

**An Evaluation of Framework for
The National HIV/AIDS Strategy**

David Dunt

Associate Professor, Co-Director, Centre for Health Program Evaluation

Jeff Richardson

Professor, Co-Director, Centre for Health Program Evaluation

Jeanne Daly

Doctor, Department of Sociology, La Trobe University

CENTRE PROFILE

The Centre for Health Program Evaluation (CHPE) is a research and teaching organisation established in 1990 to:

- undertake academic and applied research into health programs, health systems and current policy issues;
- develop appropriate evaluation methodologies; and
- promote the teaching of health economics and health program evaluation, in order to increase the supply of trained specialists and to improve the level of understanding in the health community.

The Centre comprises two independent research units, the Health Economics Unit (HEU) which is part of the Faculty of Business and Economics at Monash University, and the Program Evaluation Unit (PEU) which is part of the Department of General Practice and Public Health at The University of Melbourne. The two units undertake their own individual work programs as well as collaborative research and teaching activities.

PUBLICATIONS

The views expressed in Centre publications are those of the author(s) and do not necessarily reflect the views of the Centre or its sponsors. Readers of publications are encouraged to contact the author(s) with comments, criticisms and suggestions.

A list of the Centre's papers is provided inside the back cover. Further information and copies of the papers may be obtained by contacting:

The Co-ordinator
Centre for Health Program Evaluation
PO Box 477
West Heidelberg Vic 3081, Australia
Telephone + 61 3 9496 4433/4434 **Facsimile** + 61 3 9496 4424
E-mail CHPE@BusEco.monash.edu.au

ACKNOWLEDGMENTS

The National Centre is supported by a grant from the National Health and Medical Research Council's Public Health Research and Development Committee and by a matching grant from the Victorian Health Promotion Foundation. The Research reported in this paper was made possible as a result of this support.

Personnel for the Preparation of this Document

Co-Directors

Dr D.R. Dunt, Department of Community Medicine, University of Melbourne.

Professor J. Richardson, Public Sector Management Institute, Monash University.

Project Officers

Dr Jeanne Daly, Department of Sociology, La Trobe University.

Meredith Temple-Smith (P/T), Department of Community Medicine, University of Melbourne

Priscilla Pyett (P/T), Department of Social Work, University of Melbourne.

Consultants

Neil Day, Centre for Programme Evaluation, University of Melbourne.

Dr Rod O'Connor, O'Connor and Associates, Melbourne

TABLE OF CONTENTS

List of Abbreviations

1	INTRODUCTION	1
1.1	Terminology.....	1
1.2	Aim of the Evaluation.....	2
	1.2.1 Summative and Formative Evaluation	2
	1.2.2 Criteria.....	3
1.3	Framework for the Evaluation.....	4
	1.3.1 Level of Focus of the Evaluation	4
	1.3.2 Methods.....	6
	1.3.3 Audiences for the Evaluation and Product of the	7
2	METHODOLOGY	9
2.1	The Strategy: qualitative methods	9
2.2	The Strategy: quantitative methods	12
2.3	Programs: qualitative methods	12
2.4	Programs: quantitative methods.....	14
2.5	Projects	15
3	PRINCIPLES UNDERLYING THE PROPOSED METHODS	16
3.1	Economic Evaluation.....	16
3.2	Prevention vs Treatment	17
3.3	On-going Research	18
3.4	Performance Indicators	18
3.5	Data and Data Collection.....	19
3.6	Confidentiality.....	20
3.7	Consultation.....	20
3.8	Multidisciplinary Research	21
4	EVALUATION OF SPECIFIC PROGRAMS.....	23
4.1	EDUCATION AND PREVENTION.....	23
	4.1.1 The Summative Evaluation of Prevention and Treatment.....	24
	4.1.2 The Summative Evaluation of Major Projects	28
	4.1.3 The Formative Evaluation of the Prevention and.....	29
4.2	TREATMENT, CARE AND COUNSELLING	30
	4.2.1 Summative Evaluation: Quantitative	31
	4.2.2 Summative and Formative Evaluation: Qualitative	33

4.3	ACCESS AND PARTICIPATION	36
4.3.1	Key Issues	36
4.3.2	Summative Evaluation: quantitative.....	37
4.3.3	Summative Evaluation: qualitative.....	38
4.3.4	Formative Evaluation	39
4.4	RESEARCH	40
4.4.1	The Summative Evaluation of the Research Program.....	40
4.4.2	The Summative Evaluation of Major Projects	41
	4.4.2.1 Key Issues	42
4.4.3	The Formative Evaluation of the Research Program Area.....	43
4.5	INTERNATIONAL CO-OPERATION	44
4.5.1	Regional Action	44
4.5.2	Key Issues	44
4.5.3	Summative Evaluation of Program.	45
4.5.4	Summative Evaluation of Major Projects	46
5	IMPLEMENTATION AND ORGANISATION OF EVALUATION	47
5.1	Proposed organisation	47
5.2	Staging	48
5.3	Selected commissioned studies	50
5.4	Selected Expert Reviews	51
6	INDICATIVE BUDGET	52
7	ANNEX	53
7.1	Personnel and Organisations to be consulted.....	53
7.2	Performance Indicators	56
7.3	Evaluation of National AIDS/HIV Strategy: The role of mathematical/statistical modelling	67
7.4	Selective Literature Review on HIV/AIDS Evaluation	71

LIST OF ABBREVIATIONS

AFAO	Australian Federation of AIDS Organisations
AIDS	Acquired Immune Deficiency Syndrome
ANCA	Australian National Council on AIDS
AZT	Azidothymidine
CARG	Commonwealth AIDS Research Grant
CUA	Cost Utility Analysis
GAS	Global AIDS Strategy
GPA	Global Program on AIDS
HIV	Human Immunodeficiency virus
IGCA	Intergovernmental Committee on AIDS
IV	Intravenous (as in drug use)
NESB	Non-English speaking background
NHMRC	The National Health and Medical Research Council
QL	Quality of life
States	Includes all Australian States and the two mainland Territories
STD	Sexually transmitted disease
WHO	World Health Organisation

1 INTRODUCTION

The National HIV/AIDS Strategy has the goals of eliminating transmission of HIV/AIDS and minimising the personal and social impact of HIV infection. The national evaluation team will have the responsibility for designing, conducting and implementing the evaluation of the Strategy.

The present document does not attempt to review the dimensions and magnitude of the HIV/AIDS epidemic, nor is it primarily concerned with a description of the existing Strategy. These are the introductory tasks of the actual evaluation. Rather, the purpose of this document is to outline how the team from the National Centre should conceptualise and implement such an evaluation.

1.1 Terminology

All major interventions such as the National HIV/AIDS Strategy develop their own vocabulary to express their unique approaches to their task. This evaluation should adopt the terminology recommended by the WHO¹. In the present context, the term 'Strategy' is used to refer to the national response by all Australian agencies, both governmental and non-governmental, to the HIV/AIDS epidemic. It subsumes all Programs and Projects. 'Programs' refer to a set of potentially broad-ranging activities oriented to a particular set of goals. In Australia, Programs have been developed in:

- . education;
- . prevention;
- . treatment, care and counselling;
- . access and participation;
- . research;
- . international cooperation.

The term 'Project' is applied to a particular component within a Program that can be specified by a particular set of similar activities, a united organisational structure or some other characteristic that limits its scope. For example, various 'needle exchange' facilities have been established as separate Projects.

¹ WHO, Global Programme on AIDS, Steering Committee on Evaluation Methods: Report of the First Meeting, Geneva, 21-23 June, 1989, WHO paper, reference WHO/GPA/DIR/89.9.
WHO, Health Programme Evaluation, Health for All Series, no 6, 1981.

1.2 Aim of the Evaluation

The most general aim of the evaluation is to provide public accountability as well as a means of improving the National HIV/AIDS Strategy. Evaluation is carried out for a practical policy-oriented purpose. Its end point should be guidance with respect to the need to maintain, alter, expand, contract, commence or terminate particular activities.

1.2.1 Summative and Formative Evaluation

There are two broad purposes that can be distinguished for most evaluative activity: **summative** and **formative**.

'Summative evaluation' involves making judgements about the effectiveness and efficiency of interventions in achieving the purposes for which the intervention was made. Other outcomes, often described as 'unintended', are also of interest here as they may measure the achievements of Programs, even though the achievements were not anticipated at their inception. A summative evaluation is mainly concerned with appraising the worth of an intervention and deciding whether the resources expended upon it are being used as effectively and efficiently as possible. In the terminology of economics, summative evaluation corresponds with cost benefit or cost effectiveness analysis in which costs and benefits are defined broadly to include all personal and social factors which have a positive or negative effect upon well-being.

By contrast, an evaluation may be described as 'formative' if its major objective is the modification and revision of interventions which are likely to continue at least in the short term.

A formative evaluation is oriented towards the improvement of an existing program, rather than making 'an end of term report'.

On occasions these distinctions become blurred. In practice, a summative evaluation may result in the modification of Programs. Decisions about what type of Programs should continue to be funded may be summative in the context of the individual Projects while the same decisions may be interpreted as making modifications, revisions and improvements at the level of the broad Program of which a special Project might be only a small part.

While recognising these qualifications, the present assignment is envisaged as being primarily a summative evaluation. That is, the questions to be answered are:

- . whether interventions carried out to date have been effective and efficient?
- . whether alternative outcomes or an alternative mix of activities would have been more effective and efficient?

As a secondary objective, formative evaluation may be undertaken with certain key Projects. The nature of the evaluation may, however, vary somewhat as described in Section 1.3.1 below.

1.2.2 Criteria

The overriding criteria for summative evaluation should be cost and health outcome where the latter is broadly defined as life and the quality of life (QL). This includes social, psychological and other dimensions of quality. It is, of course, fully recognised that there are often major impediments to the use of these criteria. In particular, it will often be necessary to evaluate processes and to establish performance indicators of processes. But it is important that the end point is not forgotten and that the relationship between process and outcomes are documented and incorporated in the assessment as far as possible. To achieve this the evaluation should commence with an exhaustive review of the state of knowledge with respect to all aspects of the prevention and control of HIV/AIDS.

1.3 Framework for the Evaluation

The framework proposed here is not definitive: it should be negotiated and refined in consultation with the IGCA and other interested groups and agencies. The framework encompasses a number of considerations. Primary among these are the issues outlined in Figure One: the purposes to which the product of the evaluation will be put, the administrative levels at which attention will be focused and the methods to be used. In addition there are a number of secondary considerations raised later in this section: the product of the evaluation and the audiences to which this is addressed.

1.3.1 Level of Focus of the Evaluation

HIV/AIDS interventions are categorised into a hierarchy of three intervention levels, from the most general to the most particular: Strategy, Programs and Projects.

In addition to defining the level in terms of the breadth or narrowness of a particular intervention there is also the administrative level, as indicated in the middle box of Figure One. The National HIV/AIDS Strategy is unique in the range of organizations it encompasses: central government at Federal level, regional government at State level, Local Government and non-government organisations and community groups. In the evaluation, it is proposed that attention should be focused on the federal and regional levels of this hierarchy.

For the overall Strategy and at the level of the six general and broad-ranging Programs, the evaluation should be summative, commenting on the achievements and shortcomings of the overall thrust of the Strategy. The goal here is to provide an objective, independent and methodologically sound appraisal of the overall outcomes of each Program.

As many aspects of the Programs consist in establishing and funding specific Projects, the summative review of a particular Program will include a summation of the effectiveness and efficiency of the components within it, although these may be made at a general and aggregated level. There will also be a consideration of possible Projects that have been omitted from the Program.

At the level below the broad Program, the evaluation will serve a somewhat different purpose. A concern will be to assist State administrations as well as selected Projects to establish their own procedures for evaluation, monitoring and improving Projects, including the further development of appropriate performance indicators. At these lower levels, the focus of the evaluation will be formative, assisting administrators and Project delivery personnel to develop the methodological skills needed to evaluate, refine and improve their own Programs.

