



WHO



UNDCP



EMCDDA

WHO/MSD/MSB/00.5
Original: English
Distribution: General

***International
guidelines***
*for the evaluation of
treatment services and
systems for psychoactive
substance use disorders*

Disclaimer

This document contains collective views of an international group of experts and does not necessarily represent the views of the WHO, UNDCP or EMCDDA.

© World Health Organization, 2000

This document is not a formal publication of the World Health Organization (WHO) and all rights are reserved by the Organization. The document may, however, be freely reviewed, abstracted, reproduced and translated, in part or in whole, but not for sale or for use in conjunction with commercial purposes

International guidelines for the evaluation of treatment services and systems for psychoactive substance use disorders

John Marsden

Institute of Psychiatry, King's College London, United Kingdom

Alan Ogborne

Alan Ogborne and Associates, London, Ontario, Canada

Michael Farrell

Institute of Psychiatry, King's College London, United Kingdom

Brian Rush

Centre for Addiction and Mental Health, Toronto, Ontario



WHO
World Health
Organization



UNDCP
United Nations
International Drug
Control Programme



EMCDDA
European Monitoring
Centre for Drugs
and Drug Addiction

Acknowledgements

We would like to thank to the following reviewers who provided valuable commentaries on various drafts of the guidelines:

Maria Lucia O. Souza Formigoni, Ph.D.

University of Sao Paulo
Escola Paulista de Medicina, Brazil

Professor Christine Godfrey

University of York,
United Kingdom

Professor Wayne Hall

Executive Director
National Drug and Alcohol Research Centre
University of New South Wales, Australia

Jennifer Hillebrand

Management of Substance Dependence
Department of Mental Health and Substance
Dependence
World Health Organization

Dr Alexander Kantchev

National Center for Addiction
Sofia, Bulgaria

Petra Paula Merino

European Monitoring Centre for Drugs and Drug
Addiction (EMCDDA)

Dr Maristela Monteiro

Management of Substance Dependence
Department of Mental Health and Substance
Dependence
World Health Organization

Margareta Nilson

European Monitoring Centre for Drugs and Drug
Addiction (EMCDDA)

Ms Christina Gynna Oguz

Chief, Demand Reduction Section,
United Nations International Drug Control
Programme (UNDCP)

Dr Charles Parry

Mental Health and Substance Abuse
National Urbanisation & Health Research
Programme
Medical Research Council, South Africa

Dr Vichai Poshyachinda

Chief of Drug Dependence Research Center
Institute of Health Research
Chulalongkorn University, Thailand

Dr Vladimir Poznyak

Management of Substance Dependence
Department of Mental Health and Substance
Dependence
World Health Organization

Dr Edle Ranvndal

National Institute for Alcohol & Drug Research,
Norway

Dr Margret Rihs-Middel

Federal Office of Public Health
Main Unit Dependency and Aids
Section Policy and Research, Switzerland

Dr Juana Tomás Rosselló

Drug Abuse Treatment Adviser,
Demand Reduction Section
United Nations International Drug Control
Programme (UNDCP)

Dr Maree Teesson

National Drug and Alcohol Research Centre
University of New South Wales, Australia

This project has been partly developed under the
directorship of Dr Mary Jansen, former WHO
Substance Abuse Department.

Foreword to the WHO/UNDCP/EMCDDA guidelines for treatment evaluation

This publication is the result of an exemplary collaboration between the World Health Organisation (WHO), the United Nations International Drug Control Programme (UNDCP) and the European Monitoring Centre for Drugs and Drug Abuse (EMCDDA). Already in 1996 UNDCP and WHO initiated joint work towards the development of tools for programme managers to carry out evaluation of their existing treatment approaches and to help them decide on the allocation of resources and access to effective treatment to their patients. A series of workbooks on evaluation of treatment were prepared and field tested in several countries, and should be seen as complimentary to the present guidelines, for training and planning purposes.

The association of the EMCDDA with this project started in 1997 in Athens with a seminar organised jointly by the EMCDDA, the Greek National Focal Point and WHO. The seminar aimed at reaching consensus among key European scientists, networks and policy makers in the area of drug abuse treatment evaluation and the role of the EMCDDA in this field. Since then, the EMCDDA became a partner of WHO and UNDCP in this project, and will now contribute to the diffusion of the guidelines and the associated workbooks across the EU.

Since the inception of the project, further international developments have occurred and need mention. The Declaration on the Guiding Principles of Drug Demand Reduction and the Action Plan for its implementation adopted by United Nations Member States in 1998 and 1999 respectively, highlight the importance and need for evidence-based practices as well as of learning from experiences. These are key pillars for the development and implementation of enhanced drug demand reduction strategies and programmes, a target to be reached by Member States by 2003 in accordance with the 1998 UN General Assembly Political Declaration on the Drug Problem. The UNDCP is committed to providing guidance and assistance to Member States in reaching this target, and views these guidelines for treatment evaluation and the workbook series as key instruments to encourage the assessment of the effectiveness of drug abuse treatment strategies and activities, and the exchange of evaluation results.

The Helsinki European Summit endorsed the EU Drug Strategy 2000-2004 in December 1999. The Strategy provides a framework for all drug-related activities in the EU over the next five years. The European Union

Action Plan on Drugs (2000-2004) set out an agreed programme for the implementation of the principles and targets set in the Strategy. The Action Plan states that The European Monitoring Centre on Drugs and Drug Addiction will establish guidelines for the evaluation of drug policies in Europe.

The World Health Organisation decided to make 2001 the year for Mental Health. One of the major objectives of the coming World Health Assembly 2001, World Health Day 2001, and the World Health Report in 2001 will be the control of mental health problems. The promotion of this objective through the WHO will also include recommendations for effective service delivery in the field of substance dependence. The present guidelines will serve as one step towards improved treatment for psychoactive substance disorders internationally.

