

**A FRAMEWORK FOR THE EVALUATION
OF NATIONAL AIDS PROGRAMMES**

**Ties Boerma, Elizabeth Pisani, Bernhard Schwartländer,
Thierry Mertens**

January 2000



MEASURE
Evaluation

Carolina Population Center
University of North Carolina
at Chapel Hill
123 W. Franklin Street
Suite 304
Chapel Hill, NC 27516
Phone: 919-966-7482
Fax: 919-966-2391
measure@unc.edu
www.cpc.unc.edu/measure

Collaborating Partners:

Macro International Inc.
11785 Beltsville Drive
Suite 300
Calverton, MD 20705-3119
Phone: 301-572-0200
Fax: 301-572-0999
measure@macroint.com

John Snow Research and Training Institute
1616 N. Ft. Myer Drive
11th Floor
Arlington, VA 22209
Phone: 703-528-7474
Fax: 703-528-7480
measure_project@jsi.com

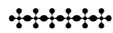
Tulane University
1440 Canal Street
Suite 2200
New Orleans, LA 70112
Phone: 504-584-3655
Fax: 504-584-3653
measure2@tulane.edu

Funding Agency:

Center for Population, Health
and Nutrition
U.S. Agency for
International Development
Washington, DC 20523-3600
Phone: 202-712-4959

WP-99-17

The research upon which this paper is based was sponsored by the MEASURE *Evaluation* Project with support from the United States Agency for International Development (USAID) under Contract No. HRN-A-00-97-00018-00.



The working paper series is made possible by support from USAID under the terms of Cooperative Agreement HRN-A-00-97-00018-00. The opinions expressed are those of the authors, and do not necessarily reflect the views of USAID.

The working papers in this series are produced by the MEASURE *Evaluation* Project in order to speed the dissemination of information from research studies. Most working papers currently are under review or are awaiting journal publication at a later date. Reprints of published papers are substituted for preliminary versions as they become available. The working papers are distributed as received from the authors. Adjustments are made to a standard format with no further editing.

A listing and copies of working papers published to date may be obtained from the MEASURE *Evaluation* Project at the address listed on the back cover.

Other MEASURE *Evaluation* Working Papers

- WP-99-16:** National trends in AIDS knowledge and sexual behaviour in Zambia 1996-98 (Charles Banda, Shelah S. Bloom, Gloria Songolo, Samantha Mulendema, Amy E. Cunningham, J. Ties Boerma)
- WP-99-15:** The Determinants of Contraceptive Discontinuation in Northern India: A Multilevel Analysis of Calendar Data (Fengyu Zhang, Amy O. Tsui, C. M. Suchindran)
- WP-99-14:** Does Contraceptive Discontinuation Matter?: Quality of Care and Fertility Consequences (Ann Blanc, Siân Curtis, Trevor Croft)
- WP-99-13:** Socioeconomic Status and Class in Studies of Fertility and Health in Developing Countries (Kenneth A. Bollen, Jennifer L. Glanville, Guy Stecklov)
- WP-99-12:** Monitoring and Evaluation Indicators Reported by Cooperating Agencies in the Family Planning Services and Communication, Management and Training Divisions of the USAID Office of Population (Catherine Elkins)
- WP-98-11:** Household Health Expenditures in Morocco: Implications for Health Care Reform (David R. Hotchkiss, Zine Eddine el Idriss, Jilali Hazim, and Amparo Gordillo)
- WP-98-10:** Report of a Technical Meeting on the Use of Lot Quality Assurance Sampling (LQAS) in Polio Eradication Programs
- WP-98-09:** How Well Do Perceptions of Family Planning Service Quality Correspond to Objective Measures? Evidence from Tanzania (Ilene S. Speizer)
- WP-98-08:** Family Planning Program Effects on Contraceptive Use in Morocco, 1992-1995 (David R. Hotchkiss)
- WP-98-07:** Do Family Planning Service Providers in Tanzania Unnecessarily Restrict Access to Contraceptive Methods? (Ilene S. Speizer)
- WP-98-06:** Contraceptive Intentions and Subsequent Use: Family Planning Program Effects in Morocco (Robert J. Magnani)
- WP-98-05:** Estimating the Health Impact of Industry Infant Food Marketing Practices in the Philippines (John F. Stewart)
- WP-98-03:** Testing Indicators for Use in Monitoring Interventions to Improve Women's Nutritional Status (Linda Adair)
- WP-98-02:** Obstacles to Quality of Care in Family Planning and Reproductive Health Services in Tanzania (Lisa Richey)

WP-98-01: Family Planning, Maternal/Child Health, and Sexually-Transmitted Diseases in Tanzania: Multivariate Results using Data from the 1996 Demographic and Health Survey and Service Availability Survey (Jason Dietrich)

A FRAMEWORK FOR THE EVALUATION OF NATIONAL AIDS PROGRAMMES

Ties Boerma(1), Elizabeth Pisani(2), Bernhard Schwartländer(3), Thierry Mertens(4)

(1) University of North Carolina, Chapel Hill, NC, USA

(2) UNAIDS consultant, Nairobi, Kenya

(3) UNAIDS, Geneva, Switzerland

(4) WHO, Geneva, Switzerland

Running title: Evaluation of national AIDS programmes

Biographies (brief version)

Ties Boerma is Director of the USAID-sponsored MEASURE *Evaluation* project and Research Associate Professor at the Maternal and Child Health Department of the School of Public Health, University of North Carolina, Chapel Hill. He is a medical demographer who has worked for 15 years in international health programmes, including 10 years in eastern Africa.