Differences between the States in the scale of the Projects is of significance here. In the States where only a small expenditure has been necessary, the marginal cost of establishing evaluative procedures weighs disproportionately more heavily. One possibility that will be explored is for the evaluation team to provide workshops on the development of specific evaluative techniques, for instance for developing performance indicators, in those areas where regional administrations and Projects are not sufficiently well-funded to have sophisticated evaluation techniques already in place. The extent to which this occurs will be dependent upon budgetary factors. Formative evaluation would be limited to Projects that were perceived as being quantitatively significant in terms of their cost or outcome and where there appeared to be substantial benefits in this approach.

1.3.2 Methods

A detailed discussion of the methods proposed for the study are set out in Sections 2 and 3 of this proposal. There are two broad and overlapping categories of assessment referred to in this proposal. 'Quantitative assessment' refers to methods which attempt to estimate the magnitude of effects. These involve mathematical modelling, interview and mail questionnaire surveys and analysis of administrative data. When they are based on statistical sampling or a total enumeration (eg from administrative data records), and when data are valid, these methods provide a known degree of precision and permit generalizations to be made about a population.

The term 'qualitative assessment' is used here to refer to other, non-quantitative but nevertheless systematic attempts to evaluate effects. It subsumes a range of qualitative methods involving unstructured data collection techniques such as in-depth interviews, group consultations and documentary analysis. Data collection methods are flexible and the interpretation of the data allows an understanding of the complexity and variability within a given context.

Qualitative and quantitative assessments are often wrongly viewed as being independent or even competitive. Since quantitative data is necessarily limited in depth and richness of detail, the principle that will underline all quantitative analysis in the present study is that it will be preceded by a qualitative assessment of the factors which are of primary importance.

Further, the distinction between qualitative and quantitative has been blurred by recent methodological developments, for example the application of numerical techniques to the analysis of unstructured data. This proposal combines methods from these two broad approaches and ensures that the methods used are congruent with the varied requirements of the evaluation strategy.

1.3.3 Audiences for the Evaluation and Product of the Evaluation

While evaluation, as a form of public accounting, should be of general interest, there are 'audiences' with particular needs and concerns. These will necessarily influence both the emphasis within the evaluation and the nature of the final product.

The National HIV/AIDS Strategy is built on a unique combination of government and private or community organizations. Considerations of the audiences for the evaluation must be sensitive to this distinction, and reporting should take the differing requirements and backgrounds of the many audiences into account. In particular, those who have contracted the HIV virus and their friends and relatives will comprise an especially sensitive audience for the evaluation.

At an early stage, consideration must be given to the product of the evaluation, that is to say the results that are expected to come out of it: reports, recommendations, changes in staff attitudes and abilities, and implementation of any recommended change.

The product of a summative evaluation is usually a report, written at a level and in a form appropriate for the intended audience. Given the range of audiences for the summative aspects of the evaluation, the product should encompass written and graphic material suitable for opinion leaders, community groups and members of the public. For government funding bodies, the product should encompass more formal reports and cost benefit considerations, as well as substantial detail about methods used in gathering and analyzing data. A summative report can be a single event and the response of the appropriate agencies need be no concern of the evaluator who provides a substantively and methodologically sound summation of the achievements and shortcomings of the Program.

It is envisaged that all summative assessments will commence with a qualitative analysis of the scope, nature and importance of the factors that have a positive or negative effect upon well-being or, to use economic terminology, costs and benefits. Where feasible these factors will be

quantified. In the case of costs, these are sometimes, (but not always) measured in dollars. Benefits can sometimes (but not always) be measured as lives saved (or life years saved) and dollars gained. When costs and benefits cannot be directly measured an attempt will be made to find a quantitative indicator or index which is believed to be closely related to the relevant outcome.

In the formative elements of the evaluation there is scope for a varied product. Workshops for administrators, Project delivery staff and consumers of the services are all options that could be canvassed. If the establishment of formalised evaluative routines is a desired outcome, some training in the development and application of appropriate methods will be required.

2 METHODOLOGY

The relationship between the purpose, level and methods of evaluation are shown in Figure Two below. This also indicates the proposed priority for each cell in the matrix.

In Figure Two the formative evaluation of the overall Strategy and of particular Projects is marked as being of lower priority. This is subject to two qualifications. First, the outcome of a summative evaluation of Projects could be viewed as a formative evaluation of Programs to the extent that particular Programs may be shown to be ineffective or that they should be modified or expanded. A similar relationship exists between the summative evaluation of Programs and the formative evaluation of the overall Strategy. Secondly, it is possible that particular Projects or activities could be identified that have not been undertaken within Australia but that appear desirable for a particular aspect of the Strategy.

2.1 The Strategy: Qualitative Methods

The first aim is to develop a comprehensive overview of what is happening in Australia as a result of the National HIV/AIDS Strategy. This will include details on the range and structure of the activities included in the Strategy. Data will be drawn, as far as possible, from existing sources.

The second aim is to compare the Australian achievement with that of other countries.

Data for the achieving the analysis of the Australian achievement will be drawn substantially from documentary sources and through consultation in a series of forums, at least one in each State. Key participants for these forums include:

- . representatives of government and non-government organisations;
- . political parties and representatives;
- . law enforcement organisations, health education and welfare groups;
- . service users and groups representing them;
- . relevant community groups;
- . service providers;
- . public commentators.

One objective of the forums is to ensure that the evaluation team has a comprehensive understanding of the position of Australia in relation to HIV/AIDS; its geographic situation (in relative isolation but proximity to Asian countries); its population structure (class and ethnic diversity; gay communities); and its laws and welfare policies. It will also consider differences

between States both in the rate of diffusion of the disease and in the existence of populations who might require special Programs for treatment or prevention.

The evaluation will identify:

- . the appropriateness of the Strategy and its objectives for Australia;
- . the extent to which the Strategy is accepted by the target populations;
- . the achievements of the Strategy: intended and unintended, positive and negative;
- . the extent to which specific initiatives have succeeded;
- . omission and shortcomings;
- . the sociopolitical context and its effect both upon the range of Programs instituted and their achievements.

The next step is to undertake a comparison of the Australian and international achievements in order to provide an overview of:

- . the range of policy options available in Australia in comparison with other countries;
- . the state of knowledge in comparison with other countries.

The relative performance of different countries will be assessed by comparing the content and emphasis of their Programs with the effectiveness of such activities as revealed in the literature and from currently available research.

2.2 The Strategy: Quantitative Methods

It is recognised at the outset that the quantitative assessment of the overall impact of the Strategy is fraught with almost insurmountable problems. The Strategy is not simply the sum of the component Programs and Projects as there will be synergistic effects and it will not be possible to quantify the impact of some of the constituent Projects. The basic methodological problem is that, while it is possible to monitor the progress of the HIV/AIDS epidemic in Australia, it is not possible to observe or monitor the counterfactual situation that would have occurred in the absence of the Strategy. Nevertheless it is believed that an attempt should be made to quantify the overall impact of the Strategy no matter how crude this attempt may be.

The approach will use mathematical modelling. At least two variants of this approach can be adopted. First, the counterfactual situation in which there was no HIV Strategy could be simulated from a base year and compared with the actual Australian experience. Secondly, and methodologically the preferred option, the simulated situation in the absence of the Strategy could be compared with the simulated outcome with the Strategy. The difference which would be attributable to the Strategy would abstract from other shortcomings of the mathematical model. See Section 7.3 for a more detailed discussion of issues relating to mathematical modelling.

2.3 Programs: Qualitative Methods

The evaluation requires a comprehensive description and tabulation of all HIV/AIDS Programs in Australia including information on:

- . objectives;
- . activities;
- . budget;
- . linkages to client groups (including number, type and location);
- . utilization of services by client groups: whether it is equitably distributed;
- . client satisfaction with a service: what they perceive as the benefits;
- . data collection (i.e. the processes and availability of relevant data collection);
- . social and political constraints on the implementation of the Programs.

Within each Program, an important source of information for the evaluation is the experience of community groups and service providers. Expertise in these groups is an important but under-utilised resource both for understanding how a Program works and for determining what its constraints are. While the knowledge held by each group is partial, it contributes to a model of the whole. It also provides a check on conclusions reached.

There is, in addition, a broad range of 'opportunistic', unstructured data (eg accounts of community experience, media accounts, the general literature, professional accounts). Synthesising this broad range of qualitative data into a coherent analysis can be simplified by the use of computer programs for the analysis of unstructured data.

As in the evaluation of the overall Strategy, where existing data are inadequate, further information will be sought by interview and consultation with key informants. Forums will be held for each Program to ensure a comprehensive understanding of goals and achievements.

The end-product will be an overview of:

- . the goals and objectives of each of the Programs and how these contribute towards the overall Strategy;
- . the state of knowledge in each Program area;
- . an assessment of effectiveness;
- . the sociopolitical context: its effects on the range of Programs which has/has not been instituted and the achievements.

The implementation of the evaluation will require consultation on methods of client-responsive data-collection which can be set in place as a result of the evaluation.

Feed-back on preliminary findings will be sought from within each Program.

2.4 Programs: Quantitative Methods

Determining the outcome of particular Programs encounters the same methodological problem that was described above for the overall Strategy. That is, the counterfactual situation in which there is no Program is very hard to simulate. However, as the goals of the Programs are more specifically targeted there is a greater scope for the development of performance indicators of their success as judged by the criteria outlined in the Annex (Section 7.2). Another important task here is to establish, either from the literature or from consensus, that there is a relationship between the indicator and the final objective, namely the length of life and the quality of life. The feasibility of the two approaches to the evaluation of the Strategy will also be examined in the context of the Program.

The final report will contain an explicit statement of the validity and reliability of the analysis for the purpose of prediction. Areas of greatest uncertainty indicate where there are the most serious gaps in our knowledge and where future efforts should therefore be concentrated. However, if the costs involved are low this may not be judged worthwhile.

Where appropriate, modifications will be proposed including evaluation methods for incorporation into new Programs or Projects and the development and implementation of new performance indicators for existing Programs and Projects. While the social and political context is recognised as an important determinant of the progress of the HIV/AIDS epidemic, it is also an important area where change may be needed in order to increase the efficacy of any National Strategy. Specific recommendation regarding legal, administrative and other structures will be made.

2.5 Projects

As noted in the overall framework there should not be an attempt to evaluate every Project in the country. This task is beyond the scope and budget of the evaluation. Nevertheless, particular Projects will be examined when they are quantitatively very costly or where apparent benefits are very large.

The qualitative and quantitative assessment of selected Projects will follow the same steps and principles as outlined for the Programs. As the objectives are more narrowly defined it will generally be more feasible to quantify effects or to obtain satisfactory performance indicators and attribute changes in these indicators to the impact of particular Projects.

3 PRINCIPLES UNDERLYING THE PROPOSED METHODS

In this Section we outline additional principles which should be incorporated in the evaluation.

3.1 Economic Evaluation

The need for explicit economic evaluation arises from the recognition that the resources available to combat the HIV/AIDS epidemic are finite. If they are used inefficiently in one area then there will be fewer resources available for other activities within the Strategy. That is, economic inefficiency within a fixed budget necessarily means that less will be achieved and that the control of the epidemic will be less effective.

In principle, Cost Utility Analysis (CUA) provides the most appropriate framework for summative evaluation of the sort to be undertaken. It is explicitly designed to compare the cost of a Program with its achievements or impact expressed in terms of likely life years saved adjusted for the quality of life. In practice the application of CUA will be limited by the availability of data and the inapplicability of the existing techniques of CUA to some aspects of the HIV/AIDS evaluation. The principles of CUA will, however, be applied wherever possible. That is, evaluation will explicitly consider cost (both social cost and cost to the health budget), life and the quality of life broadly defined.

As there is now an exceedingly large literature on the measurement of quality of life, at an early stage of the evaluation there should be a review of this literature to determine the most appropriate measurement instruments for each stage of the evaluation.