In the past, evaluation of treatment services has not been recognised as a means to enhance the efficacy of existing services. However, due to recent developments in health and social care systems, and budget constraints, the need for more systematic evaluations has increased. Evaluations not only help to justify financial support for treatment services; they are also essential to increase and maintain best clinical practice. The success of future evaluations will partly depend on the availability of information on how to implement and plan evaluations. For this reason, practical and comprehensive guidance on evaluation is necessary. This publication is especially addressed to policy makers, programme managers and service providers in the field of drug demand reduction who will find the guidelines a unique source of reference. The guidelines will help decisions to be made about critical aspects, especially at the moment of conceptualising, planning and commissioning the evaluation of treatment services. They should also serve as a valuable reference in designing evaluation and to make optimal use of available resources. In order to improve drug demand reduction policies in the area of assistance to drug abusers, the guidelines are also aimed at disseminating critical, scientifically supported knowledge in this regard.

We would like to thank the authors of the guidelines for providing a compact and succinct overview of different types of evaluation. We would also like to express our thanks to all experts and policy makers who helped to review the guidelines by providing their valuable comments on earlier drafts.

Contents

2	Acknowledgements
3	Foreword
5	List of boxes
6	Preface
7	Background
7	Structure of the guidelines
8	Section 1: Setting the stage for evaluation
8	Communication with interested parties
9	Resources
9	Setting objectives
12	Getting specialist advice
12	Ethical issues
13	Section 2: The main types of evaluation
13	Needs assessment
15	Process evaluation
16	Client satisfaction studies
17	Outcome evaluation
18	Naturalistic, observation studies
18	Experimental, controlled designs
19	Economic evaluations
19	Cost analysis
19	Cost-effectiveness analysis
20	Cost-utility and cost benefit analysis
21	Section 3: Special issues
21	Reliability, validity and sensitivity of measures
21	Reliability of self-report data
21	Time periods for measures
22	Assessing non-treatment factors that may affect outcome
22	Subject recruitment and follow-up
23	Sampling and sample size
23	Statistical testing and confidence intervals
24	Section 4: Reporting results
24	Increasing the impact of evaluation
26	Summary and conclusions
27	References

List of boxes

- Box 1**
- 9 Components of a treatment programme
- Box 2**
- 10 Logic model for a hypothetical detoxification centre
- Box 3**
- 11 Logic model for hypothetical youth services programme
- Box 4**
- 13 The five major types of evaluation
- Box 5**
- 14 Major types of activities in need assessments
- Box 6**
- 15 Questions and issues addressed in process evaluation
- Box 7**
- 16 Client characteristics considered in process and/or outcome studies
- Box 8**
- 17 Major activities in process evaluation
- Box 9**
- 18 Questions and issues addressed in client satisfaction studies
- Box 10**
- 25 Structure and content of a formal report from an outcome evaluation

Preface

These guidelines describe methods for the evaluation of treatment services and systems for substance use disorders. The guides are intended to be a companion resource to the World Health Organization (WHO), United Nations Drug Control Program (UNDCP) and European Monitoring Centre on Drugs and Drug Addiction (EMCDDA) workbook series on the evaluation of costs and effects of treatment for substance use disorders, co-ordinated by Dr. Maristela Monteiro and Dr Brian Rush (WHO, 2000).

The main audience is policy makers, commissioners of treatment services and treatment agency personnel who want to know more about research evaluation and commission or undertake evaluation studies. We hope that there will be much to interest the international research community as well.

The guidelines are intended to be of particular interest for countries where research and evaluation in this area is not widely developed. Our aim is to offer a concise description of the main evaluation methods and to help the reader to select the best type of study to answer specific questions. The goal is to encourage readers to the use of evaluation to help develop and maintain effective and efficient treatment services and treatment systems.

These guidelines are not intended to be a detailed, step-by-step instruction manual on how to do research. The interested reader should obtain the related series of workbooks which cover this in detail (WHO/UNDCP/EMCDDA, 2000). We have referred to particular workbooks in several places for further reading and this document is a natural companion to the series. There are also other valuable sources of information which can be accessed via the internet. For example, valuable information can be found at the Cochrane Collaboration site for information on treatment effectiveness (see <http://www.cochrane.org>); an evaluation instrument bank can be accessed at the EMCDDA site (see <http://www.emcdda.org>); and a series of best practice guides on treatment and evaluation issues can be obtained from the US Center for Substance Abuse Treatment (see <http://text.nlm.nih.gov>).

We have tried to prepare the guidelines using a non-technical language. Nevertheless, some evaluation issues of a more technical nature have been covered in Section 3: Special Issues. Those who wish to read a more general account of evaluation methods and issues should skip this section. Finally, we do encourage those planning an evaluation to obtain expert advice wherever feasible. Getting appropriate advice early on in the planning of an evaluation is very important and can ensure that appropriate resources are identified, that the correct evaluation design is selected, and the data is analysed and correctly presented.

We hope to find the guidelines useful and warmly welcome feedback and suggestions.

