的原始收集单位（PSUs）或“群”的定义的解释	74
9-2 举例说明抽样框架的建立过程	77
9-3 两阶段抽样方法在基于固定场所的性工作者中运用的解释	78
9-4 在 MSM 中进行行为调查整群抽样方法应用的解释	79
9-5 在注射吸毒者中运用目标滚雪球抽样方法的解释	80
9-6 举例说明在学校班级调查年轻人	81
9-7 举例说明在教室外年轻人的学校调查	81
9-8 流动人口目标群体亚人群分群定义的说明	83
9-9 样本量计算的解释	83
9-10 对于定位一个已确定具有特定特征的目标人群的数量所需样本量的解释	84
12-1 软件包的选择	112
12-2 书写定性研究结果的小技巧	112
12-3 定性研究报告建议大纲（专题小组讨论、重要信息员访谈、快速人种学研究）	113

第一章 艾滋病项目评估的作用

第一节 监督与评估的基本方法和框架

评估艾滋病预防和关怀项目是一个长期的挑战，认识它的重要意义可以改进目前的工作，并有助于将来工作的顺利开展。

项目评估的定义和评估指标有许多种，如评估性研究、项目评估等，Michael Quinn Patton 对评估艾滋病预防和关怀项目的方法进行了准确的描述——我们非常广泛地使用评估这个词来囊括任何通过系统地收集数据来试图增加人类活动有效性的做法。当人们检查和判断所完成的工作和效果时，这就是进行评估。当这种检查是系统、完整、认真地收集和分析数据时，人们便是在进行评估性研究。评估是应用性研究，或实践性科学。这使评估性研究不同于基础科学研究。应用性评估研究的目的是指导行动、谨慎地决策和运用知识来解决人类和社会的问题。应用性评估研究的效果由是否能指导人们行动、是否能使干预更加有效，以及决策者和政策制定者是否能实际运用等来判断。

本评估方法以运用为核心，强调了不同水平信息的使用者和关键出资者的利益，包括出资方、实施国家、项目实施机构等。本评估方法通过社会学和流行病学研究来寻找改进艾滋病预防和关怀项目设计和实施的方法。

本小节首先描述了几个对于设计项目评估起关键作用的考虑因素，然后通过讲解评估的主要类型、讨论几个与评估设计和改进项目效果测量能力有关的主题，为读者提供了一个综合的评估框架。

1. 基本考虑因素

在艾滋病项目的评估决策过程中，选择一个合适的角度进行项目的评估是非常关键的，因为这决定了实际评估过程背后的指导思想。许多评估理论家和实践者提出了各种概念性的方法，如什么是评估、基层工作人员和出资者的关系是什么、谁对项目作出价值判断以及判断评估过程的标准是什么。这些方法在概念上是不同的。

对原则性方法的争议主要在于如何使用最好的方法来测量和解释变化。它强调了一系列多层面的方法学，层面的不同决定了选择方法的不同，例如：目前各界一致同意定性和定量的数据是有价值的、用于项目评估是合理的，实际上，这些方法不能单独使用，而应该结合起来使用。

收集什么数据和多少数据，取决于评估中如何在质量和信息的利用间作出权衡和选择，使用多种数据收集方法（包括定性、定量）的评估过程比单一的方法更能满足不同的评估需求，同时，当艾滋病流行快速增长时，评估的优先领域必须关注资源的有限性。评估者的主要任务应使研究方法与具体评估内容和资源的可获得性相一致。

在作评估项目设计时，还需要评估研究者发挥主观积极性，而不仅仅是一名咨询者。项目评估者应该是

一名中立的、客观的科学观察者。他们可以协调不同的出资方，能通过参与式的评估方法提高他人的能力，能在众多不同的决策者中积极广泛地宣传评估结果。为了实现实际需求和最佳方法间的平衡，参与式的评估设计和数据收集是非常关键的，使关键出资方参与设计过程，并建议他们尽可能有效地使用有限的资源。

确保关键出资人的积极支持和参与对于外部资助的项目非常重要，他们对不同数据收集系统得到的结果都很有兴趣。通常这些项目使用本国的机构收集数据，收集到的数据有时效性和较好的质量。无论如何，参与者们（包括实施机构、东道主国家、当地出资机构的代表）应努力对各种数据达成一致的需要意见。