It is perhaps worth emphasising that in economic evaluation the term 'cost' does not mean the expenditure of dollars. Rather, costs should include all social, psychological, environmental or financial factors that impinge upon an individual's well-being. While some of the less tangible factors have on occasions been omitted from purported economic studies, this should be seen as a criticism of such a study. Of course, the practical problems associated with measurement are not confined to economic studies.

When costs and benefits are correctly defined, a Program should normally continue only if all costs are less than all benefits. With a finite budget the principle that should be applied is that the net social benefits per dollar of the budget should be maximised. This implies that the priority of Projects would normally be determined by the following ratio:

$$(\text{Total benefit} - \text{total (social) cost})/\text{budget cost}$$

The priority implied by this ratio could, of course, be overridden by issues of distributive justice or by other political considerations. Further, when costs or benefits cannot, with any amount of ingenuity, be quantified they should be noted as 'intangibles' to be assessed qualitatively by decision-makers.

These principles will be incorporated at all levels of the evaluation. Even in those activities which resist quantification most stubbornly, it is important to know the expenditure upon the activity to determine the benefits that have been foregone from other parts of the Strategy which are restrained by a lack of finance.

3.2 Prevention vs Treatment

It is worth emphasising that, with the application of these principles, no preference would be given to preventive or to palliative Programs per se. While the maxim that 'prevention is better than cure' has prima facie appeal, it must be accepted that those who are infected by the HIV virus also require treatment. The benefits received by such persons are no more or less worthy of consideration. Cost Utility Analysis suggests that competing Programs should be compared in terms of their benefits (ideally, Quality Adjusted Life Years gained) per unit of cost. This, of course, implies that prevention would automatically receive higher priority since victims of the HIV epidemic are generally young and preventive measures gain a large number of life years. As noted above, while these principles are accepted, there will be little opportunity for the application of formal Cost Utility Analysis in the assignment.

3.3 On-going Research

In any evaluation exercise a difficulty that arises is that the Project or Program under review will change and evolve over time. New Projects, and many of those in the HIV/AIDS Strategy are new, will be moving down a learning curve both in terms of the nature of the output and the efficiency of the organisation. Assessment and policy conclusions would be invalid if they measured the performance of an infant Project and suggested its cessation.

For this reason each of the summative evaluations will attempt to incorporate three components. The first consideration is what a particular activity could potentially achieve if it were operating to its potential: **efficacy**. The second consideration is the actual performance to date: **effectiveness**. The final consideration is the means for setting in place a system of **on-going monitoring**, for example, through the use of performance indicators.

3.4 Performance Indicators

The development of performance indicators will comprise a substantial part of the quantitative elements in the evaluation and is essential for an ongoing assessment. Indicators are seen as summary statistics that give a measure of the success of any Program or Project in meeting its objectives. Performance Indicators are norm referenced that is to say the statistics are produced so that they can be compared with a specific reference point in order to make a qualitative judgement about the success of an intervention. Normally they are based on interval data, but there are many examples of ordinal, nominal or more descriptive data being scored and used as indicators.

The ideal indicator is a direct measure of a Program's objectives - lives saved, or an index of suffering. Where this outcome is not easily measured, or where it is influenced by factors from outside the Program, an attempt should be made to move back along the causal chain which exists between the intervention and the outcome and to identify an indicator based upon more reliable data and not subject to extraneous influences. Gains in data quality may have to be traded off against considerations of validity.

For example, attitudinal change resulting from media campaigns is very difficult to measure. Where measures show no appreciable difference before and after a campaign, it could be due to failure of the attitudinal measure adequately to discriminate. It may be possible to measure the levels of perception of the advertisement (eg proportions who recall it in various levels of detail). Other methods, eg small group analysis, can be used to establish the responses to the advertisement and their effect on attitude change. In this way an indicator can be developed for an outcome (change in attitude) by measuring a logically prior phenomenon (perception of the advertisement).

In order to be successful indicators must meet a number of criteria, such as validity, reliability, adequate discrimination and so on. There are other criteria that require joint consideration by Program management and delivery staff and the evaluation team, including the significance and appropriateness of the indicators, the selection of suitable norm reference points and the ready availability of data.

Performance indicators are very time bound: a change in Program objectives or management, or even minor changes affecting the collection of data, can require modification or development of new indicators. If the evaluation results only in the establishment of a set of indicators, the benefit could be very short term. If staff are involved in the development of performance indicators, the benefits are more likely to endure over the long term.

A detailed discussion about performance indicators may be found in Section 7.2.

3.5 Data and Data Collection

Evaluation will use centralised data sources of both quantitative and qualitative data, where available, for both Programs and Projects.

Since data will come from a variety of sources, many of which are not scientifically controlled, comparison between data will be undertaken with considerable care. However, a cost effective evaluation requires the minimisation of new data collections and the use of existing information where possible. For this reason the approach to data collection will:

- . identify the specific purpose for which the data would be used in the evaluation;
- . explore existing sources to determine their suitability for the task;
- . undertake new collections where there is no valid and reliable alternative source.

It is unlikely that a large scale nationwide survey will be needed. Instead particular sections may be added on to omnibus surveys conducted for other purposes. Such exercises are often cost ineffective in the sense that they produce large volumes of descriptive data which have no particular relation to the objectives. It is, however, probable that new data collections will be undertaken as preliminary enquiries do not reveal very extensive data sources relevant to our task.

All data will, if possible, be collected in a format compatible with international data collections where these exist.

3.6 Confidentiality

There is a need to preserve strict confidentiality of all sensitive data. Proposals for new collections of such data will be subject to the inspection of an independent ethics committee.

3.7 Consultation

It is a recurring theme of this submission that the proposed methodology requires ongoing consultation at almost every point with individuals, providers and organisations involved in the National HIV/AIDS Strategy. It is, however, worth emphasizing that the starting point of a summative evaluation must be extensive consultation with all relevant bodies to determine the range and nature of the activities undertaken and the social costs and outcomes to be included in the assessment. Formative evaluation is, to an even greater extent, dependent upon the advice, co-operation and participation of those whose activities are to be influenced.

Section 7.1. contains a list of personnel and organisations to be consulted.

3.8 Multidisciplinary Research

The breadth and complexity of the task implied by these factors is greater than the skills incorporated in any single discipline. This implies that an effective evaluation must necessarily be multidisciplinary. Using a multidisciplinary team ensures the availability of a large range of methodological skills. Whether focusing on the National HIV/AIDS Strategy or its component Programs, our approach will be to identify key objectives and practical constraints to evaluation (time, finances and the accessibility of data) and then to select that method which can assess in the most rigorous manner, the achievement of objectives. Thus, a multiplicity of techniques will be available for evaluation of any facet of the Program at a particular stage of development.

We believe the evaluation should be conducted by the National Centre for Public Health and Health Care Evaluation in Monash University and the University of Melbourne. The proposed team has expertise in health economics, epidemiology and sociology. Members of the team have experience in mathematical modelling of the spread of the virus and in the qualitative techniques of case study, documentary analysis and the use of unstructured interview material. The team includes a Health Administrator and members of the Centre for Programme Evaluation at the University of Melbourne which specialises in research and teaching across a wide range of evaluation methods. Using a multi-faceted methodological approach means that more than one method can be applied to a given problem. Comparisons between the results of the different methods of analysis is the best means of ensuring a rigorous evaluation especially in areas which are difficult to assess either because the data are perceived to be 'soft' or the interpretation of the data is somewhat ambiguous.

4 EVALUATION OF SPECIFIC PROGRAMS

Each of the Programs which constitute the Strategy requires a different approach, detailed below. The approach to the Education and Prevention Programs is sufficiently similar for these to be addressed together.

4.1 Education and Prevention

Prevention clearly encompasses Education and it may be difficult to separate the effects of specific components of either Program. Methods appropriate to the evaluation of the Education Program will very often be appropriate also to the Prevention Program. Indeed, it is probably inevitable that methods of data-collection and analysis for these areas will overlap. Although we recognise the distinctions between the Education and Prevention Program areas outlined in the White Paper, these two Program areas are considered together for the purpose of the evaluation.

The overall questions that the evaluation must address in relation to Prevention and Education are:

- Has the Prevention Program worked? Is there evidence of a reduction in the transmission by blood, semen, organ and tissue donation which is attributable to the Program? Is there evidence of the minimisation of personal and social impact?
- Has the Education Program been effective as judged by:
 - knowledge of risks to self and others?
 - attitudes towards people with HIV?
 - attitudes towards people at risk of HIV?
 - changes in practices that put people at risk of HIV?

4.1.1 *The Summative Evaluation of Prevention and Education Programs*

Sources of Data for Evaluation

- . Literature review (social and behavioural, educational, psychological, program development and implementation, epidemiological, etc.).
- . Expert opinion and consultation with workers in the community who have considerable knowledge of their own areas. The evaluation should aim to integrate the knowledge of specific areas, including anecdotal information, into a composite picture of the whole area under study. People to be consulted include researchers, epidemiologists, psychologists,

mathematicians, counsellors, social workers, educators, representatives of ethnic and rural communities, workers in Family Planning Centres and Community Health Centres, youth workers, church workers and so on.

- . Government, Catholic and Independent schools would be consulted and possibly surveyed to examine their policies and the success of their HIV/AIDS education and prevention efforts.
- . Discussion with workers who handle blood and blood products, surgeons, emergency staff, health workers and workers in institutions who are likely to come into contact with infected blood.

The qualitative evaluation will:

- . investigate knowledge, attitudes and practices via small-scale surveys, in-depth interviews and focus groups, particularly in hard-to-define populations such as men in rural areas who engage in male-to-male sex, clients of prostitutes, recreational injecting drug users and homeless youth.

More specifically, the qualitative evaluation will:

- . examine knowledge about HIV in relation to its seriousness, prevalence, modes of transmission, risk behaviour, safe behaviour, when people are at risk, how to protect self and others, dealing with people with HIV and infection control practices (what, how and when);
- . examine attitudes towards people perceived to be at risk, people with HIV, sexual practices, drug use, needle exchange, condoms, laws, testing, discriminatory practices and sex education;
- . examine perceptions of risk;
- . examine practices relating to sexual behaviour, use of condoms, drug use, needle exchange, discrimination and infection control;
- . investigate the provision of information and education, counselling and testing services;
- . investigate areas of possible discrimination practised by surgeons and other health care workers.

The quantitative evaluation may need to develop innovative performance indicators to measure the incidence and prevalence of HIV infection since the accuracy of current measures of sero-conversion can be questioned: testing is largely both voluntary and anonymous, involves an unspecified level of underreporting and an unknown number of people have multiple tests. It

will still be possible to use this data since we can expect little or no change in these practices and the inaccuracies should remain relatively constant over the period of the evaluation (unless there are changes in testing policy or changes in practices are recommended to particular groups).

Family Planning Centres, Action Centres and STD Clinics can provide data on decreases (if any) in the teenage pregnancy rate, incidence of other STDs and Hepatitis B, which are indicators associated with safe sex and clean needle use.

There will be a need for a combination of methods to assess the incidence of new cases, mode of transmission in new cases, incidence of risk behaviour, maintenance of safe sex, maintenance of safer drug using behaviour.

Data sources could include:

- . mandatory testing of blood, sperm, organ and tissue donations, in vitro fertilisation programs, of prisoners on exit from prison, and other data collections such as testing of people with haemophilia;

Large-scale community surveys are more appropriate for measuring perceptions of personal risk and for measuring knowledge of and attitudes towards people with HIV/AIDS, rather than for sensitive, more personal topics. Surveys of secondary school students could also be conducted with the co-operation of State Education Departments (similar to studies conducted by researchers at the University of Melbourne).

Evaluation of National Campaign to reduce Discrimination

Such surveys might measure:

- . percentage of the community with accurate knowledge of the antidiscriminatory policy of the Strategy towards people with HIV/AIDS;
- . percentage of health care workers (including surgeons) with accurate knowledge of the antidiscriminatory policy of the Strategy towards people with HIV/AIDS;
- . percentage of health care workers who know the national criteria for blood transfusion and universal precautions.