Please write to:

World Health Organization (WHO)
Management of Substance Dependence
Department of Mental Health
and Substance Dependence
Avenue Appia 20
1211 Geneva 27
Switzerland

Background

The misuse of illicit drugs, prescription drugs or alcohol is a world-wide problem and there is an international commitment to providing prevention and treatment services. Estimates have suggested that the annual global population prevalence of illicit drug use is between three and four percent; prevalence estimates for heroin and cocaine use are approximately 8 million and 13.3 million adults respectively (United Nations International Drug Control Programme [UNDCP], 1997). Prescription drug abuse is known to be high in some countries (Adams et al., 1993) and alcohol-related problems can far exceed those involving all other drugs in term of prevalence and costs to society (Single et al., 1996). Several different types of services and treatments have been developed for substance abuse disorders. These can be broadly categorised as follows:

- “Open access” services:
 - advice, education, information, early intervention programmes;
 - needle/syringe exchange programmes (in some nations/regions)
- “Structured” services:
 - prescribing interventions (inpatient and outpatient/community settings);
 - community-based psychosocial counselling and relapse prevention;
 - residential rehabilitation programmes

These five categories are really only part of the picture. Social reintegration and support services may sometimes be available to support treatment gains from the above services. Space does not permit us to give a full account of the various modalities of treatment and the array of different providers. These concern the specific approach or philosophy of a treatment – such as maintenance or reduction regimes in substitution treatment for opiate users, or the style of psychosocial counselling (e.g. cognitive behavioural), or the philosophy of residential rehabilitation (e.g. 12-step; therapeutic community).

It is quite common for a treatment programme to contain several different therapeutic components linked together (e.g. methadone maintenance with cognitive behavioural psychotherapy). It is also quite common for a person with substance-related difficulties to receive health and social care services from several service providers during the course of their

treatment. Support and assistance may also come from the family and friends and other support networks. These practical aspects of treatment have implications for how evaluation research in this field is constructed and the questions addressed.

There is reason for optimism about investing in the treatment of people with substance-related problems. Many careful studies have shown that treatment can result in short, medium, and sometimes long-term reductions in substance use, improvements in health and reduced demands placed on health and social services among those treated (see reviews by Landry 1995; Miller and Hester, 1986; WHO Expert Committee on Drug Abuse, 1996; Roberts and Ogborne, 1999; National Institute on Drug Abuse, 1999). There is also good evidence that some types of treatment can reduce criminality and the associated costs of law enforcement among substance abusers who maintain their habits by crime (Tims, 1995). The economic benefits of treatment can also exceed treatment costs (Holder and Blose, 1992; Holder et al., 1991; Centre for Substance Abuse Treatment, 1999) and by the same criteria, treatment compares favourably with interdiction and other law enforcement efforts directed at individual drug users (Rydell and Everingham, 1994).

Structure of the Guidelines

The guidelines have four sections. Section 1 sets the stage for evaluation research and describes the importance of assessing the feasibility of a study in terms of the questions to be addressed, the resources required, ethical considerations and consultation with interested parties. Section 2 summarises the key features of the five main types of research evaluation: needs assessment, process evaluation, client satisfaction studies, outcome evaluation and economic evaluations. Section 3 describes seven special issues of a more technical nature: the reliability, validity and sensitivity of measures, reliability of self-report data, time periods for measures, assessing non-treatment factors that may affect outcome, subject recruitment and follow-up, sampling and sample size, and statistical analysis and testing. Section 4 offers guidance of report writing and ensuring results from evaluation studies are effectively presented to key audiences.

Section 1

Setting the Stage for Evaluation

An evaluation is done when one wants to assess the costs, effects or impact of a treatment or treatment system. Evaluation therefore provides feedback to key audiences to help them with various types of decisions. A healthy culture for evaluation is one that encourages the routine collection and dissemination of information to help improve how services are delivered. Nevertheless, in many countries, the evaluation of treatments for substance use disorders is in development. In fact, evaluation may well have had a minor role in influencing what treatment services have been developed in a country. Some policy makers and planners have relied on personal experience, opinions and testimonials rather than research data or published studies. This has led to disparities in the development and management of treatment services and to the support of interventions of uncertain effectiveness or efficiency. As we shall show, evaluation research can help identify which treatment services are needed as well as assess the efficiency and effectiveness of existing services.

Prior to starting any type of evaluation, it is a good idea to assess the feasibility of the desired research (see Wholey, 1994). A feasibility assessment should examine the following:

- results from previous research;
- the specific questions to be addressed;
- the size and scope of the study;
- the human resources which will be needed;
- the skills needed by the research team;
- the financial resources needed;
- the timetable for the study

Communication with interested parties

Discussions with interested groups – such as service providers, clients of treatment programmes, and treatment funding agencies – can be an invaluable means of learning about the extent of support for the study, the direction it should take and practical issues concerning its implementation. In most situations those with the greatest interest in evaluation (the stakeholders) will be treatment programme personnel, representatives from the community, funding bodies and government. It may be very helpful for some stakeholders to serve as members of an advisory committee for the study.

Good communication with key stakeholders throughout the implementation of an evaluation is vital and they should be involved in an early discussion of findings and implications. It is also important that a clear understanding of the information requirements and interests of the funding body is secured. Discussions with all relevant stakeholders should be held at the outset and their views and concerns sought throughout the study. These discussions will help to formulate the central questions to be addressed. The aim is to clarify who wants to know what, by when, with what degree of precision, and at what cost. Each stakeholder may have unique experiences and perspectives that can contribute to the overall understanding of the issues and to the design and implementation of useful evaluations.

Different groups may, of course, have different ideas or emphases on what to evaluate. For example, policy makers and service purchasers may be most interested in costs and efficiency, while service provider staff may be more interested in assessing the benefits of a new treatment. Naturally, the number of questions which are worth looking at may over-stretch the time and resources available. If this is the case, it is essential that the evaluation team look at the questions that have the highest priority.

Resources

Some evaluations can be done quite quickly and do not require major resources. In fact, relatively simple and modest studies can be of great value in learning more about how services operate and how they can be improved (Rossi and Freeman, 1982). Other types of evaluation are complex and may need a substantial investment of funds to be done well. It is also important to ensure that those to be involved will have enough time to devote to critical tasks and that the timetable for the project is realistic. In some cases resources are allocated and set aside when a treatment programme is in the planning phase. This is the ideal situation and policy makers are encouraged to support evaluation by earmarking adequate resources for the evaluation of all new treatment initiatives. More commonly, evaluations require special grants from governments and other domestic and international funding bodies.