Elizabeth Pisani is a freelance consultant who has done extensive work in AIDS, including a major contribution to the UNAIDS Report on the global HIV/AIDS epidemic in 1998 and 1999. Elizabeth is trained as a journalist and demographer and is based in Nairobi, Kenya.

Bernhard Schwartländer is Chief of the Monitoring and Evaluation division of UNAIDS. He is an epidemiologist with main research interests focusing on the epidemiology of infectious diseases and evaluation research.

Thierry Mertens is an epidemiologist who has been working at WHO since 1992. He was chief of surveillance, evaluation and forecasting of the Global Programme on AIDS (GPA) of WHO in Geneva from 1992 to 1995, and played a major role in developing the GPA package for monitoring and evaluation of national AIDS programmes.

Address for correspondence:

Ties Boerma, Carolina Population Center Suite 304, 123 W Franklin Street, Chapel Hill, NC 26516, USA.

Abstract

At present, UNAIDS, WHO, and USAID are coordinating a global initiative to improve monitoring and evaluation of national programmes (UNAIDS/WHO/MEASURE Evaluation, 1999). This initiative builds upon the experience in monitoring and evaluation of a wide range of developing countries and aims to develop guidelines for a sound monitoring and evaluation system.

Experience in countries such as Thailand and Uganda has shown that there is a great need to demonstrate the overall impact of national AIDS programmes. Showing that programmes change behaviour and consequently reduce new infections is an essential component of building public support to sustain or expand the current budget and activity level of the national AIDS programme. One of the shortcomings of existing monitoring and evaluation efforts by national AIDS programmes is that they exist independently of a clear framework which links programme efforts and behavioural trends to HIV-related outcomes in a logical way.

This paper proposes a simple conceptual framework for monitoring and evaluation of AIDS programs, using the intermediate or proximate determinants conceptual framework used in the study of fertility and child survival. By specifying the programme outcomes as proximate determinants based on epidemiological theory the conventional input-output-outcome-impact monitoring and evaluation framework becomes conceptually clearer. The framework underscores the need for selection of indicators at different levels of the framework, ranging from context and programme inputs to health impact and mortality and emphasises the central role of the proximate determinants. The combination of monitoring data at different levels of the framework with disease surveillance data provides the most practical basis for the evaluation of national programmes.

Introduction

It is almost two decades since the emergence of the AIDS pandemic was first recognised. Since then national AIDS control programs have been established in many countries. A range of health and behavioural interventions has been implemented to reduce the spread of HIV and other sexually transmitted diseases and to minimise the consequences of AIDS for individuals, families and communities. While many interventions have been shown, individually, to be effective, it is still not clear what mix of interventions is most appropriate – and most cost-effective – in different epidemiological and socio-cultural settings.

Cohort studies of special populations and randomised controlled trials in community or clinical settings have been used to evaluate the effectiveness of specific interventions aiming to prevent HIV or other STD transmission (e.g., Grosskurth et al., 1995, Wawer et al., 1999). Such studies are vital. There is, however, an equally urgent need to demonstrate the overall impact of national AIDS programmes. AIDS programmes, finance ministries and donors alike need such information to ensure adequate funding for prevention and care activities. Showing that programmes change behaviour and consequently reduce new infections is an essential component of building public support to sustain or expand the current budget and activity level of the national AIDS programme. A few well-known country examples are often taken to show that successful national programmes can change the course of the epidemic. In Uganda, a rapid decline in HIV prevalence among young pregnant women has been observed in sentinel antenatal clinics in urban areas. Survey data on sexual behaviour were used to assess whether behavioural changes may have caused the decline in HIV prevalence and to what extent these changes in behaviour could be attributed to programmes (Asiimwe-Okiror et al., 1997). In Thailand, a national effort to promote safer sexual behaviour was followed by a rapid decline in visits to female sex workers and an increase in condom use. Shortly thereafter, STDs and new HIV

infections also fell, a fall recorded in a multitude of data sources (Nelson et al., 1996, Hanenberg et al., 1994, Mason et al., 1998, Bunnell et al., 1999, UNAIDS, 1998). In Senegal, epidemiological data indicate that HIV prevalence is low and stable, while socio-behavioural and programme information suggest a positive effect of a timely national response (Meda et al., 1999). In Jamaica, the analysis of a series of national knowledge, attitude, behaviour (KAP) surveys, programme information and HIV/STD surveillance among various population groups suggested that the national programme has slowed the spread of HIV and STD (Figueroa et al., 1998).

The extent to which countries have made an effort to monitor the implementation of their national AIDS programme and to evaluate their impact has varied considerably. Some tools do exist to help in these monitoring efforts, most notably a standardised set of Prevention Indicators developed by the Global Programme on AIDS of the World Health Organization (WHO, 1994, Mertens et al., 1994). This set of 10 Prevention Indicators, aimed at monitoring the progress of HIV-prevention programmes, was accompanied by data collection protocols to ensure that information could be collected in a way that was comparable over time and across populations. These indicators aim to capture trends in knowledge about AIDS, condom availability and use, the prevalence of non-regular sexual partnerships, the quality of STD treatment practices in health facilities and the prevalence of HIV and STD among young adults. Many countries have used the prevention indicators at some time (e.g., Mehret et al., 1996, Chilongozi et al., 1996), often adapting them to local circumstances. Only rarely have they been measured repeatedly over time and very few have been based on nationally representative samples. Another key data source on knowledge, attitudes and sexual behaviour at the national level is the Demographic and Health Surveys program (DHS). An increasing number of national DHS surveys include an AIDS module that provides information on HIV-related knowledge, attitudes and sexual behaviour, and is administered to women and men of reproductive ages.