HIV Testing (Screening)

Screening has a variety of possible effects:

- . early detection of HIV infection which facilitates treatment;
- . early detection to prevent unintentional transmission of the virus to a partner; and screening for bloodbanks;
- . medico-legal requirements (which are outside the terms of reference here).

As a consequence it is not easily classified under a single Program. The first consequence above may be considered as part of a Treatment Program whereas the second is part of Prevention.

The important issue to be resolved is the target group for screening. At one extreme, the entire population could, in principle, be screened. At the other end of the spectrum, it could be limited to voluntary, high risk groups. Along this spectrum there are varying levels of persuasion and compulsion, and various target groups.

Despite the difficulty in classifying screening, the same principles govern its assessment as govern the assessment of other Projects. Each potential target group in the population defines a different potential Project. The benefits of screening and, in principle, the quality adjusted life years gained through early detection and reduced transmission of the virus must be compared with the social costs of the Project and the alternative uses of the expenditure.

Action Research aimed at Hard-to-reach Groups

Some level of Action Research is recommended as the best way of acquiring knowledge about hard-to-define and hard-to-reach populations such as men engaging in male-to-male sex but not identifying as gay or homosexual, homeless youth, people from NESBs. The following methods are suggested:

- . provide worker to Gay and Married Men's Association (or equivalent group) to conduct research through advertisements in newspapers which carry advertisements for sexual services and contacts;
- . conduct (or commission) study to question men on 'beats' or at homosexual venues (Dr Chapman, Westmead Hospital, who has conducted such a study, has been specifically contacted regarding this proposal);

- . provide resources to tap into major studies (such as the Victorian IV drug users cohort study) already under way (Dr Crofts, Fairfield Hospital, the co-ordinator of the Victorian IV drug users cohort study, has been specifically contacted and has expressed his willingness to co-operate);
- . conduct Action Research amongst street kids or support workers to carry out research;
- . advertise in national magazines for youth;
- . provide workers or money to allow people from NESB to carry out their own research;
- . provide workers or money for Aboriginal populations to carry out their own research.

4.1.2 The Summative Evaluation of Major Projects

Qualitative evaluation will be crucial since changes in knowledge, attitudes and practices will be difficult to attribute to the Education and Prevention Program areas and even more difficult to attribute to individual Projects.

Issues of particular relevance include:

- . the accuracy and sensitivity of media reporting and representations of HIV/AIDS;
- . the effects of differing State legislation;
- . a comparison of HIV/AIDS information produced by various government and non-government organisations;
- . the availability, utilisation and appropriateness of specific services universally and in particular locations.

4.1.3 The Formative Evaluation of the Prevention and Education Programs

Performance indicators will need to be developed in relation to the questions:

- . Have education and prevention services been available universally (in every State and Territory, in rural areas and to specific sub-groups, such as people from different ethnic backgrounds, homeless people etc)?
- . Has information been properly disseminated, comprehensive, understandable, clear and accurate, and specifically comprehensible and sensitive to sub-groups?

- . Are appropriate and manageable evaluation procedures in place or needed?
- . Is there a case for increased mandatory testing in order to increase knowledge of prevalence, mode of transmission, risk behaviour, location and identity of infected individuals?

4.2 Treatment, Care and Counselling

This Program area consumes over 50% of Commonwealth HIV/AIDS expenditure and is therefore an important focus for evaluation. It constitutes a wide spectrum of services including acute hospital-based care for the management of crises, hospital out-patient care and community-based care. These services are delivered to individuals with established or suspected HIV infection and primary care givers.

Where appropriate, a cost-utility approach will be used. Changes in survival and QL (ideally quality adjusted life years saved) at the beginning and end of the triennium will be compared with previous points in time. Where possible, international comparisons will also be made.

The difficulties in collecting the relevant data on Performance Indicators and then attributing to them the differences in the treatment program at the beginning and end of the triennium are considerable.

However, data on survival and at least some measures of quality of life at the two points should be available. Cost data will need to be comprehensive and specifically include services not specifically designated as HIV/AIDS-related as well as those so designated.

One problem is that the quality of life and the cost of services consumed by individuals with AIDS varies considerably. This is not only on an individual basis but also in terms of the particular stage of the disease. Thus studies of QL and cost may need to be made specific to the particular stages of the disease as a basis for making a study of them across the disease in toto (i.e. from first diagnosis to death).

Attribution is difficult for the reasons noted in other Program areas (i.e. change in clinical status may reflect non-treatment factors). In addition, however, the Treatment Program is complex and multidimensional (eg it is hospital and community based; drug based and counselling based). It will be difficult to specify that a treatment Program effect is related to a particular part of that Program.

Existing data sources (eg multicentre drug trials) offer opportunities and limitations in collecting data suitable for evaluation. Additional data collection to supplement existing sources will be inevitable. For this Program area there is a near conjunction between Program and major Project in a number of States since one Specialist Centre - or organisation - delivers the large majority of a

relevant service (eg hospital in-patient). Where this occurs they can be evaluated together with the proviso that alternatives to, and gaps in a major Project are considered across a State.

4.2.1 *Summative Evaluation: Quantitative (Preliminary Performance indicators)*

Parameters set out below will serve as a basis for discussing performance indicators with service providers, relevant populations, government officers and other evaluation experts.

Benefits

This could require the collection of data concerning:

Survival

- . from first infection to death;
- . from onset of symptoms to death.

Quality of life

- . simple indices of disability (activity and mobility) such as the Karnovsky Index (more usually used for individuals with cancer, but applied to AIDS at Fairfield Hospital, Melbourne);
- . The Spitzer QL measure could also be introduced as a Performance Indicator, subject to the agreement of the treatment agencies.

Cost

This is likely to be the marginal cost per capita of providing services for particular groups of individuals at different stages of the disease - including the opportunity costs of the primary care givers and volunteers. These average costs will be computed from:

- . the costs of the type of care received within one designated stage of the disease (eg hospital intensive care, other hospital care and community based care);
- . the length of time during which this type of care was received.

Other data (performance information rather than indicators) could also be collected:

- . perceptions of individuals with HIV/AIDS (and their primary carers): their outlook, grief reaction, emotional and physical symptoms and so on. This data is qualitative in nature and is properly regarded as performance information;

- . patient satisfaction with, and expectation of treatment;
- . the level of knowledge relating to the management of HIV/AIDS, among health workers, particularly those providing care to individuals with the disease;
- . the level and geographic distribution of relevant services such as Specialist Centres (including AZT treatment centres), home nurses and general practitioners with special knowledge of the disease, specialist HIV/AIDS counsellors, services for the treatment of complications of the disease and special disability-related benefits (eg for domiciliary home care and equipment).

Methods of Data Collection for Performance Indicators

Data collection will be based on:

- . Published studies especially multi-centre drug trials; reports of Specialist Research Centres and CARG-funded research.
- . Hospital and home nursing organisation clinical records and unpublished case series.
- . Both can be used to obtain data on survival (bearing in mind lead-time bias), indices of activity to be used in constructing QL measures and on service provision that may be relevant in determining costs of particular types of services.
- . Other organisational records (eg accounting and financial records) can be used in determining costs of particular types of services.
- . Questionnaire (or interview) surveys of stratified samples of individuals with HIV/AIDS and their primary carers for measures of QL and client satisfaction. Intrusion of privacy will be minimised if these are conducted through organisations providing relevant care, or through relevant community groups.
- . Questionnaire surveys of health workers (including volunteers). These can be used to assess education and training about the management of HIV/AIDS. Workers would be contacted under the auspices of relevant organisations (Royal Colleges, Royal Australian Nursing Federation etc.).
- . Government reports and working documents, especially reports relating to evaluations of HIV/AIDS services.

4.2.2 Summative and Formative Evaluation: Qualitative

These will be conducted simultaneously and be based on a review of the relevant literature and consultations with the following groups and individuals:

- . specialist AIDS physicians, other relevant health care workers such as general practitioners and home nurses providing treatment, and volunteers. Health workers with little direct contact with treating the disease will specifically be consulted;
- . organisations representing relevant population groups.

These consultations may occur within State Forums and with particular individuals and groups. Key issues proposed as a starting point for consultations about both existing services and desirable changes include:

- . difficulties in diagnosis;
- . Centralisation/ Decentralisation;
- . specialised/ generalised basis of services - particularly AZT therapy - specialisation or mainstreaming;
- . availability (current and projected) and quality of comprehensive services eg counselling, palliative services;
- . attitude of Specialist treatment Centres as well as other general hospitals and treatment services as perceived by individuals with HIV/AIDS;
- . intensity and length of treatment, particularly hospital treatment (too much, not enough);
- . cheaper institutional care options (eg home, hospice, nursing home);
- . difficulties in the interface between hospital- and community-based care;
- . the availability and quality of education and training programs for the management of AIDS:
 - complications of AIDS, including the provision of intensive care services;
 - associated illnesses and risk-factors;
 - management of 'normal' diseases in the presence of AIDS, including surgical treatment;
 - counselling, emotional support and empathy;
 - avoidance of needlestick injury;

- 'burn-out' of full-time specialist staff;
- relationship of treatment centres with organisations representing HIV risk-factor groups (particularly in regard to the advisory role of the latter in the provision of services).

4.3 Access and Participation

Although there is overlap with other Program areas, eg education, the National HIV/AIDS Strategy sees the issues of access and participation as sufficiently important to be the focus of a separate Program area. The major area of concern is discrimination against people who are HIV seropositive or people from groups who may be imputed to carry the virus. Relevant areas are employment, financial support, housing and insurance (including superannuation). Regardless of whether the discrimination is based on legal or social norms, the matter is of key importance to the success of the HIV/AIDS Strategy.

There are two arguments in favour of maintaining access and participation for anyone infected with HIV. The first is ethical and argues that individuals with HIV/AIDS should not be the targets of discrimination on the grounds of their infection. The second argument is a pragmatic one, namely that non-discrimination will encourage participation in Prevention and Education Programs.

4.3.1 Key Issues

- . How adequate is the definition of Access and Equity in the Commonwealth's Strategy?
- . Is there discrimination against people who are HIV seropositive, in the key areas identified in the Strategy, or in other areas?
- . How effectively are the rights of people who are HIV seropositive balanced against their obligations to prevent the spread of the disease?
- . Are there any important groups suffering discrimination which are not identified in the Strategy, or any types of discrimination not adequately covered (eg persons of non-English speaking background, and Aborigines and Torres Strait Islanders)?
- . How are the human rights aspects of testing procedures handled, especially with respect to informed consent, the testing of prisoners and applicants for migration to Australia?
- . What are the implications of the confidentiality of test results, and the need for large scale anonymous testing for the monitoring of the disease?
- . What legal constraints impact on the HIV/AIDS Program?

4.3.2 *Summative Evaluation: Quantitative (Preliminary Performance Indicators)*

Quantitative methods are relevant for assessing attitudinal change towards specific groups, particularly those seen as being at high risk. Opinion poll data will be reviewed or commissioned at several points during the evaluation to assess attitudinal change, especially public perception of groups as being at high risk of infection. Questions could be commissioned with omnibus or telephone surveys by commercial marketing or survey organisations. These would be done on a nationwide basis using pre-coded questions. It would be very difficult, however, to assign causes to any attitude changes detected.

Further opportunities for quantitative analysis may lie in assessing the penetration of particularly difficult-to-reach groups by HIV/AIDS educational material. Although this overlaps extensively with the evaluation of the Educational Programs, the matter of access and equity is significant here. Studies have been made of the penetration of specific language groups by HIV/AIDS advertising² and these could make a starting point for a comparative statistical study.

² Rigby, Ken and Rosenthal, Doreen, An evaluation of the national campaign against AIDS for people of non-English speaking background, unpublished report for Department of Community Services and Health, February, 1990.