Setting objectives

When thinking about the objectives for an evaluation of a treatment programme, it is important to be clear about how it operates and what it is expected to achieve. A “logic model” can be of help in representing these elements as a diagram (see Rush and Ogborne, 1991). A logic model describes the different components of a treatment programme and what these are designed to achieve (see definitions box 1). When fully developed, logic models can lead naturally to the design of management information systems for the monitoring of services. A logic model can also be useful when inducting new staff and for communicating with funding bodies, board members, clients and others with an interest in the programme. Two examples of logic models are shown in boxes 2 and 3.

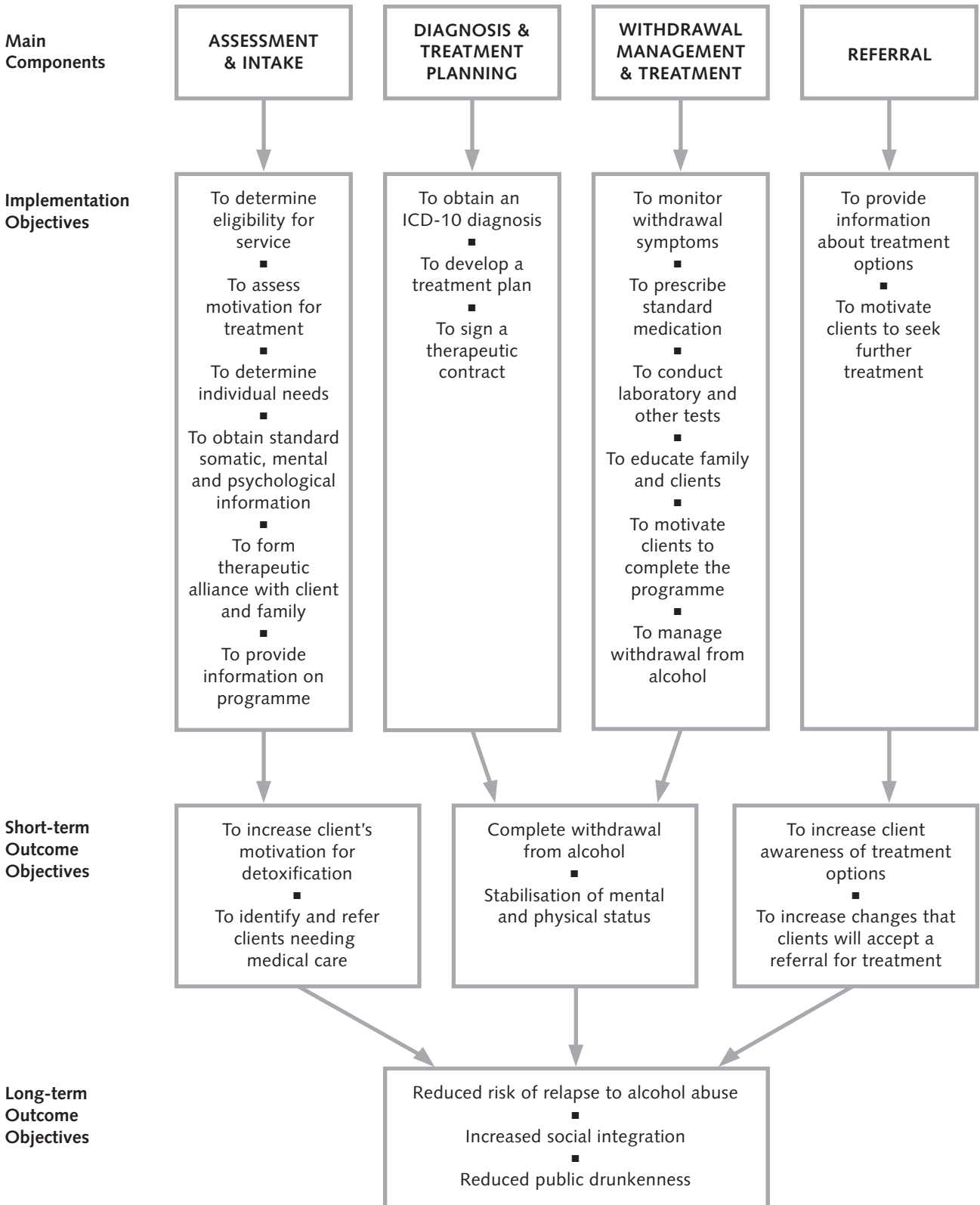
Components of a treatment programme

Box 1

ELEMENTS	DESCRIPTION
Programme or system components	Activities which are directed to the attainment of specific goals
Implementation objectives	What a programme or system seeks to achieve
Short-term outcome objectives	Changes which the programme or system seeks to induce in its target clients in the immediate or short term.
Long-term outcome objectives	Changes which the programme or system seeks to induce in its target clients in the longer term.