One of the shortcomings of existing monitoring and evaluation efforts by national AIDS programmes is that they exist independently of a clear framework which links programme efforts and behavioural trends to HIV-related outcomes in a logical way. This paper proposes a simple conceptual framework for monitoring and evaluation of AIDS programs, using the intermediate or proximate determinants conceptual framework used in the study of fertility and child survival. This conceptual framework can guide the selection of indicators and data collection in the evaluation of national programmes. Indicators selected according to the framework will produce a sequence of information which, presented together, will provide more convincing evidence of the link between programme effort and ultimate outcome. Such national level programme evaluations are more difficult and more ‘contaminated’ than evaluation research focusing on specific interventions in defined local populations. However, in the sometimes dispiriting world of HIV prevention, the value of clearly demonstrated success at a national programme level cannot be overestimated (Mertens and Carael, 1997, Mertens et al., 1994).

Evaluation of national programmes

In the case of national AIDS programmes, the ultimate goals will be to reduce the spread of HIV, to improve care for those infected, and to minimise the social and economic impact on affected families and communities. Progress towards those goals is measured in a number of ways. Surveillance (the routine tracking of disease or, less commonly, risk behaviour) and programme monitoring (the routine tracking of priority information about a programme and its intended outputs and outcomes) are needed to track the implementation of programmes and the progress of the HIV epidemic itself. To evaluate the effectiveness of a programme, it must be possible to link a particular output or outcome directly to a particular intervention or programme.

Ideally, there are three phases in the evaluation of programmes at the national level. The programme's content, scope and coverage should be assessed, together with the quality and integrity of implementation. If findings from this process evaluation show that there is indeed an active and appropriate programme implemented on a scale large enough to produce any results, then outcome evaluation activities are designed to determine the programme's short-term outcomes – generally changes in behaviour, including treatment-seeking behaviour. If there is adequate evidence that the programme has achieved or is achieving its short-term objectives, then an evaluation of the longer-term impact on HIV infection can be undertaken.

In practice, this sequence has rarely if ever been followed in the evaluation of national AIDS programmes in developing countries. In most countries the HIV surveillance system forms the core of the monitoring and evaluation system – that is, the programme success is judged primarily on trends in HIV prevalence among pregnant women or groups at high risk of contracting or passing on the virus. Lower prevalence tends immediately to be equated with successful prevention. This is problematic for several reasons. Firstly, a drop in HIV prevalence does not

necessarily mean that new infections are falling. Changes in mortality, migration, sampling bias, fertility (in case of antenatal women), and even sexual activity may affect prevalence without affecting incidence (Boisson et al., 1996). Secondly, even if new infections are falling, it is not automatically possible to attribute this to programme success. It may be due to the natural course of the epidemic or other factors unrelated to programme efforts.

To draw any kind of link between changes in prevalence and programme effort, it is necessary at a minimum to examine intermediate behavioural variables which are the object of programme effort, such as condom use with casual sex partners, and to ensure that they are consistent with input and output variables such as the number of condoms distributed. Countries such as Uganda, Senegal, Thailand and Jamaica that are considered to have successful programmes have, to some extent, been able to show consistency between indicators of programme input and output, of changing behaviour and of impact (UNAIDS, 1998, Kilian et al., 1997; Asiimwe-Okiror et al., 1997; Meda et al., 1999; Bunnell et al., 1999). The link between trends in infection rates and behavioural change is most convincing in the case of Thailand where a very large data base allowed detailed and disaggregated analysis of trends in infection and behaviour. In Jamaica HIV and STD trends are monitored in different population groups, while five national surveys provide data on changes in sexual behaviour during 1988-1996. In Uganda and Senegal a much more modest behavioural database exists, derived from selected urban areas. Fairly limited data on the programme inputs, outputs and context were used to try to assess the extent to which changes are attributable to national and subnational programmes. Community-based research studies also contributed to assessment of trends in HIV and sexual behaviour in Uganda.

A framework for monitoring and evaluation

Input-output-outcome-impact frameworks, described in the introduction of the WHO Prevention Indicators (Mertens et al., 1994) and elsewhere (Rugg et al., 1999) are useful for the monitoring and evaluation of programmes. In such frameworks, programme inputs such as money and staff time may result in outputs such as trained staff, improved availability of services, etc. If these outputs are well designed and reach the targeted population, the programme is likely to have positive short-term outcomes (e.g., increased condom use), which should lead to long-term impact of programmes (e.g., reduced HIV incidence). Here, this traditional framework for monitoring and evaluation is expanded to better specify programme outcomes as intermediate variables between programme efforts and context and ultimate health impact using the proximate determinants model used in research on the determinants of fertility and child survival.

Proximate determinants: fertility and child survival

Models with proximate or intermediate determinants have been used extensively in the study of fertility. Almost half a century ago Davis and Blake (1956) developed an analytic framework for the comparative sociology of fertility. The key to the framework was a set of "... intermediate variables through which any social factors influencing the level of fertility must operate." (p. 211). This approach was expanded by Bongaarts (1978) who replaced the term intermediate variables by proximate determinants: "The proximate determinants of fertility are the biological and behavioural factors through which social, economic and environmental variables affect fertility. The principal characteristic of a proximate determinant is its direct influence on fertility." (Bongaarts and Potter, 1983: 1). According to Bongaarts the four most important proximate determinants of fertility are the variables marriage (entry into sexual union), breastfeeding, contraception and induced abortion. All four determinants are behavioural in the

sense that they are purposive actions, but they also have direct biological consequences: If the determinants change there is a direct effect on human fertility. Bongaarts also presented the relationship between proximate determinants and fertility as a simple statistical model. He assumed the existence of a natural fertility level, the maximum number of children a woman can have under optimal conditions. Natural fertility (about 16 children per woman) is subsequently reduced by the operation of the proximate determinants. Due to its conceptual clarity and the ability to quantify the effects of the proximate determinants on fertility, the Bongaarts' model found wide application in the study of fertility. It was also used to evaluate the impact of family planning programmes at the national level (Bongaarts and Potter, 1983, Bertrand et al., 1996). Decomposition of changes in the total fertility rate can be related to changes in the proximate determinants and thus the proportion of decline in fertility attributable to increases in contraceptive prevalence (and possibly programme factors) can be estimated.