2. 国家项目的综合评估框架

艾滋病预防和关怀项目需要在实施的不同的阶段进行评估。表1-1介绍了综合项目评估的框架，所有的项目评估阶段都要统筹考虑，以便为项目提供一个完整的评估意见。因为没有一种单独的数据收集方法能为改进项目的绩效或影响政策改变提供所有的信息，为满足各种评估需要，必须寻求多种评估途径和使用多种方法（定性的和定量的）。

表 1-1 综合评估的框架

评估的类型	不同评估能回答的问题
形成评估 (确定概念和设计)	干预是否有必要 谁需要干预 干预如何开展
过程评估 (监测投入和产出，评估服务质量)	所计划的活动实际实现了何种程度 服务提供是否良好
效果评估 (评估产出和影响)	观察到的产出是什么 产出说明什么 项目的不同之处是什么
成本－效益分析 (可持续性问題)	项目的优先领域是否被改变或扩大 资源应如何重新分配

3. 形成评估

形成评估应在项目尚未大规模实施时的计划和再计划阶段进行。此时，项目出资者对如何实施还没有一个确定的想法，项目的可塑性最大。

形成评估探讨了干预项目的需求，为明确干预的基本目的和目标提供了必要的信息，帮助项目设计者采取有效、可行的干预策略并对如何实施作出谨慎的决定。形成评估还可用于实际正在开展的项目，为管理者提供反馈信息，根据环境变化调整项目的目标。形成评估可以发现不能接受的或无效的干预措施、设计和概念。

因为艾滋病问题的紧迫性，许多干预项目在没有充分准备的情况下未做形成评估便匆忙实施，尤其在以社区为基础、以减少艾滋病的性途径传播为目的的干预设计中。在没有形成评估的情况下，容易导致整个项目的失败。在许多情况下，干预取决于干预实施对象以外的人们的想法，有关行为改变的文献充满了理想的设计，但

在实际运用中都失败了。这主要是因为这些干预设计对于目标人群来说是不可接受的，或表达方式不被目标人群理解和认可。可喜的是，这种情况正在改变，因为形成评估在预防不切实际的项目设计中越来越多地被运用。

形成评估运用综合研究的方法，为项目设计者快速提供了相关信息，这些方法包括：

- 现有信息的综述
- 专题小组讨论
- 个人深入调查
- 参与性观察
- 使用结构性问卷的短期定量调查

对于形成评估最多的方法学批评在于它缺乏外部可信度和普遍性，因为形成评估的结果来自小范围快速评估过程、试点研究，人们不能推广应用。尽管有这些局限性，但是，形成评估研究却能经常识别不能接受或无效的干预措施、设计和概念。即使在项目设计阶段做了充分的形成评估，也不能保证干预项目最终有效，可能会因为项目仍然没有充分实施而无效。

4. 过程评估

一旦项目活动拉开序幕，就需要检查它是否正确、按时、在预算内实施。过程评估解决的基本问题是：计划的实施到了何种程度、提供了什么服务、对谁、何时、何种频度、多长时间、什么内容等等。投入（所需的人、财、物、时间等基本资源）和产出（表现为已宣传社区、已培训人员、实施服务单位等方面的较快服务改进）都是过程评估的关键因素。这些关键因素通常用定量的数据回答。涉及项目如何和为什么起作用或失败的定性证据也是过程评估的重要内容。过程评估要求紧扣数据、获得项目的细节，不但是预期的结果，还有未预期到的结果，通过对预期活动进展过程的了解有助于解释干预的产出。但是，过程评估并不能反映干预是否有效。

在疾病流行动态的变化中，过程评估通过提供调整实施措施和项目目标必需的信息，在改进和调整干预项目方面发挥了重要的作用。以过程为核心的评估贯穿于项目实施始终，应使用不同的方法来评估所提供的服务，并进行排序，从对服务记录和常规报告系统的回顾、关键信息访谈、服务接受者访谈、便衣雇员的直接观察，到关于服务覆盖率、服务接受障碍的定量人口调查等。定性和定量调查结果之间的相互补充能提供综合的信息（详细的过程参见第二部分第三节到第七节）。