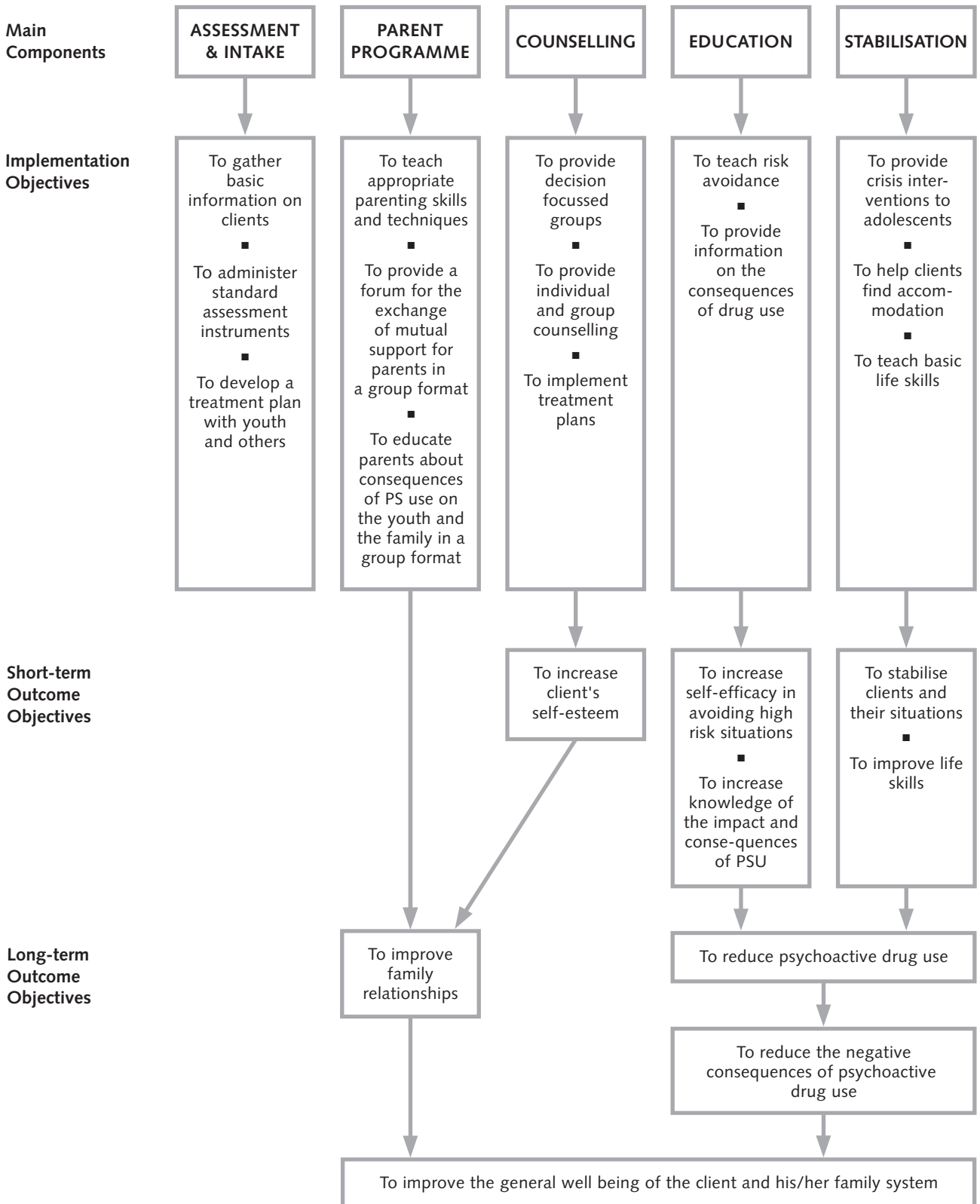
Logic model for a hypothetical detoxification centre

Box 2



Logic model for hypothetical youth services programme

Box 3



Getting specialist advice

On occasions, assistance should be sought from specialists from countries where similar evaluations have been conducted. A certain amount of technical expertise is also required. Particularly by the following areas:

- Economic analysis;
- Effectiveness and cost-effectiveness;
- Treatment system performance;
- Statistical analysis;
- Study planning and conceptualisation;

Although readily available in some countries, other countries may have limited expertise in some of these areas. However, there are usually research institutions and universities whose staff can assist in the planning or conducting evaluations and preparing. Also students from social science, social work and other relevant courses may be willing to assist in data collection and analysis to gain work experience in the evaluation field. It is, however, important to recognise that the involvement of university or college faculty and students may have implications for the timing of evaluation.

Ethical issues

Evaluation can raise ethical concerns. It is critical to ensure that everyone involved understands the purposes of the evaluation and gives informed consent to participate. The concerns of potential participants should be addressed at this stage, particularly issues relating to confidentiality. Anyone asked to provide information for an evaluation or to undergo other procedures for research purposes (e.g. random assignment to a new form of treatment) should be asked to sign an informed consent form if they wish to participate, following adequate information and resolution of any questions. Confidentiality of information should be assured and respected. If there are circumstances under which confidentiality could be breached (e.g. records could be subpoenaed by the courts) then this must be clearly explained in advance. An appropriate ethics review body should consider all studies that involve human subjects.

Section 2

The Main Types of Evaluation

In this section, we describe five main types of evaluation: (a) needs assessment (b) process evaluation (c) client satisfaction (d) outcome evaluation and (e) economic evaluation. These may overlap to greater or lesser degrees. As indicated in Box 4, each type of evaluation aims to answer certain types of questions and the decision to undertake a particular type of evaluation will be determined by the questions of interest.

Needs assessment

The goal of a needs assessment is to estimate the current and/or future nature, prevalence and incidence of substance use disorders in a specific population. This information is then used to help to guide the development and operation of treatment and support services. Ideally, a needs assessment should be undertaken before a treatment or network of treatment services is planned and implemented.