Mosley and Chen (1984) used the same approach to develop an analytical framework for the study of the determinants of child survival. This framework integrated knowledge and methods from demography and epidemiology to promote interdisciplinary communication. The basic feature of the Mosley-Chen framework is the specification of a set of proximate or intermediate determinants that directly influence the risk of child morbidity and mortality. The proximate determinants include reproductive, hygienic, childcare (injury prevention), feeding and disease care practices. Socio-economic determinants, such as mother's level of education or household income, operate through the proximate determinants to influence child morbidity (including growth) and mortality. Van Norren further specified the proximate determinants and showed how the proximate determinants framework can be applied to monitoring and evaluation of primary health care programmes (Van Norren and Van Vianen, 1986, Van Norren et al., 1989). The Mosley-Chen framework has had considerable influence on international child health research. The translation into a statistical model – where the effects of the proximate determinants on child

survival can be measured – is much more complicated for child mortality than for fertility. A child's death is often the ultimate outcome of a series of illness episodes (translating into increased frailty of a child), while the proximate determinants of fertility affect a single biological event (conception) (Mosley and Chen, 1984). Becker and Black (1996) developed a statistical model to show how the proximate determinants may affect child health and mortality. As in the model of fertility, they determined a 'natural' level of disease incidence and case fatality – the disease incidence and case fatality in the absence of interventions such as good water, sanitation, breastfeeding or good delivery practices. Interventions were classified into those reducing disease incidence, frailty or case fatality. Multiple disease conditions interact through the mechanisms of competing risks and acquired frailty, resulting in high child mortality. Simulations showed how certain interventions, such as measles vaccination or reduction of vitamin A deficiency, potentially have a large, direct and indirect, effect on child survival.

Proximate determinants: HIV and STDs

The principal biological measures for HIV-prevention programmes are incidence and prevalence of HIV and other STDs. Epidemiological theory can help identify a set of proximate determinants in this context: behaviours that directly affect HIV or STD incidence and can be changed by interventions. The proximate determinants lie on the causal pathway that links a programme and its context to health impact. For HIV/STD transmission, as for other infectious diseases, the basic reproduction rate of infection – defined as the average number of susceptible people infected by an infected person over his/her life time – is of central importance (e.g., May and Anderson, 1987, Anderson, 1992). Models of transmission suggest that this reproductive rate of an infection is determined by (1) the efficiency of transmission during exposure between susceptible and infectious partners, (2) the risk of exposure of susceptible to infectious persons, and (3) the

duration of infectiousness. Most interventions in AIDS and STD control programmes address one or more of these components of the reproductive rate of infection (St Louis and Holmes, 1999).

A framework

Figure 1 presents an expanded version of the conventional monitoring and evaluation framework, including the proximate determinants concept. The introduction of the proximate determinants into the monitoring and evaluation framework implies a narrower and theoretically underpinned specification of the short-term outcomes. Epidemiological and biomedical theory dictates that if programmes or other societal changes succeed in favourably altering a proximate determinant (e.g., increase in condom use with casual partners), HIV or STD incidence will necessarily fall. This specification helps clarify the distinction between programme outputs and outcomes. For example, improved knowledge about HIV transmission is often an important goal of prevention programmes, but it is not a programme outcome in this framework. Improved knowledge is a programme output as it has no direct influence on the health outcome – HIV or STD incidence in this case. Improved knowledge affects the health outcome only if it leads to changes in sexual behaviour: reduced efficiency of transmission (e.g., by adoption of safer sexual practices including condom use) or reduction of the risk of exposure (e.g., in the number of sexual partners).

The left side of the conceptual framework includes the underlying factors that must operate through the proximate determinants in order to have health impact. This includes socio-economic and cultural factors, the policy context, health system issues, and specific factors related to the interventions of the AIDS programme itself. The ultimate effect of the program outputs on the proximate determinants is dependent on the context. For example, the effects of a training program for health workers to improve STD management is affected by the overall status of the

health system. Whether or not condom promotion and distribution leads to actual use may be affected by the religious and cultural context. The extent to which increased knowledge leads to changes in sexual behaviour depends on a range of individual- and population-level contextual factors. The context also includes the level of political commitment and the amount and type of program resources.

A different but important contextual factor is the extent to which the population has been exposed to increased HIV-associated mortality. The response to interventions in terms of effect on proximate determinants is likely to be larger if AIDS mortality is high and recognised. This may occur at the individual and at the population level. If an individual knows a family member who has died of AIDS, his or her knowledge of HIV transmission and AIDS is likely to be better, but it may also be more likely that actual sexual practices, i.e. the proximate determinants, change. In a community where mortality of young adults has doubled and where it is known that this is due to AIDS, young people are likely to be more receptive to behaviour change messages.