Taking the preceding point more broadly, if the term 'access' is taken as 'access to specific HIV/AIDS Programs or Projects', other effective indicators become available. For example, it may be possible to measure the impact of any Program by expressing the number of clients it reaches as a proportion of the total target population. Where the target population is not known, it may be imputed, and the proportions, although unlikely to be accurate in the absolute sense, may be legitimately used for comparison between Programs. Alternatively, rather than taking the target population as a base and analysing it by Program impact, it is possible to take as the base the number of persons affected by the Program and analyze these to describe the profile of Program participants. Either way, taking a broad definition of access will permit the application of satisfactory quantitative techniques.

4.3.3 *Summative Evaluation: Qualitative*

Information sources will include sectors of public administration charged with the monitoring of complaints, for instance the various ombudsmen's offices, and relevant units charged with dealing with complaints within more specific administrative structures, for instance State Health Departments. Community groups, especially those likely to suffer discrimination or disadvantage, will provide valuable sources of information, and appropriate consultation will be sought with such organizations. The frequency of complaints at both levels will be sought.

Given the importance of legal sanctions in this area, and the primacy of the State responsibilities, attention will be paid to the contribution made by State legislation and the role of cooperation between State and Federal bodies.

In general, the equity implications of a range of policy initiatives will be sought in regard to instances of inequity, discrimination or unfairness that indicate some substantial flaw in policy or its implementation. Conversely, aspects of Programs that significantly enhance fairness, or are seen to discourage discrimination, will also be highlighted.

This aspect of the evaluation will focus on the broad Program area at both Federal and State levels. Unless specific and narrowly based problems are found to have major equity or discrimination implications, it is seen as unlikely that there will be any significant emphasis on specific Projects. An exception may exist in the need to assess the role of general practitioners.

As the first point of contact for many people infected with HIV will be a general practitioner, the information available to, and the practices of, this group of care givers is of particular relevance for access and for education. The evaluation will include both a summative and a formative assessment of the knowledge and practices of GPs and other front line care givers.

4.3.4 *Formative Evaluation*

In this Program area, a summative evaluation may imply the need for reform. If there are found to be major shortcomings in access, for instance, or significant areas of discrimination are uncovered, specific amendments to the Strategy and its implementation will be suggested.

4.4 Research

The research Program, covering both social and biomedical aspects, is to be evaluated in terms of its capacity to:

- . generate data necessary for directing the Strategy;
- . contribute to the development of the necessary vaccines, tests, therapies and other educational and preventive strategies for managing the infection.

4.4.1 The Summative Evaluation of the Research Program

Sources of Data for Evaluation

- . Consultation with the research and professional communities such as the AIDS Special Research Centres, NHMRC, CARG, ANCA and IGCA committees as well as leading academics.
- . Consultation with professional and government bodies and groups with high risk of HIV infection.
- . Review of the literature including professional journals, government publications and the international literature on both HIV/AIDS and evaluation, including WHO publications.

The qualitative evaluation will:

- . assess whether the goals of the Research Program area as stated in the White Paper are appropriate for the Australian context;
- . assess the balance between and the contribution to the Research Program of various research areas: biomedical research, epidemiology and the social and behavioural sciences;
- . in comparison with the international effort, determine whether there are omissions, gaps or unnecessary duplication;
- . assess whether the level of funding for research is appropriate. Is there a need for more funding or for more selectivity in the research funded?

- . assess the level of dissemination of results to those groups who require them for policy formulation or implementation;

The quantitative evaluation (preliminary performance indicators) will assess:

- . the contribution to the development of tests, drug treatment and vaccines;
- . the contribution to biomedical and epidemiological knowledge of the epidemic;
- . the generation of surveillance data on the levels of HIV and fully established AIDS;
- . the development of computer data-bases of all results for use in evaluation, and for contributing to multinational data-bases.

Realistic and appropriate performance indicators will be developed for each activity in cooperation with the specific research community and with those independent experts whose role it is to review the research of the Special Research Centres. Performance indicators may include measures of productivity such as scientific citations and numbers of publications in selected journals.

4.4.2 *The Summative Evaluation of Major Projects*

A qualitative evaluation will also be conducted at the level of major Projects including the Special Research Units. For major Projects information on the following issues will be sought from the researchers themselves.

4.4.2.1 Key Issues Concerning the Program and Major Project

- . Which procedures for the evaluation of research activity are in place at present, how useful and effective are they?
- . What is the composition and effectiveness of Project and Scientific Advisory Committees to Special Centres?
- . Which Peer Review processes are being used, both in funding particular research Projects and in reviewing the Special Centres?
- . What is the balance between pure and applied research; between professional and community-based research? Are there impediments to research in any significant area?
- . Have important areas been omitted?

- . Which grant-giving procedure works best, specific purpose grants (eg CARG) or general purpose grants (eg NHMRC)? Should these bodies adopt similar or different procedures? Should Projects be commissioned, based on national needs? What should be the composition of these grant-giving bodies?
- . Centralisation vs De-centralisation. In which areas is it better to centralise research in Special Centres and where it is best divided between smaller teams and individual researchers?
- . What is the availability of research personnel, both at senior and junior levels? Are there adequate training procedures, career structures; what impediments exist?
- . What is the availability of evaluation personnel? What can be done to support or foster these?
- . Is the research of use to the international community or is its primary function to maintain and support the skills and knowledge of the Australian research and medical community?

4.4.3 *The Formative Evaluation of the Research Program Area*

This aspect of the evaluation will focus on the institutional structure of the Research Program area and the contribution which these have made to the achievements of Australian HIV/AIDS research. The evaluation will examine:

- . the committees which administer funding;
- . the criteria under which funding is allocated;
- . the balance between the various research efforts.

4.5 International Co-operation

The objective of international cooperation is to initiate and to participate in appropriate international HIV/AIDS Programs, recognising the global impact of HIV/AIDS and the need for concerted action.

The leading role in multilateral action against HIV/AIDS has been taken by the World Health Organization, which, in February, 1987, set up the Global Program on AIDS (GPA) which is the architect of the Global AIDS Strategy (GAS).

The question may be asked then to what extent and in what way does Australia contribute to the GPA? Is the Australian HIV/AIDS Strategy in conformity with the WHO guidelines on HIV/AIDS and if not, due to what local circumstances?

4.5.1 Regional Action

The White Paper stresses the requirements for cooperative ventures in the Pacific and East-Asian region within which Australia plays a leading role in the development of HIV/AIDS programs. Although there are few cases reported in the immediate region, the virus is considered likely to spread through tourism, alcohol and drug abuse and insecure blood supplies.

4.5.2 Key Issues

- . What contribution have Australian medical and social policies made to the development of the global HIV/AIDS Programs?
- . How far does the Australian HIV/AIDS Strategy accord with the prescriptions made by the GPA with regard to planning and implementation, data collection and processing, and evaluation?
- . Is the implementation of specific Projects in the Australian context in accordance with WHO GPA guidelines.
- . Has Australia met its responsibilities to its immediate neighbours in the East and South East Asian Region.

4.5.3 Summative Evaluation of Program.

The evaluation will:

- . summarize the achievements and shortcomings of the international aspects of the Programs;
- . assess the relative costs and benefits of Australia's international involvement;
- . identify gaps or omissions;
- . identify areas of overlap. Avoiding unnecessary duplication may allow for considerable savings in resources.

The main thrust of the evaluation of this Program area will be summative and qualitative. Data sources used will include publications and lists of personnel, administrative information, and possibly small and specifically targeted surveys of participants in specific international activities. These sources will be supplemented by more narrative, qualitative approaches: extensive interviews with key personnel, reviews of literature and consultations with key figures in the international HIV/AIDS Programs, in particular GPA personnel.

It is difficult to specify appropriate quantitative performance indicators in advance. There are questions about the availability of data, and the validity and reliability of measures. In consultation with appropriate HIV/AIDS personnel indicators might be developed by examining the following areas:

- . the extent of comparability of Australian monitoring processes with those specified by GPA;
- . the extent and nature of Australian participation in appropriate international conferences;
- . the extent and nature of Australian collaboration with regional Programs, eg training programs for medical staff.

4.5.4 Summative Evaluation of Major Projects

The major focus will be on the Program but particular Projects within the international cooperative Program may be of special interest either because of their substantive significance or because they consume extensive resources. In these instances, it may be desirable for the evaluation to range more broadly and suggest modifications and improvements if appropriate. In these instances, the evaluation may be more formative than summative.

5 IMPLEMENTATION AND ORGANISATION OF EVALUATION

5.1 Proposed organisation

The framework for the evaluation that is outlined above is multidisciplinary in nature; indeed given the nature of the epidemic and the national strategy developed to meet it, it could not be otherwise.

It is proposed therefore that the evaluation be managed by one group with the appropriate multidisciplinary skills. We would propose that this group, the PHRDC National Centre for Public Health and Health Care Evaluation is the most appropriate group to undertake this role. The skills of its directors and staff encompass the relevant disciplines of health economics, epidemiology, public health evaluation and the social sciences. As a National Centre, it is well able to implement such an ambitious evaluation successfully.

It will nevertheless be desirable for the central management group to disperse funds to commission specific, selected population surveys of a specialised nature (eg for IV drug users) and to commission specific expert reviews, again of a specialised nature (eg the range of, and characteristics of national strategies adopted in other countries). Details of such surveys and reviews are set out below.

The group managing the evaluation should have both a central core, as well as a dispersed presence in the various States and Territories. It is proposed however that the majority of funds for staff be vested in a central core to ensure its viability and that staff in the States and Territories operate on contract/ part-time/ as required basis in conjunction with central staff who would travel at regular intervals to the States/ Territories to support them. (New South Wales, as the State with the largest number of individuals with AIDS/HIV may require a more long- term arrangement in terms of the dispersed staff).

Staff of the management group will need to assume specific responsibilities:

- . for particular levels of the strategy - the National Strategy, Major Program areas, selected Major Projects;
- . for particular program areas;
- . for particular States.

It is proposed that the management group should consist of:

- . a Director (Senior Research Fellow);
- . two Research Fellows;
- . an Administrative assistant/Word Processing Clerk;
- . one Research Fellow (i.e. Full-time equivalent dispersed between States and Territories).

This group will need to play four main roles:

- . to provide program support and liaise with individuals/groups engaged in commissioned work;
- . to receive and interpret all available data sources (published articles, government reports and working papers, commissioned reviews, etc);
- . to consult widely with relevant professional, government, community and special interest groups;

and as a result

- . integrate and interpret all these sources of data to form and produce an evaluation that is both summative and, to a lesser extent, formative in nature.

5.2 Staging

It is strongly recommended that the budget for the Study be restaged so that the monies are distributed more evenly between the three years. Indeed they should be distributed in descending rank order of amount, to the first year, then third year and last second year. This is because empirical commissioned surveys are likely to be conducted in the first year (baseline) and the last year (outcome of triennium's program). In addition consultations and commissioned reviews will necessarily be largely though not entirely concentrated in the first year (preferably in the first three to six months) to acquire the necessary inputs for an informed evaluation and one that is well accepted by the various interest groups. (The main outline of the Budget is set out below.)

Main Stages of Evaluation:

Year 1

Appointment of Director and Staff of Evaluation group.

Consultation with interested parties

- . State Forums (and Evaluation Workshops);
- . Key Individuals;
- . Important groups, agencies and organisations (International consultation by mail/phone/fax etc. only at this stage).

Commissioning

Studies (selected) with necessary program support/liaison:

- . Base-line study - see below;
- . Expert Reviews (selected) - see below.

Clearing House, and Data Repository

Establish, synthesise and interpret:

- . Published and unpublished studies, conference papers;
- . Intergovernmental and governmental (eg ANCA and IGCA) working papers, institutional data sources (eg hospital morbidity);
- . Submissions from community groups.

Interim Report - Interim Evaluation - Year I

Year 2

Consultation- ongoing/new developments

- . Overseas inspection tour (GPA, Geneva, US, Africa etc.).

Commissioning

- . Studies - analyse base line/ prepare follow-up study for Year 3;
- . Expert Review - review and synthesise.

Clearing House - ongoing/new developments

- . Review and synthesise

Interim Report and ?Interim Evaluation - Year II

Year 3

Consultation - ongoing/new developments

- . Second round State Forums.