The five major types of evaluation

Box 4

TYPE OF EVALUATION	EXAMPLES OF QUESTIONS ADDRESSED
Needs assessment	What is the nature and extent of substance use disorders in the community? How well do current services address the needs of the population at risk?
Process evaluation	What is the programme/service/system expected to accomplish? To what extent is the treatment implemented as planned (programme fidelity)? Are services of good quality? What are the views and perceptions of the programme staff on the value and quality of the service? What are the characteristics of the clients of the service? Which clients do not complete treatment?
Outcome evaluation	What proportion of clients complete treatment? What proportion of clients reduce their drug use after treatment? Are clients better off than before treatment? Is our treatment better than no treatment or better than other treatments? Why does the programme achieve the outcomes that it does?
Client satisfaction	Do clients feel that the treatment met their needs? What did clients like least or find unhelpful? Why do some people drop out?
Economic evaluation Cost studies Cost effectiveness Cost utility Cost benefit	What is the average cost per client? How do two or more treatments compare in terms of cost per unit of outcome? Given their relative costs, how do two or more treatments compare with respect to the quality and quantity of the life of their clients after treatment? Does the programme yield more benefits than costs valuing everything in money terms?

- Defining the geographic boundaries of the population(s) whose needs are to be considered (local, regional or national)
- Use of existing population data on the prevalence and incidence of specific problems in the target population and in sub-groups
- Use of indicators to provide indirect estimates of the prevalence and incidence of specific problems in the target population and in sub-groups
- Special population surveys that focus on the issues of concern
- Focus group discussion with key stakeholders (commissioners, clinicians, treatment providers and service users) to explore what they want from services
- Reviews of service arrangement in other, similar jurisdictions
- Reviews of the literature on the cost-effectiveness of particular types of interventions
- Identifying current treatment services and auditing their capacity, use and waiting lists
- Comparing existing arrangements with generally agreed-upon needs or with different perspectives on existing needs
- Preparation of recommendations for increasing treatment coverage, purchasing efficiency and service effectiveness.

There is no universally agreed definition of 'need' or how one should assess it. From the perspective of an individual, "felt need" is what the person wants (Bradshaw, 1972) and from an epidemiological perspective, need can be defined by the ability of that population to benefit from effective health care services (Stevens and Raftery, 1994). Sometimes 'need' is used to refer to what people in a community believe should be provided or what problems professionals think should be treated. A needs assessment can address the gap between existing services and those needed to adequately treat or manage known problems in the target population (Marsden et al., forthcoming). This view of needs explicitly links the research evidence for particular interventions into health and social care planning and delivery.

Some needs assessments use statistical methods to estimate the size of the population in need of treatment from survey, treatment utilisation and/or other indicator data (see DeWit and Rush, 1996). Needs assessments can also draw on public and professional opinions about what works and what should be done and they are often commissioned by funding agencies wanting guidance on how to allocate existing or new funds. The main types of activities

involved in needs assessments are summarised in box 5.

Ideally, needs assessment should be an ongoing and flexible process that is sensitive to changes in the levels and types of problems within the areas of concern. Change may be seen in the following areas:

- New trends in substance use and consumption patterns in the target population;
- Movements in the geographical distribution and concentration of psychoactive substance use;
- Fluctuations in the demand for services;
- New barriers to accessing treatment encountered by special groups;
- The relationship between substance use and other conditions (notably HIV infection, and blood borne viral hepatitis);
- National, regional and local policies in response to drugs;
- The organisation of health and social care services;
- Changes in the resources available for treatment; and
- New evidence for the effectiveness of treatment interventions.

Process evaluation

Process evaluations look at how a treatment or a programme operates. Issues such as the efficiency and quality of treatment services are addressed and an evaluation made of the extent to which a treatment has been implemented as intended (fidelity assessment). Process evaluations also try to identify areas where improvements can be made. Process evaluations are not directly concerned with changes in clients that may result from the treatments received. They focus on how treatment services or systems operate and the ways in which resources are used to produce outputs (e.g. number of clients assessed; numbers treated). Specific questions for a process evaluation can include the following:

- The extent to which a service or system is conceptually or theoretically sound;
- Whether the clients served are those intended;

- Whether clients experience difficulty in accessing the programme;
- The extent to which the programme staff are adequately trained and the extent to which they are satisfied with their work.

Process evaluations may also consider internal and external coordination, organisational and service management issues and case management and record keeping (see box 6 and 7). The steps that may be taken when conducting a process evaluation are shown in box 8:

Questions and issues addressed in process evaluation

Box 6

- Focus on single programme, service or agency
- Number and characteristics of clients served (see box 7)
- Nature and extent of treatment(s) provided
- Ease of access to target groups
- Costs of service
- Programme logic
- Staff qualifications
- Quality of records
- Discharge planning and aftercare arrangements
- Case management
- Internal and external co-ordination
- Compliance with established standards
- Characteristics of environment that impact programme delivery
- Focus on a network or system of services or agencies
- Characteristics of clients served by different agencies (see box 7)
- Referral and cross-referral patterns
- Case management across services
- Intra-agency co-ordination in planning for direct client service
- Characteristics of broader environment that impact the function on the system

Process evaluations are closely related to programme audits and continuous quality improvement (CQI) initiatives. CQI focuses on customer satisfaction, accurate measurement of activities, ongoing improvement of services and operational processes, and meaningful involvement of people at all levels in the organizational process. Other ingredients, including programme logic models, integrated care pathways, client satisfaction surveys and routine outcome monitoring, can all be part of CQI. The results of process evaluations may have important implications for individual staff and managers.

Space limitations preclude a discussion of routine treatment utilisation and audit systems. These provide valuable information about the operation of a treatment system. Setting up monitoring systems that routinely provide information on the operation of a treatment service or system can also reduce the costs of process evaluations.

Client satisfaction studies

The assessment of client satisfaction can add an important “consumer” perspective to evaluations of treatment services and systems (Lebour, 1983). Without such a perspective, problems such as high drop out or low success rates might be related only to client factors when, in reality, these problems may reflect shortcomings in the services provided or in the treatment system. However, it is important to recognise that evidence of client satisfaction is not, in itself, sufficient to establish the effectiveness of a treatment. Clients with no base for comparison may be satisfied with services that more objective measures show are of poor quality. In the short-term, clients may also indicate high satisfaction with services that are well run yet quite ineffective. Likewise, clients may express dissatisfaction with services that are successful nonetheless in reducing their substance use.