Even though AIDS programmes primarily focus on HIV prevention and STD control, care and support for people who are infected and reduction of the impact of the epidemic on families and communities are increasingly important components. These interventions, too, may affect a proximate determinant of infection. Better care for an HIV-infected person means a longer, healthier life. But it also may mean that infected people remain in the pool of infectious partners, increasing the chances of epidemic spread.

Indirectly, care and support interventions may affect HIV-prevention efforts. Essential aspects of care and support also feed back into the context level. For example, programmes may attempt to reduce the stigma surrounding HIV infection, because stigma may lead to active discrimination against HIV-infected people. But addressing stigma has a wider implication for prevention

efforts. Where HIV is highly stigmatised, for example, people may avoid condoms simply because they do not want anyone to think they are concerned about their own HIV status. A reduction in stigma surrounding HIV produces a more favourable context in which programme inputs and outputs might affect behaviours – i.e. the proximate determinants. In this example stigma reduction could contribute to an increase in condom use, and therefore have an impact on HIV incidence.

Indicators

The selection of a set of indicators for monitoring and evaluation of a national programme is essential. Good indicators need to be relevant to programmes, feasible to collect, easy to interpret and able to track changes over time. Tracking changes in indicators over time will help programme managers and decision-makers tell how successful the national programme is in meeting its goals.

Evaluation of national programmes should rely on sound monitoring of programme context, input, output, the proximate determinants of HIV infection, and impact. The analysis and interpretation of trends in indicators at different levels form the basis for evaluation of the national programme. Taken together, monitoring indicators can act as an evaluation of the national response as a whole. They give programme managers and decision-makers an idea of whether the sum total of all HIV-related efforts in a district, region or country is making any difference in terms of slowing the epidemic spread of HIV or reducing its impact on individuals and families affected.

The national monitoring and evaluation strategy depends on the type of epidemic, the programmatic response and the availability of resources. As a first step, programmes should monitor their inputs and outputs. Unless these change, any change in outcome can not in any case be ascribed to programme effort. Input and output indicators are often relatively easy and cheap to collect; where they register change, they indicate the need for monitoring and evaluation at the outcome or impact level. Indicators should be chosen to measure change in areas of programme effort. Since most national AIDS programmes tailor their responses to the state of the epidemic in their country, it follows that the appropriate indicators will also differ according to epidemic state. UNAIDS currently distinguishes three broad types of epidemics, using numerical proxies of

infection levels in various population groups to determine the epidemic state. In generalised epidemics, HIV prevalence is over one percent in the general population. In concentrated epidemics, HIV prevalence is over five percent in any sub-population at higher risk of infection (such as drug injectors, sex workers, men who have sex with men), but below one percent in the general adult population. In low-level epidemics, HIV prevalence is below five percent in any group.

It is beyond the scope of this paper to present and discuss all priority indicators for each level of the conceptual framework by type of programme effort and type of epidemic. The following discusses selected key issues related to selection of indicators at different levels of the framework for monitoring and evaluation (see also Table 1 for examples).

Inputs and outputs

The choice of input and output indicators clearly depends on what programmes aim to do. Increasing condom use is often a key programme objective and monitoring includes indicators of the distribution, availability and quality of condoms. Similarly, monitoring of programmes that aim to deliver specific services, such as voluntary counselling and testing, reduction of mother-to-child transmission, treatment of bacterial STDs and care and support for people and families affected by HIV/AIDS, should include measures of knowledge of the programme among intended beneficiaries, availability and quality of services, and service utilisation.

Without anti-retroviral treatment the inevitable outcome of HIV infection is illness and premature death, and interventions are aiming to mitigate the effects of HIV/AIDS on individuals, families, communities and societies. A limited number of indicators have been suggested to monitor the effects of interventions that focus on the improvement of the quality of care and support during

HIV infection, especially during the terminal stages, and monitoring of the societal efforts to ensure non-discrimination (Mertens and Carael, 1997). With regard to medical care these indicators include training in management of HIV/AIDS patients, availability and use of anti-retroviral drugs, availability and use of drugs against opportunistic infections, adequacy of nursing care, etc. Initial experience with indicators that measure the quality of care showed that such data are difficult to collect and difficult to interpret.

There are currently few indicators defined in the area of psychological and social support, partly because programme activities in these areas tend to be weaker than in prevention. Indicators of the political, legal and attitudinal context in which a programme operates are also little tested. At the societal level, indicators are being developed around the presence and application of non-discriminatory acts, legal rights, measures to ensure adequate support, etc. (Mertens and Carael, 1997). Indicators of political context are also being field tested, using a methodology that has been used to assess national family planning effort scores (Mauldin and Ross, 1991). The index is based on the opinion of a number of “experts” in a country.

Input and output indicators are also being developed for areas of care and support ranging from the availability and quality of voluntary counselling and HIV testing services to the coverage of home-based support services for people suffering from chronic illness likely to be related to HIV. These indicators are in varying stages of development: protocols for some are still being fine-tuned while others have been tested in a number of countries and contexts.

Proximate determinants

At the national level a number of indicators are needed to track changes in the proximate determinants. Each indicator has its advantages and disadvantages, but it is beyond the scope of this paper to discuss these issues (see UNAIDS/WHO/MEASURE Evaluation, 1999). The most important intervention to reduce the *efficiency of transmission* is promotion of condom use, especially in casual or commercial sexual interactions. In surveys, data on condom use during the last sexual act by type of sexual partner are frequently collected. Other risk behaviours targeted by prevention programmes that can be tracked by indicators include unprotected anal intercourse and needle sharing – behaviours that increase the chances of infection if exposure occurs. Indicators also focus on reduction of critical co-factors that increase infectiousness such as the improved treatment of other STDs and the provision of suppressive chemotherapy that reduces infectiousness – providing anti-retroviral drugs to reduce mother-to-child transmission is a common example of the latter.