Commissioning

- . Follow up studies/ analysis and comparison with baseline.

Clearing House - ongoing/new developments

Evaluation - Formal Synthesis, interpretation, decisions, recommendations, write-up and production of final Report.

5.3 Selected Commissioned Studies

These have been outlined in the particular program areas above they will be drawn from:-

Socio-behavioural Studies

- . These could include Client Satisfaction, Quality of life;
- . Mail Questionnaires of relevant health care workers.

Health Economic Studies

- . Detailed costings, particularly for hospital treatment.

Academic Studies of Particular Groups

- . eg. School aged adolescents (eg school-based/ Youth- magazine based);
- . Intravenous drug users;
- . Beats and Venues (Men who have sex with men).

Community Knowledge and Attitude

This could be by contact with the ABS, market research company (in an omnibus or telephone survey) or university-based group such as the Social Survey Data Archive at the Australian National University.

Community based studies of particular groups

eg. Action Resource centre, Gay and Married Men's Association, (GAMMA) Aboriginal organisations, Youth Organisations, Non English Speaking Background organisations.

5.4 Selected Expert Reviews

These could include :

- . epidemiological characteristics of the epidemic internationally as well as in Australia;
- . range and characteristics of national strategies in different countries;
- . literature reviews of the 'state of the art' in particular areas eg treatment, prevention and research;
- . commentaries on particular program areas in the National Strategy.

6 INDICATIVE BUDGET

It is proposed that an indicative Budget (of \$1.6 million, 1990 rates) might be distributed as follows:

Year I	635,000
Year II	410,000
Year III	<u>605,000</u>
	1,650,000

	Year I	Year II	Year III
Staff (Central and dispersed inc. oncosts and superannuation).	255,000.00	255,000.00	255,000.00
Consultation, State Forums	80,000.00	30,000.00	80,000.00
Commissioned Studies/Reviews	200,000.00	50,000.00	170,000.00
Office Supplies (establishment and operating)	20,000.00	20,000.00	20,000.00
Institutional Support	80,000.00	80,000.00	80,000.00
TOTAL	635,000.00	470,000.00	605,000.00

7 ANNEX

7.1 Personnel and Organisations to be consulted

Members of Parliamentary Liaison Group
Royal Australian College of General Practitioners
World Health Organization
National AIDS Counsellors' Association
Haemophilia Foundation of Australia
Human Rights Commissioner
Members of Intergovernmental Committee on AIDS
Partners living with AIDS
Australian Teachers' Union
Ethnic Affairs Commission
NHMRC Special Unit
Office of Multicultural Affairs Commonwealth Department of the Prime Minister and Cabinet
National Catholic Education Centre
Ansell International

In Each State and Territory

Positive Women's Groups
Health Ministers
Directors of Blood Banks
Pharmacy Associations
AIDS Information Committee in Universities and Colleges of Advanced Education
AIDS Carers
Aboriginal Health Units In Health Departments
Representatives of Aboriginal Communities
Representatives of Ethnic Communities
Practitioners in Drug Treatment Agencies
Department of Corrections Health Services or Equivalent
Women's Health Organisations
AIDS Treatment Centres
Medical Director Family Planning Associations
Ethnic Health Policy Analysis
Education Departments
Prostitutes Collectives and Sex Industry Organisations
Private Venereologists
AIDS Program Manager from Health Departments

Representatives of People Living with AIDS
Manager Unit for STDs and AIDS Health Departments
Ethnic Health Workers
Action Centre or Equivalent
STD Counsellors
District Health Councils
STD Centres/Clinics
Health Educators
Representatives of Australian Federation of AIDS Organisations
Venereology Societies
Counsellors and Clinicians treating HIV patients in major hospitals
Gay and Married Men's Associations
Chief Medical Advisor Unit for STDs and AIDS Health Departments
Health Promotion Units in Health Departments
VIVAIDS or equivalent
IV drug users organisations
HIV/AIDS researchers
Health Reporters for major newspapers and journals
Student Representative Organisations
Country Womens Association Members
Rotary Group Members
Country Party Group Members
Prison Officers' Associations
Anti-Discrimination Boards
Royal Australian Nursing Federation
Police Departments

Other Key Informants

Dr. Dennis Altman
Mr. Hans Heilpern
Professor Roger Short
Dr. Alex Wodak
Ms Elizabeth Reid
Justice Michael Kirby
Dr. David Cooper
Dr. Michael Ross
Dr. Julian Gold
Dr. John Kaldor
Dr Robert Douglass
Dr. Ron Thomas

Dr. Anne Mych
Prof. Bob Connell
Prof Beverley Raphael
Prof. Peter Karmel
Mr Adam Carr
Mr Phil Caswell

Historical Figures

Professor Ron Penny
Ms Ita Buttrose
Professor David Penington
Dr. Jonathan Mann
Dr. Neal Blewett

7.2 Performance Indicators

Neil Day, Centre for Programme Evaluation, University of Melbourne.

7.2.1 Introduction

The White Paper suggests the use of performance indicators in Chapter 12 on Performance Measurement and Accountability³. The indicators are prefaced by the remark that they result from the development of program budgeting in the 1987-88 financial year. Recipients of grants are asked to specify how a Project will be evaluated before funding is made available. Moreover the high salience of performance indicators also results from the emphasis on data collection in the Global Program on AIDS.

A number of issues essential to the development of effective performance indicators are considered in this section.

Performance Indicators measure the achievements of a program against the yardstick of its objectives. Thus, before indicators can be developed, objectives must be clearly and simply articulated.

³ AGES, *National HIV/AIDS Strategy*, Canberra, August, 1989.

7.2.2 Objectives

Often formulations of objectives are in broad and general terms: sometimes referred to as 'mission statements'. At other times, objectives may be couched in very specific terms outlining desired outcomes in detail. Some objectives are aspirational, setting long term goals that are worthy, but highly unlikely to be achieved solely as the result of the program. Others are pragmatic, establishing outcomes that are less grand, but more possible in the short term and more likely to result directly from the interventions made as part of the program.

Another dimension on which objectives vary is in the substantive area of policy to which they refer. Many objectives deal with the process of the programs, rather than their outcomes.

For the purposes of developing Performance Indicators, objectives should ideally be:

- specific, rather than general,
- pragmatic, rather than aspirational, and
- relate to outcomes of programs in preference to processes, at least in the longer term.

7.2.3 Performance Indicators

The term 'indicator' usually refers to a simple combination of two or three variables constructed so as to enhance and simplify measurement of a particular factor. More elaborate constructions using more variables (scales, component scores etc.) are also commonly used.

The term 'performance indicator' appears in the literature to refer to a type of statistical indicator with the following characteristics:

- has a norm or reference point (other time, place or program);
- used to make qualitative judgement;
- used for resource allocation, decision-making or accountability.

Also, the term 'Performance Indicator' is occasionally applied to more complex measures (scales, clusters, component scores etc) based on elaborate statistical manipulations of many variables.

Performance indicators are sometimes analysed according to the audience or purpose for which they are intended. For instance, the term 'management indicator' is occasionally used to distinguish performance indicators measuring system level objectives and used to make major resource allocation decisions.

The term 'Performance Information' has been used to emphasize that certain objectives cannot effectively be measured by Performance Indicators. In some instances the nature of the

objective implies a more narrative and descriptive measure. For example, quality of life of HIV sufferers may be considered a qualitative matter better described narratively than measured.

Occasionally subjective assessments like these are scored and used like statistics, but the essentially subjective and qualitative nature of the judgement persists.

7.2.5 7.2.4 Performance Indicators: Criteria for Development and Selection

Statistical Considerations

Validity

Are the measures used biased in any way, eg consistently understating the number HIV seropositives for instance? Do the indicators measure the attainment of program objectives and not other matters?

This question is more complex than appears at first. It is evident that some proposed Indicators measure outcomes well outside the scope of a particular program. Increases in the use of safe sexual and drug taking practices, for instance, might be the result of a particularly harrowing episode in a popular television serial rather than a consequence of an expensive advertising campaign.

Performance Indicators may measure outcomes from the program, but not those that are part of the objectives. For example, some researchers attribute falls in the general levels of sexually transmitted disease to aspects of the HIV/AIDS strategy, although this is an unintended outcome of the programs.

The GPA stresses the need for a few simple indicators but a single Indicator may measure only a part of an objective of the program. The validity of a group of indicators may need to be checked to ensure they provide a comprehensive coverage of all aspects of the objective.

Reliability

Reliability relates to the precision or reproducibility of the indicator. Will the results vary randomly as the result of error (eg test errors or carelessness or inconsistency in reporting and recording of incidents on which indicators are based)?

Discrimination

Is the indicator sufficiently sensitive to changes in the outcomes of the program to register these as changes in the measure? Conversely, will minor and insignificant changes in outcomes result in disproportionately massive shifts in the value of the indicator?

7.2.6 *Performance Indicators: Substantive Considerations*

Significance

Does the indicator measure some aspect of the objectives that is of substantial significance?

Appropriateness

Is the measure appropriate to the context of the program and its objectives?

7.2.7 *Norm Reference Considerations*

Given that Performance Indicators imply comparisons, it is important to be clear what suitable norms or standards are available for comparisons.

The type of norm needs to be specified: will the indicator be compared with figures from

- . other times: (previous years? the preceding month?)
- . other places: (similar program in other States? GPA data from other countries?)
- . other programs.

In considering all these aspects, the major concern is comparability. Are the measures based on data that is truly comparable? Many major problems in this area are hidden by ambiguities in language, units of analysis and time spans may be different, data collection methods can vary in ways that impact severely on the figures. These matters have to be checked very carefully to ensure comparisons are valid and that differences in the indicators are solely due to variations in the extent to which objectives are achieved.

These matters are of particular concern where international comparisons are involved. GPA is putting a heavy emphasis on the quality and comparability of data.

7.2.8 *Reporting Considerations*

Matters such as form of presentation, ease of interpretation and timeliness of data should influence the choice of performance indicators.

7.2.9 *Other Considerations*

Non-manipulable Indicators

In devising indicators and the data collection systems on which they are based, it is desirable to leave as little discretionary power as possible. This not only reduces the very rare incidents of conscious 'cheating', but, more importantly, avoids the understandable tendency for people to present facts in the best possible light when faced with a choice in the means of presentation.

Objective Displacement

Although this matter relates closely to validity, discussed above, it is not uncommon for organizations to place a higher value on those elements of objectives that are measured by performance indicators than those that are not. This is very much the case in instances where the indicator focuses on the quantity of a product, where the real concern is to produce high quality (eg increase in life expectancy for people who are HIV seropositive may be easier to measure, but arguably less significant than quality of life. However, if duration is the only measure used, it is easy to appreciate that resources may be diverted towards longevity and away from quality of life.)

Cost

A key consideration is the ready availability of data in a form suitable for use in Performance Indicators. There are usually high costs associated with collecting data and processing it into a form suitable for analysis. Wherever possible, data should be collected for several purposes, so that the most cost effective means are available to ensure high quality. This may imply careful selection of indicators, in terms of data that is already available. But it may also suggest changes to procedures used for collection of data for other purposes.

On the other hand, given the importance of performance evaluation, the costs of NOT collecting appropriate data in terms of lost efficiency and effectiveness must also be taken into account. This is particularly the case in an information deficient area such as AIDS.

7.2.10 General Issues Affecting Indicators in HIV-AIDS

There are some more general considerations that should inform the development of performance indicators.

Simplicity

Performance Indicators provide simple and easily interpretable statistics that show how effectively programs are achieving their objectives, and permit comparative assessment of performance. Indicators lend themselves to presentation in easily understood graphical form, and give an accessible insight into the performance of complex programs.

On the other hand, the simplicity of Performance Indicators can be deceptive may be bought at the expense of oversimplification of the achievements of complex programs.

This suggests caution in the development and use of indicators, and special attention to the circumstances in which they are appropriate.