Client satisfaction surveys can be used for routine or periodic “check-ups” on the quality of services from the clients' perspectives. They can also be used to assess client reactions to changes in service delivery. For example, a service may have planned changes to increase its efficiency but be concerned that these could lead to decreased client satisfaction. Client satisfaction surveys are most useful when they are

Client characteristics considered in process and/or outcome studies

Box 7

- Number of clients served
- Age, gender (intake)
- Source of referral (intake)
- Treatment history (intake)
- Motivation for treatment and readiness for change (intake)
- Post-intake treatment from index treatment and elsewhere
- Marital/partnership status
- Education
- Work situation/income levels
- Accommodation status and stability
- Involvement with justice system
- Health and social service use
- Alcohol and drug use levels and patterns
- Level of alcohol/drug dependence
- Adverse consequences of alcohol/drug use
- Mental health problems
- Physical health problems
- Self-esteem
- Coping skills
- Social relationships

designed to meet specific objectives and when they use appropriate methods and measures. Sampling procedures, timing, cultural acceptability and sensitivity of the questions to various levels of satisfaction must all be considered carefully.

The most common method for assessing client satisfaction is to use a confidential self-administered questionnaire. This can be given to clients at discharge or follow-up. In some cases, confidential personal or telephone interviews, or focus groups may be more appropriate. If interviews or focus groups are used, it is preferable to have them conducted by someone who is not directly connected with the

- Review of original planning documents concerning the need for the service/system, the objectives, the specific treatment methods used and the resources that were thought to be required
- Focus groups or interviews with programme participants or others such as community planners, managers, staff, community agency representatives and families
- Development and critique of programme logic models
- Chart audits
- Analysis of information from built-in management information systems
- Observation of the programme(s) in action, for example, to rate the extent and quality of implementation and integration with other programmes;
- Surveys of the population for whom the programme or treatment system was intended and of actual programme participants.

service. This may be an independent evaluator, volunteer or former client trained to take on this role. In all cases clients should be assured that their responses would not in any way affect their present or future treatment.

It is important to take cultural differences into account regarding expectations of feedback on public and private services. In jurisdictions where consumerism is firmly established, frank verbal or written feedback may be freely given. However, direct negative feedback in some cultures may be considered impolite and complaints may only be shared with intimate acquaintances. Direct and challenging questions may also be culturally inappropriate. Experiences with (and attitudes toward) the use of questionnaires, interviews, focus groups and other methods of enquiry also differ between cultures. Methods for soliciting client feedback must take into account the prevailing cultural norms and seek to ensure the use of appropriate methods that assess client beliefs and opinions. In addition to using confidential survey methods there is increasing interest in establishing channels through which service users can have an active input into service delivery. Some of the specific issues that could be addressed in client satisfaction studies are indicated in box 9.

Outcome evaluations

Outcome evaluations look at whether clients of a treatment service or system have changed over time and how much of this can be attributed to the care received. Although such evaluations sometimes involve single clients, usually one or more groups of clients that have received a certain treatment within the system are involved. In the past, most outcome studies of people treated with substance use problems used post-treatment abstinence from substances as a primary indicator of treatment success. However, modern studies focus more on patterns and levels of post-treatment substance use and also consider post-treatment functioning in other life areas. Three other problem domains are usually assessed: health risk behaviours, health problems and various aspects of personal/social functioning, including employment, family and other personal relationship problems and criminal behaviour. Box 7 summarises some of the specific issues commonly considered. Further details and information on specific measurements and procedures are provided in the WHO/UNDCP/EMCDDA workbook on outcome evaluation. We recommend that expert advice is sought in the selection of measures for particular studies, especially in cases where new measures are to be developed. Particularly attention should be paid to issues of reliability and validity (see Section 3: Special issues).

When conducting an outcome evaluation, it is important to develop a theory or at least a framework for accounting for why a treatment process achieves a measured outcome (Chen, 1990). Process evaluations are the logical partner to outcome evaluation and good quality outcome evaluation should always have a process component. There are several different designs that can be used in outcome evaluations.

Naturalistic, observation studies

In naturalistic or observational studies clients are assessed on a set of measures at several points in time – usually before, during and after they complete a treatment programme of interest. These studies are useful for the study of treatment systems where clients engage in multiple treatments of varying intensities and duration. There have been several major multi-site naturalistic post studies conducted including the Drug Abuse Treatment Outcome Study (DATOS) in the United States and the National Treatment Outcome Study (NTORS) in the United Kingdom (<http://www.datos.org>; <http://www.ntors.org.uk>). These sorts of study can be used to show if the desired client outcomes are

achieved and to identify which clients change most or least. They can also show how outcomes vary with the amount or type of treatment received. However, observational designs cannot definitively conclude that any observed changes were caused by the treatment(s) of interest. This is because they do not control for many of the other factors that might explain the changes observed (e.g. non-treatment factors, maturation, cyclical changes).

In some circumstances it may be feasible to obtain one or more comparison groups of cases that are similar to the clients of a target treatment but who have not been treated or have been treated by some alternative means. If the groups are initially very similar, then it can be assumed that any differences over time are at least partly due to differences in treatment. Suitable comparison groups might be cases treated in a different programme or cases included in studies reported in the literature. However, it is important to ensure that the treatment and comparison groups are as similar as possible and not specifically selected to give them an advantage in any one case (e.g. screened for motivation or social stability).