5. 效果评估

评估艾滋病预防项目的有效性需要定量的测量。定量测量能评估项目目标已达到的程度，效果评估通常回答的问题是：观察到的产出是什么？产出说明什么？项目有什么不同之处？

考虑到艾滋病预防项目不同实施阶段的情况，随着时间的推移，小年龄组队列性活动开始活跃，因此，建议对短期和中期项目效果（产出）和长期项目效果（影响）进行分层。表1-2提供了项目不同阶段产出和影响的测量举例：艾滋病相关态度的转变、危险行为的减少、保护性行为的采取、性传播感染的减低等，都被认为是对减少艾滋病的性途径传播项目设计而言的最适宜的短期或中期产出的测量指标。长期效果的测量内容包括艾滋病流行趋势、可持续性问题、社会应对的改进等。

效果评估与过程评估相结合，过程评估有助于评估者理解干预如何和为什么取得效果，换句话说是什么使之发生了改变。如果没有项目实施过程的评估而仅做产出或影响指标的分析，将会导致干预效果的错误结论。

项目的目的和目标应仔细定义，以便选择适宜的产出或影响测量方法来评估项目效果。效果评估通常以指标为基础，用指标对干预的产出和影响进行定量测量。由于多种干预措施联合运用能有效促使危险行为改

变，所以调查不应只针对一种干预效果来设计（参见图 1-1 和案例 2-2），而应针对接受综合干预人群组的行为趋势进行综合测量设计，单种干预的测量通常只在严格和花费大的对照实验中使用。

6. 成本－效益分析

成本－效益分析也是对项目的效果进行分析的方法。通过加入每个效果单位（如每个地区艾滋病改变的数量）成本的测量来扩展分析，通过比较不同干预的成本和效果能帮助找到优先领域，作出资源分配决策和进行项目设计。（第十七节，艾滋病预防和关怀项目实施成本－效益分析将对这种评估提供更多细节）

表 1-2 潜在项目产出 / 影响测量

项目产出 (短期和中期效果)	项目影响 (长期效果)
■ 艾滋病相关态度的变化 ■ 性传播感染率趋势 ■ 艾滋病 / 性传播感染有关危险行为持续性改变 ■ 个体和社会艾滋病易感性的减少	■ 艾滋病 / 性传播感染有关危险行为改变 ■ 社会支持和社区应对的增加 ■ 艾滋病感染 / 发病率的趋势 ■ 社会准则的持续改变

7. 结果归因分析，所观察的变化是否属于预防干预的结果

任何一项艾滋病预防项目的最终目标都是减少新发感染数。然而，由于流行随时变化，以及当流行变化为当地特定的状态时，我们对行为和流行因素如何影响流行模式缺乏认识，项目评估的内在因素确实复杂，许多与干预结果无关的因素可以导致艾滋病感染率或发病率在观察过程发生变化，如：

- 死亡率，尤其在成年人中流行
- 在高危人群中的饱和效应
- 通过亲朋经历的艾滋病影响而导致行为改变
- 与流行有关的人口流动模式
- 数据收集和解释的样本偏倚或错误

确定所观察到的艾滋病感染、发病的变化是否属流行的自然式结果或归因于干预结果，这是非常关键的评估问题。当面对艾滋病有关疾病人群增长时，做行为改变评估后如果发现导致危险减少的证据最能反映真实情况。例如，如果有一名亲戚或朋友感染了艾滋病，会使青少年推迟性关系的发生，或促使有偶遇性伴的人使用安全套。人类性行为受许多因素的影响和塑造，艾滋病预防项目涉及的只是其中之一。

对于行为改变是否能归因于预防项目的问题，在出资方和项目执行者间可能会产生分歧，特别是早期有艾滋病流行的国家。他们在此问题上的不同思考角度也会导致对项目评估过程有的看法根本不同。从公共卫生角度看，观察到的结果是否属干预所致的公共卫生意义是同样的。最重要的是性行为变得更安全、艾滋病感染实质性地减少。

从成本－效益分析或政策角度看，确定引起所观察的性行为变化的原因非常重要，如果不是特别经过设计的干预所引起的变化结果，干预的成本就可以考虑投到更合适的其他领域。为了有效分配资源，在压力不断增加的情况下，预防项目应提前考虑哪一种方式对于不同流行背景的特定目标人群更有效。因此，效果评估非常关键，因为它能回答一个基本问题：此项目是否起到明显作用？