In particular, more qualitative outcomes from programs, such as quality of life and sexual behaviour of individuals, may require other methods either as well as, or instead of, usual statistical indicators.

There is a great danger that an over-emphasis on simple performance indicators will lead to the devaluing of many important objectives of HIV/AIDS programs. An unvaried and remorseless reliance on Performance Indicators could lead to the displacement of the overall objectives of the policy to only those concerns that can be adequately summarized in a glossy bar-chart.

Use of more elaborate statistics in the development of indicators, eg component scales, while increasing the reliability lead to difficulties in interpretation by the less numerate.

7.2.11 Performance Indicators and Performance Information

From the previous point, there is a danger that an over-reliance on simple performance indicators could lead to serious misrepresentation of the achievements of many Projects. The evaluation will consider the concept of 'Performance Information' which allows the supplementation of simple indicators with other forms of assessment (for example narrative reports, case studies) that would ensure that performance appraisal was based on consideration of a variety of information collected through a range of appropriate methods.

It is acknowledged that Performance Indicators are, by their very nature, crude and insensitive indications of program performance. They do serve useful purposes however, for example indicating areas of apparent decline and shortcoming. To obtain deep understanding of Programs and to ensure effective evaluation, Performance Indicators need to be supplemented by a range of other methods of analysis.

7.2.12 Comparability

The use of Performance Indicators always involves reference to some benchmark: for instance another time (last year, last month) or another place (other State, other country) or another Program (eg community group activity vs counselling Programs).

In order for the comparison to be valid, the data and its use to construct indicators must be truly comparable. It is probable that this uniformity can be established between administrations, Programs and Projects only in the longer term, and as a result of some jurisdictions altering significant aspects of the way they record and process data.

The danger is that misinterpretations of this type could be hidden by the fact that many differences are transparent. They are disguised by the use of identical terms, with differing meanings, or hidden by seemingly identical procedures which contain minor differences that have maximal impact on particular indicators.

Careful development of indicators and detailed monitoring in future years would be required to ensure true comparability.

7.2.13 Nomenclature

A specific and pernicious example of the preceding point is nomenclature. The terminology used in corrections management is sufficiently uniform for most purposes. However, given the level of precision required in the specification of performance indicators, nomenclature causes many problems. Unfortunately, it appears likely that some of these differences in usage are transparent, and may only emerge as a result of detailed analysis of the indicators themselves.

An example is the term 'Program'. In some areas this refers to an activity at a high level of generality: eg 'Education Program' in Chapter 3 of the White paper⁴. In other applications the term applies to specific interventions: eg 'safe sex' Programs taking place within the broader 'Education Program'. Still more confusingly, specific components within the 'Education Program' may appear eponymous: eg 'Condom Awareness Program'. It well maybe that the promiscuousness of the term indicates it is nearing the end of its useful life, though while it atrophies it can make the task of developing uniform performance indicators particularly difficult.

7.2.14 Implications for Data Collection

⁴ *National HIV/AIDS Strategy: a policy information paper*, AGES, Australia, 1989: the "White Paper" p.26.

In order to develop truly comparable indicators, it may be necessary for some administrations to modify existing data collection and analysis procedures. Although the matter is too detailed to be explored in depth here, it is apparent that, in some areas, data could not be used in comparative indicators without some such change.

7.2.15 Procedures for the Development of Performance Indicators

The development of indicators is a complex matter. In order adequately to canvass the areas outlined above it is necessary to combine considerations of statistical rigour with the requirements of management at system level and the needs of Project delivery personnel. The matter of the dual structure of some management activities is also significant as the needs of regional and central management may need to be reconciled.

The varied uses of performance indicators needs to be carefully considered: are indicators required for monitoring and refining Programs, making resource allocation decisions or other purposes?

Another matter of concern is the level of the Programs under consideration: indicators for specific Projects, while too numerous for detailed review within the evaluation, must be collected in a form that allows for aggregation across a range of Projects within any particular Program area.

Methodological matters are of key concern, especially the appropriateness and adequacy of qualitative measures. Although most indicators are of a quantitative nature, consideration should be given to the development of rigorous performance information measures in the qualitative area.

The foregoing suggests that the process of the development of indicators should be an iterative one, with draft measures being continually refined. It further suggests that there be extensive consultation during the process of development.

Finally, an important product of the evaluation should be the development of a cadre of workers at all levels with skills in the development and analysis of performance measures. Only in this way can the evaluation have any long term impact on performance measurement.

7.2.16 References

WHO, *Development of indicators for monitoring progress towards health for all by the year 2000*, Geneva, 1981, WHO Publications

Working Group on Health Promotion Indicators, *Report on WHO Consultation*, Edinburgh, 1986, (unpublished paper).

Windsor, R.A *et al.*, *Evaluation of Health Promotion and Education Programs*, Mayfield, Palo Alto, California, 1984.

Elinson, Jack & Siegmann, Athilia E. (eds.), *Socio-medical Health Indicators*, Baywood, New York, 1979.

7.3 Evaluation of National AIDS/HIV Strategy: The Role of Mathematical/Statistical Modelling

Dr John Carlin, Department of Community Medicine, University of Melbourne. Recipient of CARG Grant with Drs Becker and Watson (La Trobe University).

7.3.1 AIDS Modelling: Background

There is now a large and growing research literature on modelling the AIDS epidemic, although relatively little has so far been done in Australia. The main motivation of most of the work is to make predictions about the future incidence of AIDS cases, and related quantities, such as AIDS-related mortality. These predictions are hotly sought after by public health authorities for planning purposes, and they are obviously of broad general interest. The prediction task is unusually difficult with AIDS because of the fact that the epidemic process is only partially observable, in the sense that individuals may remain infected (and infectious) with HIV for many years without knowing of their infection.

The terms of reference for much work in this area were set by a major report produced in early 1988 by a working group of prominent statisticians and epidemiologists in Britain (the 'Cox Report' [1]). The difficulty in making accurate predictions was highlighted by a follow-up report [2] fifteen months later in which the 'best estimates' for future incidence were approximately halved. In Australia, a preliminary report was published by Solomon, Wilson and Doust from the National Centre for Epidemiology and Population Health in mid-1989 [3]. This group is about to publish a follow-up report, while the local group of Becker, Carlin and Watson (headquartered at the La Trobe Statistics Department) began work under a CARG grant in early 1990 and will have a first paper ready shortly.

Methods used to predict future trends in the incidence of AIDS tend to be classified into three types. These methods are similar in the sense that they all fit some function of calendar time to AIDS incidence data, but differ in the degree to which attempts are made to incorporate into the model knowledge about the mechanisms that generate the data. At one extreme, a simple exponential, polynomial or other mathematically convenient curve is used and no attempt is made to use knowledge about the infectious nature of AIDS or the incubation period between infection with HIV and contraction of AIDS. Predictions are made by extrapolating the fitted

curve into the future. An early effort of this kind in Australia was made by Whyte et al [4]. It is well known that such extrapolations must be treated with extreme care. One cannot estimate the prevalence of HIV infection with this approach.

At the other extreme are transmission models that attempt to describe the exact nature of the spread of the disease (eg [5]). Construction of these models requires a great deal of information and/or assumptions about sizes of risk groups, types of sexual and needle-sharing practices, etc.

Apart from the work of Anderson and colleagues, few serious attempts have been made in this area, because of the complicated nature of the models and the quantity of information (much of which is simply unavailable) required in formulating them.

The method of back-projection fits in between these extremes and has attracted most attention from statistical modellers. The basic aim is to use the known incidence of AIDS cases up to the present time to estimate (back-project) an incidence or intensity of HIV infections in the past. The basic method incorporates only elementary assumptions about the way the data are generated, basically in the form of a probability distribution for the time from infection to clinical AIDS. The model is sufficiently detailed to allow estimation of the unobserved HIV infection incidence curve up to the present, as well as forming a basis for projections of the future AIDS incidence. In doing the latter in principle one needs to make assumptions about the future course of HIV infections, about which the modelling provides no information, but this is not a great problem for short-term prediction, since most AIDS cases in the near future will arise from infections occurring some time ago.

The method can be enhanced in various ways, for instance by allowing the incubation time distribution to vary in time, and by incorporating other pieces of information, for example knowledge of HIV incidence and prevalence that may be obtained from the results of the voluntary antibody testing Programs. (The latter is fairly limited, although efforts are being made to improve it.)

Despite all the methodological developments (including those of Becker et al), it remains in practice difficult to estimate past incidence of HIV infection precisely, because of the long and variable induction period between infection and affliction with AIDS (even leaving aside the difficulty of estimating the distribution of this incubation period).

7.3.2 Role in Evaluation

It should be clear that modelling cannot provide definite answers about current or future rates of HIV infection and future AIDS incidence. However, it is an essential tool for making meaning out of the limited epidemiological information that is available, and it provides guidance on the types of data that would be of most use in enhancing our knowledge of the epidemic's likely direction.

It is argued that an evaluation strategy must be cognisant of the best current information on the state and likely course of the epidemic, even if such information is accompanied by considerable uncertainty. It would certainly be of interest to examine successive models produced (say) year after year using the same techniques, to see whether trends in behaviour and infection patterns are reflected. Unfortunately, it would be very difficult to say whether changes are due to interventions produced by the national strategy - most changes to predictions arise simply from the greater understanding of past phenomena that comes from having observed a bit more of time's progress.

7.3.3 References

Cox, D.R. et al. Short-term prediction of HIV infection and AIDS in England and Wales, Report of a Working Group, London: HMSO (1988).

Report of a Working Group. Acquired immune deficiency syndrome in England and Wales to end 1993: Projections using data to end September 1989. Communicable Disease Report, PHLS, London, January 1990.

Solomon, P.J., Doust, J.A. & Wilson, S.R. 1989, Predicting the course of AIDS in Australia and evaluating the effect of AZT: a first report. Working Paper No.3, NCEPH, ANU, Canberra, 1989.

Whyte, B.M., Gold, J., Dobson, A.J. & Cooper, D.A. Epidemiology of acquired immunodeficiency syndrome in Australia. *Med J Aust*, 145, 65-69 (1987).

Anderson, R.M., Blythe, S.P., Gupta, S. & Konings, E. The transmission dynamics of the human immunodeficiency virus in the male homosexual community in the United Kingdom: The influence of changes in sexual behaviour. *Phil Trans R Soc (London)*, 325, 45-98, (1989).

7.4 Selective Literature Review on HIV/AIDS Evaluation

Dr Rod O'Connor, O'Connor and Associates, Melbourne.

No studies specifically evaluating a country's national HIV/AIDS Strategy were accessed. A number of studies evaluating particular parts of the Strategy were located and are reviewed here.

7.4.1 *The Need for Accurate and Predictive Modelling*

Understanding of the dynamics of the AIDS epidemic is critical. As discussed by Brandt (1989), health policy issues associated with AIDS are complex and controversial, including health care, research priorities, education, testing and confidentiality, and estimates of the future numbers of symptomatic persons and of the mix of severity are necessary for developing the capability to provide and fund health care services. Models may also suggest other data needs and research directions. An important policy role of models is to permit evaluation of potential intervention steps and to project negative consequences.

As well as future case numbers, econometric modelling is required to evaluate or project the costs-of-illness to the individual and to society; and to examine the efficiency and cost-effectiveness of medical and social services provided to patients. Most of the standard theoretical and statistical approaches are not applicable because of the scarcity, incompleteness and non-representative nature of available data.

Hays (1989) has discussed specific methodological approaches concerning AIDS and HIV epidemiology, medical cost estimation techniques, evaluation of the cost-effectiveness of social support programs and valuation of human life, and has proposed methods for estimating both the direct and indirect costs of illness.

Drummond and Davies (1988) have reviewed the evidence on the costs of treating AIDS, comparing European data with those from the U.S.A. They also investigate the reasons for variations in cost estimates, the likely impact of changing treatment patterns and the relative 'value for money' from treatment of AIDS, compared with other health care interventions.