A range of health education interventions aims to reduce the risk of *exposure of susceptible to infected persons*. Interventions aiming at reducing the exposure of susceptible individuals to infected partners include the promotion of abstinence and later sexual debut, monogamy and reduced rates of sexual partner change. Other interventions aim at modifying behaviour of known infected persons, for example through voluntary counselling and testing programmes. Models of STD/HIV transmission and population-based surveys have highlighted the importance of sexual mixing patterns or sexual networks in determining the epidemiology of STDs. The rate of new sexual partner acquisition occupies a central role (Anderson 1999: 28). Perhaps one of the most important issue in measuring sexual behaviour through surveys is how to filter relationships to get an idea of levels of risk involved. This question becomes more vexed as prevalence in the general population rises and the lines between “high-risk” partners such as sex workers and “low-risk”

partners such as husbands become blurred. The matter of central interest is not numbers of partners but patterns of sexual networking, and this is all but impossible to analyse with simple indicators. To date, the most common way of dividing relationships into high and low risk has been using a simple measure of time: any (non-marital) relationship that has lasted or is expected to last for more than a year is classified as “regular”, while any other relationship is classified as “non-regular”. Alternatively, relationships can simply be divided on the basis of cohabitation or marital status. Sex with any non-cohabiting partner is considered to be higher risk than with a cohabiting partner, regardless of the duration of the relationship. Furthermore, attempts are now being made to collect information on the last three partnerships in large-scale surveys.

For HIV/AIDS, the third proximate determinant – the *duration of infectiousness* – is often thought to be least amenable to interventions. Anti-retroviral treatment may affect HIV incidence by shortening the duration of infectiousness, although there is no evidence of this effect yet. Care programmes may also have the opposite effect, increasing the span of sexual activity of an HIV-infected person. It is also important to make the distinction between levels of infectiousness. For the spread of other STDs – notably bacterial STDs – improved access to, utilisation and quality of medical treatment shortens the duration of infection. STDs are an important co-factor for the efficiency of transmission of HIV (the first proximate determinant) and effective treatment of STDs reduces the period of heightened infectiousness for an HIV-infected individual (or, indeed the period of heightened susceptibility to HIV in those who have STDs but not HIV). Active case finding or partner notification can also help reduce STD prevalence, with a positive effect on infectiousness of HIV.

Impact

The direct health impact in the framework is HIV or STD infection, while other impacts include mortality and socio-economic status of affected families. Obviously, HIV incidence is the best indicator of new infections, although this is virtually impossible to track in routine surveillance. In certain circumstances, HIV prevalence among young adults may be used to estimate incidence (Zaba et al., 1999). Several studies have assessed the biases that may be associated with HIV surveillance of specific populations, especially women attending antenatal care clinics (Boisson et al., 1996). With regard to the latter, important biases include utilisation of antenatal care, selection for sexual activity and HIV-associated reduction of fertility (Zaba and Gregson, 1997). At the national level, bias is often also introduced by the selection of antenatal HIV sentinel surveillance sites. Rural areas are often severely underrepresented.

If a good clinic-based reporting system exists, e.g., an STD sentinel surveillance system, such data can be used to estimate trends in incidence of specific STDs, as was done in Thailand. Data on syphilis testing (RPR test) among antenatal women can also be used for surveillance purposes.

Finally, adult or child mortality trends can provide information on the health impact of the epidemic and the impact of anti-retroviral treatment. If no good vital registration data exist – as is the case in most developing countries – household surveys can be used to measure adult or child mortality trends. Since such surveys require large samples and are costly, two surveys are often held many years apart (e.g., five or ten years). Data on the prevalence of orphanhood can also be obtained in such household surveys.

Statistical model

Can the effects of changes in the indicators of sexual behaviour (e.g., an increase in condom use in commercial sex acts) on HIV and other STD incidence be quantified? To quantify the effects of the proximate determinants of fertility, a natural level of childbearing in the absence of interventions was assumed and the fertility effects of changes in contraceptive use or other proximate determinants could be 'decomposed' using empirical data (Bongaarts, 1978). In the study of child survival, a natural level of disease incidence and case fatality in the absence of interventions were used and the complex effects of specific interventions or combinations of interventions can be assessed, but only through modelling (Becker and Black, 1996). The specific nature of current STD and HIV epidemics – with often a very heterogeneous distribution of disease and risk behaviour within the population and emphasis on the stage of the epidemic – seems less suited to define a natural incidence of STD and HIV. The level and stage of an epidemic may also affect the magnitude of the impact that a change in a proximate determinant has on the change in HIV incidence. Furthermore, HIV is a chronic communicable disease rather than a one-time outcome such as pregnancy or child death. A statistical model would have to involve a complex feedback loop with variable assumptions about the infectiousness of infected partners. In this case, the basic reproduction rate of infection is the best indicator of impact, determined by the three groups of proximate determinants defined above. Mathematical models have been developed to study the dynamics of STDs and HIV in populations with different patterns of sexual behaviour and level of interventions and use high-low assumptions about the components of the reproductive rate of infections (Anderson, 1999). At present, no simple statistical formulae are available to decompose changes in HIV or STD incidence into changes in the proximate determinants.