项目效果评估中,有一个困扰的问题,就是如何从观察到的大量产出和影响中区分出哪些是项目带来的。这种估计需要运用各种思辩的方法而不是得出片面性的结论。使用的普遍原则是:设计越严格,结果估计越可信。研究设计预期得到的层次分明的证据能增加所下结论的可信度,即所观察到的行为改变在多大程度上是由于干预项目引起的,而不是其他因素的结果。

表 1-3 为按提供证据说服力递减的顺序排列的研究设计。在行为产出评估中,不能使用没有对照组的非实验性观察方法。但是,在项目提供服务效果方面,没有比较组的设计仍然是有效的,但对于测量项目的效果而言还不够确定。这种设计的原因和结果的逻辑关系不十分紧密,因为不能对时间横断面意义上的行为变化作出有说服力的解释。

在一些情形下,如评估某个干预项目或特定因素的暴露情况时,需要确定结果与暴露的联系程度,在对服务或干预对象缺乏联系或参与干预的对象不能准确报告的情况下也可有限度地使用结果归因的评估方法。

对项目评估数据应谨慎解释。大多数情况下,项目和评估过程作为一个整体并不是一个严谨控制的实验性调查。评估确定项目效果真实程度的能力经常受限于时间、资源以及设计方案的局限性。许多因素能干扰、混杂干预项目结果。偏倚会通过 与艾滋病有关的混杂因素、使用的测量选择和实施评估人员而产生。评估中最难回答的问题之一是如何使用测量方法进行评估,定义交互和重叠影响因素也是困难之一。这就是为什么许多项目难于将结果归因于干预行动。在某种程度上,我们不必担心如何进行归因分析,而应关注测量发生的变化。

表 1-3 研究设计：证据的等级层次（按证据说服能力递减）

1. 实验性研究 (1) 随机对照研究
2. 半实验性研究 (1) 非随机对照研究
3. 观察性研究 (1) 现况对照的队列研究 (2) 历史对照的队列研究
4. 没有对照的横断面调查 (1) 目标人群的重复调查 (2) 干预前－后调查

8. 多视角测量的方法

通过使用多种数据来源、不同研究者、多种角度解释单独一组数据或对同一个项目、问题或主题运用多种方法来作到多视角测量。

在没有严谨对照实验的情况下,数据的多视角测量可以运用在干预与观察到的行为变化间有实质性联系的情况下。例如,关于安全套销售的过程评估数据、同伴教育的力度、媒体宣传的质量和覆盖情况等。可以结合

对行为产出数据的分析来了解干预取得的效果。因为结果是由多种干预措施与环境、个人因素综合作用产生的。来源于行为调查的定量数据应该与对目标人群进行定性调查研究后得到的发现结合起来分析。定性调查可用专题小组讨论、目标人群访谈、快速人群研究等方式进行，以使观察到的定量数据得到更合理的解释。

艾滋病项目评估中需要测量的许多领域是敏感和隐私的，如性行为或对艾滋病感染者/病人的个人态度等调查。合法性和可靠性是对性行为调查研究的关键，因为性行为不能被直接观察到，性行为自我报告在没有其他证据的情况下通常被出资方认为是不合法和不可靠的，对于他们来说性行为数据是敏感的和与文化道德违背的。

增加合法性和可靠性的最好方法之一是使用所有可以利用的生物、行为和过程数据而得到多种角度的行为数据，以便来解释更多的性行为内容（第十一节对此主题提供更多细节内容）。

在许多国家进行的大量艾滋病研究中，第二手信息是多视角测量的方法。最常使用的多视角分析技术包括快速的人群分析结合使用分布地图、参与观察和深入访谈等半结构信息收集方法；专题小组讨论与个人深入访谈、目标人群访谈等相结合的方法，被广泛运用来获得出资方对目标人群行为调查认可的多视角测量方法。（第十二节对定性评估工具提供更多信息）。

非常重要的一点是行为改变干预实施时间要长、范围要广，以实现对个人行为、社区社会准则以及最终流行的影响。泰国的例子说明在全国范围内实施专门的干预措施能导致艾滋病流行和发病在目标人群中实质性下降。这也是运用多视角定量数据分析的一个例子：通过哨点监测系统、行为监测数据系统收集性传播感染/艾滋病的流行趋势的定量数据，为解释观察到的流行趋势提供必要的实施干预效果信息。人们对国家项目需要监测危险行为趋势正日益达成共识。（行为数据使用参见第八节）