7.4.2 *HIV Testing and Under Reporting*

Central to strategy formulation (and accurate modelling) is accurate and reliable HIV testing and reporting. Valdiserri et al (1990) have pointed out that ensuring high quality in human immunodeficiency virus type 1 (HIV-1) antibody testing is an essential component of the organised public health response to epidemic HIV-1 infection, and describe the Centers for Disease Control Model Performance Evaluation Program to assess and improve the analytic quality of HIV-1 antibody testing, and to gather information about HIV-1 antibody testing

practices. The responses to a questionnaire completed by 1050 program-participant laboratories (both within and outside the United States and consist primarily of hospitals, bloodbanks, health departments, and independent laboratories) in September 1988 suggest that at several stages in the HIV-1 antibody total testing process, laboratory practices (including the interpretation of Western blot patterns) were found to be variable, and standardisation of these practices was recommended to improve quality.

Regarding the issues of informed consent and confidentiality, a West Australian study has been reported by Grove and Mulligan (1990). A survey was undertaken of all of the consultant staff members of Perth's major teaching hospitals together with all the fellows of The Royal Australian College of General Practitioners in Western Australia in order to define their views on the issues of informed consent, compulsion in relation to surgery, and confidentiality in a particular circumstance, when testing for infection with the human immunodeficiency virus (HIV). They found that the majority opinion of the most senior members of the medical profession is that specific, informed consent should not always be required, that there is great support for compulsory testing, and that confidentiality may be broken under certain circumstances. They argued that these views should be recognised by administrators and legislators when framing measures to control this infection.

As detection is vital, so the accurate central reporting of numbers is also important (to allow estimation of the spread). In a Canadian study, Calzavara et al (1990) estimated the rate of underreporting of AIDS to the Federal Centre for AIDS (FCA). The rate of underreporting was estimated to be between 12% to 18%. The specific diagnosis was not related to the rate of underreporting, and differences in the observed rates of underreporting were suggested to be the result of differences in the reporting responsibilities of physicians involved in the studies.

An earlier Canadian study (Johnson et al, 1989) found under reporting rates and argued that the deficiency in AIDS case reporting could adversely affect the long-term planning of health care resources and the development of programs to prevent and control the spread of AIDS.

7.4.3 Education

As described by Van Dam (1989), in the absence of a vaccine effective treatment, prevention of infection with the human immunodeficiency virus (HIV) is presently the only feasible option for the control of AIDS. Most transmission is related to high risk behaviours, and attempts to change these behaviours are key objectives of many AIDS Control Programs. Health Education, or Information, Education and Communication (IEC) strategies have been developed to achieve these objectives. Yet few evaluations of health education Programs have been carried out, and the intimate relationship between AIDS and other sexually transmitted diseases is often ignored. Behaviour has a multitude of determinants, and consequently no simple approach focusing on one aspect of behaviour can be expected to be successful. Not only does knowledge

need to be imparted, but attitudes, beliefs, social-and peer-group norms need to be addressed. The development of IEC Programs is hampered by the specific routes of transmission of the HIV, involving sensitive areas of human behaviour, and by a lack of knowledge about these behaviours. Nevertheless, Van Dam argues that a number of seemingly successful Programs have been carried out, with active participation of the target groups appearing to be a key factor in success. However rigorous documentation and strict evaluation is proposed as necessary to provide justification and to maximise the potential of this strategy.

Strategies that have been reported include that by Brown et al (1989), who conducted school-based education campaigns. The seventh and tenth grade classes (n = 313) in one Rhode Island school district were surveyed regarding AIDS knowledge, attitudes, and coping skills before and after a two-session educational program. Following instruction, students reported more knowledge, greater tolerance of AIDS patients, and more hesitancy toward high-risk behaviours, but the changes were modest. These data were seen to lend support to the potential usefulness of school-based education programs, especially if behaviour and coping skills are emphasised in the education. However it was recommended that Evaluation of such programs should include assessment of change in subgroups to determine components of programs that can be tailored for more effective education.

In a study of tertiary students, Dommeyer et al (1989) reported on an AIDS education campaign at a California college campus. A pretest-posttest design was used to determine whether the AIDS-related attitudes of students, faculty, and staff were affected by an AIDS Awareness Week. The results showed that the awareness week was successful in exposing the campus community to AIDS information. It was only marginally effective in changing AIDS-related attitudes because pretest attitudes were already at desirable levels, only a fraction of the campus community attended the highly motivating events of the campaign, and faculty did not actively support the goals of the campaign.

Rhodes and Wolitski (1989) have also reported on the effectiveness of commercial videotape presentations in changing AIDS knowledge and attitudes.

Of considerable importance when considering educational issues is the knowledge and attitudes of health care professionals. Walton et al (1989) have reported that many health care providers have misconceptions and fears about AIDS and possess negative attitudes about caring for people with the disease. Walton et al argue that merely transmitting the basic facts about AIDS is insufficient to prepare health care workers to deal with AIDS patients, arguing for continuing education for health care professionals and educators and offering a model for the development of educational programs.

Such workers include occupational health nurse, who Nyamathi and Flaskerud (1989) suggest are in a unique position to influence the attitudes, knowledge, and behaviours of workers in

promoting a safe and productive work environment and reducing high risk activities. In an analogous way, Alperin and Richie (1989) suggest that Social Workers are particularly involved in working with Community-based AIDS service organisations, being fully enmeshed in these new community health care agencies, performing a variety of social work and administrative functions. Hence it is argued that social work educators need to add information about community-based AIDS service organisations to the existing health care curricula.

7.4.4 *Issues Relevant to Needle Exchange Programs and Condom Use*

Needle exchange

Hart et al (1989) have reported on the monitoring of visits made and syringes dispensed and returned at the needle exchange of the Middlesex Hospital, London, UK. A sample of clients were interviewed 1 month after entry to the scheme and again 3 months later to evaluate changes in injecting and sexual risk behaviours for HIV infection. Clients were asked to donate saliva for anti-HIV immunoglobulin G (IgG) antibody capture radioimmunoassay (GACRIA). It was found that the rate of lending and borrowing of used injecting equipment fell, both compared with rates prior to entry to the scheme and during the period of study, while frequency of injecting did not increase and there was reduced incidence of abscesses. There was also a highly significant correlation between multiple sexual partners and condom use and a reduction in the proportion of clients with multiple partners. The scheme was found to attract clients, reduce injecting-related risk for HIV infection and to have high equipment return rates. Saliva testing was acceptable to clients, with continued monitoring of anti-HIV in saliva indicated.

In a further survey, Hartgers et al (1989) interviewed a group of 145 injecting drug users (IVDUs) in Amsterdam about their drug use, participation in a needle-exchange Program and needle sharing. It was concluded that a needle exchange is an effective prevention Program against the spread of HIV infection, but that efforts have to be made to reach the group of younger short-term injectors and those IVDUs who are not in contact with methadone-maintenance Programs. There were indications that regular injectors in particular exchange needles, and young male injectors were more at risk of borrowing, independently of exchanging. Hence it was argued that an exchange Program should be complemented with other prevention approaches, i.e. intensive counselling and the spread of leaflets with information on cleaning used needles with bleach.

Condoms

The use of condoms has assumed a central position in the current strategy to prevent sexual transmission of the human immunodeficiency virus. The effectiveness of condoms in disease prophylaxis is dependent, to a degree, on their correct use. Condom manufacturers routinely

include information on condom use either printed on the actual package or in an enclosed package insert. Richwald et al (1988) examined the readability of 14 different sets of instructions included with 25 brands of condoms. The readability formulas, when applied to instructions for condom use, estimated that, conservatively, 8 of the 14 instructions required at least reading at the level of a high school graduate and none required less than a 10th grade level.

They concluded that clearly written instructions and simple concepts could assist current and future condom users in the correct use of condoms and improve the effectiveness of condoms.

7.4.5 Conclusion

Information relating to HIV/AIDS programs is accumulating rapidly. Because of the proliferation of information about AIDS and the importance of disseminating information in the fight against AIDS, professionals must now rely on current AIDS information often found only in computerised sources. Of considerable use in this regard, Branch (1988) identified computerised sources of AIDS information, including databases, and bulletin boards. Four major AIDS-specific databases are reviewed and compared using three cross-database searches.

7.4.6 References

- Branch K., Computerized sources of AIDS information. *Med Ref Serv Q.* 1988;7(4):1-18.
- Van Dam C.J., AIDS: is health education the answer? *Health Policy Plan.* 1989;4(2):141-7.
- Walton R.A.; Rukeyser J.; Mastroianni P.; Doyle S.; Hunter A.; Mautner B.; Whitaker P, Education to meet the challenge of AIDS. *J Contin Educ Health Prof.* 1989;9(1):43-9.
- Hay J.W., Econometric issues in modeling the costs of AIDS. *Health Policy.* 1989;11(2):125-45.
- Drummond M., Davies L., Treating AIDS: the economic issues. *Health Policy.* 1988 Aug;10(1):1-19.
- Valdiserri R.O.; Taylor R.N.; Hearn T.L.; Schalla W.O.; Muir H.W., Centers for Disease Control perspective on quality assurance for human immunodeficiency virus type 1 antibody testing. Model Performance Evaluation Program. *Arch Pathol Lab Med* 1990 Mar;114(3):263-7.

Grove D.I.; Mulligan J.B; Consent, compulsion and confidentiality in relation to testing for HIV infection: the views of WA doctors. *Med J Aust* 1990 Feb 19;152(4):174-8.

Calzavara L.M.; Coates R.A.; Craib K.J.; Schechter M.T.; Le T.N.; Nault P.L.; Elmslie K., Underreporting of AIDS cases in Canada: a record linkage study. *Can Med Assoc J* 1990 Jan 1;142(1):36-9.

Feinfeld D.A.; Kaplan R.; Dressler R.; Lynn, R.I., Survival of human immunodeficiency virus-infected patients on maintenance dialysis. *Clin Nephrol* 1989 Nov;32(5):221-4.

Dommeyer C.J.; Marquard J.L.; Gibson J.E.; Taylor, R.L., The effectiveness of an AIDS education campaign on a college campus. *J Am Coll Health* 1989 Nov;38(3):131-5.

Brown L.K.; Fritz G.K.; Barone V.J., The impact of AIDS education on junior and senior high school students. A pilot study. *J Adolesc Health Care* 1989 Sep;10(5):386-92.

Hartgers C.; Buning E.C.; van Santen G.W.; Verster A.D.; Coutinho R.A., The impact of the needle and syringe-exchange Program in Amsterdam on injecting risk behaviour. *Aids* 1989 Sep;3(9):571-6.

Nyamathi A.; Flaskerud J.H, Effectiveness of an AIDS education program on knowledge, attitudes, and practices of state employees. *AAOHN J* 1989 Oct;37(10):397-403.

Alperin D.E.; Richie N.D., Community-based AIDS service organizations: challenges and educational preparation. *Health Soc Work* 1989 Aug;14(3):165-73.

Johnson R.J.; Montano B.L.; Wallace E.M., Using death certificates to estimate the completeness of AIDS case reporting in Ontario in 1985-87. *Can Med Assoc J* 1989 Sep 15;141(6):537-40.

Hart G.J.; Carvell A.L.; Woodward N.; Johnson A.M.; Williams P.; Parry J.V., Evaluation of needle exchange in central London: behaviour change and anti-HIV status over one year. *Aids* 1989 May;3(5):261-5.

Rhodes F.; Wolitski R., Effect of instructional videotapes on AIDS knowledge and attitudes. *J Am Coll Health* 1989 May;37(6):266-71.

Richwald G.A.; Wamsley M.A.; Coulson A.H.; Morisky D.E., Are condom instructions readable? Results of a readability study. *Public Health Rep* 1988 Jul-Aug;103(4):355-9.

Schafer I., AIDS help--evaluation of a public health movement: *Off Gesundheitswes* 1989 Aug-Sep;51(8-9):567-70.

Brandt E.N., Policy implications of modelling of the AIDS epidemic. *Stat Med* 1989 Jan;8(1):137-9.