The effects of specific interventions and programmes or changes in the context on the proximate determinants also need quantification. For instance, the extent to which increased knowledge leads to actual behavioural changes can vary from almost none to very substantial. Modelers have used a range of assumptions to make estimates of the effects of interventions on the proximate determinants and HIV incidence, but these have been difficult to confirm empirically. Special software is available to make such estimates of potential impact of interventions at the specific populations (e.g., simulAIDS, iwgAIDS, Bernstein et al., 1998, AVERT, Rehle et al., 1998).

Conclusion

Evaluation of national AIDS programmes relies on sound monitoring of programme context, input, output, the proximate determinants of HIV infection, and impact. The analysis and interpretation of trends in monitoring indicators at different levels form the basis for evaluation of the national programme. Taken together, monitoring indicators track the success of the national response as a whole. They give programme managers and decision-makers an idea of whether the sum total of all HIV-related efforts in a district, region or country is making any difference in terms of slowing the epidemic spread of HIV or reducing its impact on individuals and families affected.

Even though the conceptual framework cannot furnish simple measures to quantify the effects of the programme efforts on proximate determinants and on HIV or STD infection, it provides useful guidance for selecting indicators to monitor and evaluate interventions. By specifying the programme outcomes as proximate determinants based on epidemiological theory the conventional input-output-outcome-impact monitoring and evaluation framework becomes conceptually clearer. The framework underscores the need for selection of indicators at different levels of the framework, ranging from context and programme inputs to health impact and mortality and emphasises the central role of the proximate determinants. The combination of monitoring data at different levels of the framework with disease surveillance data provides the most practical basis for the evaluation of national programmes.

At present, UNAIDS, WHO, and USAID are coordinating a global initiative to improve monitoring and evaluation of national programmes (UNAIDS, WHO, USAID and MEASURE Evaluation, 1999). This initiative builds upon the experience in monitoring and evaluation of a wide range of developing countries and aims to develop consensus guidelines for a sound

monitoring and evaluation system. An additional goal is to select a set of common standard indicators that can be used to monitor and evaluate programmes or programme components. Instruments to collect information on these indicators are also developed. The conceptual framework presented here is used to guide the selection of indicators and assist national AIDS programmes to develop a comprehensive strategy. This strategy should ultimately lead to the best possible answer to the question of whether or not national programmes make a difference.

Acknowledgements

This paper has benefited greatly from comments by Basia Zaba, Amy Tsui, Amy Cunningham and Sharon Weir and from discussions with the participants of UNAIDS/WHO/ MEASURE Evaluation meetings on monitoring and evaluation of national AIDS programmes in Chapel Hill, North Carolina, USA, August 3-7, 1998; in Nairobi, Kenya, November 21-24, 1998; and in Talloires, France, May 22-26, 1999. The work by Boerma was funded by the USAID-sponsored MEASURE Evaluation project.

References

- Anderson R. 1992. Some aspects of sexual behaviour and the potential demographic impact of AIDS in developing countries. *Social Science and Medicine* **34**: 271-280.
- Anderson R. 1999. Transmission dynamics of sexually transmitted infections. In: Holmes KK, Sparling PF, Mardh P-A, et al. (eds). *Sexually Transmitted Diseases*. Third Edition. New York: McGraw Hill: 25-37.
- Asiimwe-Okuor G, Opio AA, Musinguzi J, Madraa E, Tembo G, Carael M. 1997. Change in sexual behaviour and decline in HIV infection among young pregnant women in urban Uganda. *AIDS* **11**: 1757-63.
- Becker S and Black RE. 1996. A model of child morbidity, mortality and health interventions. *Population and Development Review* **22**: 431-456.
- Bernstein RS, Sokal DC, Seitz ST, Auvert B, Stover J, Naamara W. 1998. Simulating the control of a heterosexual HIV epidemic in a severely affected east African city: computational modeling of single vs. combined intervention strategies. *Interfaces* **28**: 101-126.
- Bertrand JT, Magnani RJ, Rutenberg N. 1996. Evaluating family planning programs: with adaptations for reproductive health. The Evaluation Project: University of North Carolina, Chapel Hill. 1996
- Boisson E, Nicoll A, Zaba B, Rodrigues LC. 1996. Interpreting HIV seroprevalence data from pregnant women. *Journal of Acquired Immune Deficiency Syndromes and Human Retrovirology* **13**: 434-439.
- Bongaarts J. 1978. A framework for analyzing the proximate determinants of fertility. *Population and Development Review* **4**: 105-132.
- Bongaarts J, Potter JE. 1993. *Fertility, biology and behavior*. New York: Academic Press. 1983.
- Bunnell R, Yanpaisarn S, Kilmarx PH et al. 1999. HIV-1 seroprevalence among childbearing women in northern Thailand: monitoring a rapidly evolving epidemic. *AIDS* **13**: 509-515.
- Chilongozi DA, Costello Daly C, Franco L, Liomba NG, Dallabetta G. 1996. Sexually transmitted diseases: a survey of case management in Malawi. *International Journal of STD & AIDS* **7**: 269-275.
- Davis K and Blake J. 1956. Social structure and fertility: an analytical framework. *Economic Development & Cultural Change* **4**: 211-235.
- Figuerola JP, Brathwaite AR, Wedderburn M, et al. 1998. Is HIV/STD control in Jamaica making a difference? *AIDS* **12** (suppl 2): S89-S98.
- Grosskurth H, Mosha F, Todd J et al. 1995. Impact of improved treatment of sexually transmitted diseases on HIV infection in rural Tanzania: randomised controlled trial. *Lancet* **346**: 530-536.
- Hanenberg RS, Rojanapithayakorn W, Kunasol P, Sokal DC. 1994. Impact of Thailand's HIV control programme as indicated by the decline of sexually transmitted disease. *Lancet* **344**: 243-5.