9. 指标的选择

设计和实施艾滋病项目评估的关键步骤之一是选择适当的指标，如果项目的目标描述已很明确，（包括从定性、定量、时间等方面），那么指标的选择过程将非常顺利。然而，即使对于定义清楚的目标，许多项目评估的指标选择依然要求仔细从理论和实践两方面结合起来加以考虑。

以下问题有助于选择指标

- 与目标有关的指标是否是一个既准确又能可靠测量的参数
- 是否还有其他可以考虑的测量方法
- 该指标需要什么资源（人、财、物）保障条件
- 在指标的内容或数据收集方法方面在其他相关的领域有否可借鉴的地方
- 是否还有其他指标有助于解释目标的结果

选择指标和设定目标通常在项目的设计和再次计划阶段已完成，执行机构和主要出资方应共同参与这项工作。目标和标准的设定应参考来自相关项目的信息，使目标设定更多考虑目标人群的实际特点、资源分配和干预的类型。

当现有的测量标准达不到指标本身要求时，测量效果的数量多少将影响到目标人群的大小，这也促使评估者选择有利于解释结果的其他辅助指标，理想情况下，指标应作到：

- 可靠：它们应该按预定的情形进行测量
- 可信：相同条件下测量时应得到相同的结果
- 具体：它们应该体现需要测量的内容
- 敏感：在预定的条件下或在被观察的情况下，应反映变化情况
- 可操作：按明确的定义和参照标准能被测量和量化

■ 费用可承受：测量指标的成本是合理的

■ 可行：所提出的数据收集方法可以实施

可靠性与指标的实际内容有本质的联系，这也取决于指标的可测量性。可信与测量指标的方法和实施测量的人有重要关系，许多被人们熟悉的艾滋病预防产出指标在可靠和可信方面对评估者提出了挑战，如安全套使用的测量。解释促进安全性行为干预的产出指标有内在的复杂性。实际上，危险性行为是用相近的数字测量的，例如：安全套使用的百分数是测量安全性暴露的比例，但这并不能反映将个体暴露在性传播疾病危险中的绝对数量，在10次与艾滋病感染者有关的性行为中有10%使用安全套仍然比100次与艾滋病感染者有关的性行为中有75%使用安全套更安全（分别是9次和25次无保护的性行为）。因此，在危险的情况下确定安全套使用的绝对数量还是重要的。在偶遇性伴的性行为调查中，人们开始注意收集安全套使用的其他有关“经常、总是”等频度的数据。

表1-4列出了不同水平的项目评估中可能使用的指标。将相关指标用于具体评估的优点是可以对有分歧的数据进行识别，例如，安全套使用的调查数据能与相关领域安全套出售的信息相比较，或者在调查人群中性传播疾病的数据能与安全套相关数据比较。

在与国内和国际社会合作中，联合国艾滋病规划署和世界卫生组织已为国家规划制定出一套评估指标，而且，因为艾滋病/性传播疾病的预防和关怀项目受许多因素影响，包括政治承诺、可使用的资源、社会—文化和经济背景，人们正在制定新的方法来达到国家艾滋病项目的目标，“AIDS Program Effort Index (API)”是国家有效应对内容的综合评分。

这套评估指标的优点是，即使在没有更多严谨监督与评估系统的情况下，它仍然能在上述领域得到有用的信息。使用关键信息评估方法还能在一些难以得到更多客观测量指标的领域进行评估，如政治支持和承诺。然而，对“AIDS Program Effort Index (API)”的客观性和可信性人们仍然存在着担心。分值主要依靠对信息的选择，但信息容易随时间发生变化，如果仅使用单一方面的分值，人们担心在某个领域的改进会被其他领域的恶化所掩盖。

表 1-4 项目评估水平的指标举例

评估类型	指标
评估过程	
投入	■ 资源分配（如国家预算的百分比） ■ 安全套在中心地区的可得性
投入	■ 艾滋病传播知识 ■ 安全套在周边地区的可得性 ■ 12-17岁人口接受艾滋病教育/性卫生教育的比例 ■ 改进质量的服务比例（如性病病例管理、自愿咨询和检测、关怀艾滋病感染者/病人） ■ 输血机构提供不间断的艾滋病筛查检测的百分比