- Kilian AHD, Gregson S, Ndyabangi B et al. 1999. Reductions in risk behaviour provide the most consistent explanation for declining HIV-1 prevalence in Uganda. *AIDS* **13**: 391-398.
- Mason CJ, Kitsiripornchai S, Markowitz LE et al. 1998. Nationwide surveillance of HIV-1 prevalence and subtype in young Thai men. *Journal of Acquired Immune Deficiency Syndromes and Human Retrovirology* **19**: 165-173.
- Mauldin WP, Ross JA. 1991. Family planning programs: efforts and results 1982-89. *Studies in Family Planning* **22**: 350-367.
- May RM and Anderson RM. 1987. Transmission dynamics of HIV infection. *Nature* 1987, 37-142.
- Meda N, Ndoye I, M'Boup S et al. 1999. Low and stable HIV infection rates in Senegal: natural course of the epidemic or evidence for success of prevention? *AIDS* **13**: 1397-1405.
- Mehret M, Mertens TE, Carael M et al. 1996. Baseline for the evaluation of an AIDS programme using prevention indicators: a case study in Ethiopia. *Bulletin of the WHO* **74**: 509-516.
- Mertens T, Carael M, Sato P et al. 1994. Prevention indicators for evaluating the progress of national AIDS programmes. *AIDS* **8**: 1359-69.
- Mertens TE, Carael M. 1997. Evaluation of HIV/STD prevention, care and support: an update on WHO's approaches. *AIDS Education and Prevention* **9**: 133-145.
- Mosley WH and Chen LC. 1984. An analytical framework for the study of child survival in developing countries. In Mosley WH and Chen LC (eds.). *Child survival: strategies for research. Population and Development Review* **10** (suppl.): 25-45.
- Nelson KE, Celentano DD, Eiumtakol S et al. 1996. Changes in sexual behaviour and a decline in HIV infection among young men in Thailand. *New England Journal of Medicine* **335**: 297-303.
- Rehle TM, Saidel TJ, Hassig S, Bouey P, Gaillard E, Sokal DC. 1998. AVERT: a user-friendly model to estimate the impact of HIV/sexually transmitted disease prevention interventions on HIV transmission. *AIDS* **12** (suppl.2): S27-S35.
- Rugg D, Buehler J, Renaud M et al. 1999. Evaluating HIV prevention: a framework for national, state and local levels. *American Journal of Evaluation* **20**: 35-56.
- St Louis ME and Holmes KK. Conceptual framework for STD/HIV prevention and control. In: Holmes KK, Sparling PF, Mardh P-A, et al. (eds). *Sexually Transmitted Diseases*. Third Edition. New York: McGraw Hill: 1239-53.
- UNAIDS. 1998. Relationships of HIV and STD declines in Thailand to behavioural change: a synthesis of existing studies. UNAIDS Best Practices collection 98.2. Geneva.
- UNAIDS, WHO, USAID and MEASURE Evaluation. 1999. National AIDS programmes: a guide for monitoring and evaluation. Chapel Hill and Geneva.

Van Norren B and Van Vianen HAW. 1986. The malnutrition-infections syndrome and its demographic outcome in developing countries. Programming Committee for Demographic Research, Publication No. 4. The Hague.

Van Norren B, Boerma JT, Sempebwa EKN. 1989. Simplifying the evaluation of primary health care programmes. *Social Science and Medicine* **28**: 1091-107.

Wawer M, Sewankambo NK, Serwadda D et al. 1999. Control of sexually transmitted diseases for AIDS prevention in Uganda: a randomised community trial. Rakai Project Study Group. *Lancet* **353**: 525-35.

WHO. Evaluation of a national AIDS programme. A methods package. 1. Prevention of HIV infection. GPA. WHO/GPA/TCO/ SEF/94.1.

Zaba B and Gregson S. Measuring the impact of HIV on fertility in Africa. *AIDS* 1998, **12** (suppl.1): S41-S50.

Zaba B, Boerma JT, White R. 1999. Monitoring the AIDS epidemic using HIV prevalence data among young women attending antenatal clinics: prospects and problems. Submitted for publication.

Table 1 Selected programme areas, illustrative indicators and data collection strategies by level of the conceptual framework for evaluation of national AIDS programmes.

	Input	Output	Outcome	Impact
Key programme area	AIDS education	Improved knowledge	Reduce number of partners	HIV incidence STD incidence
Indicators	Number of broadcasted radio programmes / messages that provide correct information about AIDS in a month	% of persons 15-49 knowing at least three ways to prevent sexual transmission of HIV	% of men 15-49 years who have visited a commercial sex worker in the last 12 months	HIV prevalence among young women (15-24 years) in antenatal clinics
Strategies to obtain data for indicators	Programme statistics	Population-based survey	Population-based survey or behavioural surveillance	Sentinel surveillance in antenatal clinics
Key programme area	Condom promotion	Increased condom availability	Increased use of condoms	HIV incidence STD incidence
Indicators	Number of condoms distributed in country	% of specified outlets with continuous availability of condoms	% of men/women who reported to have used a condom during their last non-marital sexual contact	Proportion of women in antenatal clinics with positive RPR test
Strategies to obtain data for indicators	Programme statistics	Survey of facilities and outlets	Population-based survey or behavioural surveillance	Sentinel surveillance in antenatal clinics (or other appropriate population groups)

Figure 1: Conceptual framework for monitoring and evaluation of national AIDS programmes.