Experimental, controlled designs

Experimental designs involve the allocation of clients to one or more treatments or to a no-treatment or waiting list (delayed treatment condition). In randomised controlled trials, a simple randomisation or group balancing procedure is used to allocate clients to the experimental or comparison treatment (usually the “standard” form of care which is usually delivered) condition. In substance use treatment, these designs are often conducted in the context of the day-to-day operation of a treatment programme and are known as randomised clinical trials or randomised field experiments (see Dennis 1994).

When properly conducted, experimental designs can produce the most convincing evidence that measured outcomes were due to the treatment studied. However the validity of experimental designs can be compromised unless care is taken to adhere to the study procedure and the treatments are implemented as planned. A common problem is that it may be difficult to ensure that those in different treatment conditions actually receive the same amount of

Questions and issues addressed in client satisfaction studies

Box 9

- Physical setting of treatment
- Helpfulness of staff
- Costs of treatment
- Accessibility of services
- Family involvement
- Waiting times
- Information
- Amount of service
- Effectiveness
- Case management and co-ordination
- Confidentiality

attention when two treatment methods are being compared. For example, clients and therapists may not be equally enthusiastic about the treatments being compared. If not, the outcomes for some treatments may be deflated because no one believed they would work or made much effort to deliver them as planned. As with comparison group studies there are technical and logistical problems with the design and conduct of experimental evaluations. Consultation with an experienced evaluator is strongly recommended if an experimental design is being considered.

Economic evaluations

Economic evaluations assess the resources required to provide treatment and the resulting benefits. A central question posed by many economic evaluations is whether the treatment or treatment system studied is an efficient use of resources (see Drummond et al., 1987). Treatment funding agencies are increasingly demanding that treatment services and systems be evaluated from an economic perspective. However, the economic evaluation of treatment is still an emerging discipline and there are only a few published studies that can be characterised as full economic evaluations.

It is important to be aware that economic evaluations are undertaken from a particular perspective and that the results can vary with the perspective taken. Generally, the perspective taken is that of society as a whole or that of the agency providing funds for treatment (usually the government). However, economic evaluations can also be conducted from the perspectives of clients or family members. Overall, economic evaluations can be complex and time consuming; since there are many technical and logistical issues to consider we recommend that consultation with an expert is well worthwhile. There are three types of economic evaluation:

Cost analysis

A cost analysis determines the overall cost of the resources used to provide a treatment activity, service or system. These may include both direct costs (such as clinical staff salaries) as well as overheads such as building and maintenance costs, administrative salaries, equipment and supplies. The amount of direct resources used for specific activities can be measured and the indirect resources apportioned. The results are then used to compute various indicators of economic performance (e.g. costs of residential treatment per client per week or episode; cost per counselling visit). This type of economic evaluation is, however, quite limited and it is important to recognise that the relative costs of treatment programmes may not reflect their relative effectiveness. Sometimes lower cost programmes may be as effective as those costing a lot more (Holder et al., 1991). However, for people who also have multiple social and mental health problems in addition to their substance-related difficulties, comprehensive, and usually higher cost interventions are likely to be more effective than more basic and lower cost interventions (McLellan et al. 1998).

Cost-effectiveness analysis

Cost-effectiveness analysis (CEA) compares the costs and outcomes of two or more different treatments with similar objectives. Examples of outcomes used for a CEA include: "number of cases that achieve a certain outcome or degree of success" or "duration of abstinence of rehabilitation following discharge" or perhaps a "reduction in substance-related problems". Multiple outcomes are sometimes considered and the results may vary with different outcomes. For example, one treatment may be more cost effective than another at reducing drug use but less cost effective at increasing employment. The first stage in CEA is usually to estimate the costs per patient outcome achieved (via a decision analysis). The second stage involves comparison of the cost-effectiveness ratios for each of the treatments studied.

Cost-utility and cost benefit analysis

Cost utility analysis (CUA) uses ratings of the quality of life and life expectancy and survival of clients who have been treated as primary outcome measures. Often these ratings are used to compute a measure of 'quality adjusted life years' (QALYs) that are used to indicate a degree of discrepancy between a client's actual outcome status and an ideal status from a client's own perspective. CUA is increasingly viewed as more appropriate for health case evaluation than cost-effectiveness analysis because the latter does not take intangible benefits or the quality of outcomes into consideration.

In cost benefit analysis (CBA) the costs of delivering treatment and the outcomes from treatment are both measured in monetary units. For example, the costs of health or social services used or the increased tax revenues paid by employed drug users. These types of studies are uncommon and there is some criticism of the appropriateness of measuring outcomes solely in monetary terms. In addition to other technical matters, these studies require attention to the issues of discounting. This occurs because the economic benefits of a programme can occur long after the programme costs are used up. The 'future values' of these benefits therefore need to be assessed in terms of present day realities and the lost opportunities to use resources in other ways. Comprehensive economic evaluations require two or more suitable alternatives to be evaluated. The most desirable design to demonstrate robust effects is the randomised controlled trial. However, the choice of design for an economic evaluation should be guided by the main questions being addressed.

Section 3

Special Issues

In this section, we discuss a number of important issues that are of a more technical nature. They apply to all types of evaluations – but there several specific matters that are of particular concern in outcome studies. Guidelines on data analysis are beyond the scope of this document and space limitations also prohibit discussion of the distinctions between qualitative and quantitative evaluation methods. Basic information on data collection and analysis and on other technical issues can be found in the WHO/UNDCP/EMCDDA workbooks. The issues are also covered in numerous basic and advanced textbooks on evaluation.

Reliability, validity and sensitivity of measures

Measures are observations or ratings of behaviour, attitudes, beliefs and the like. The principal requirements of a measuring device (such as a questionnaire) are that it should measure what it is supposed to measure (validity) and do so consistently (reliability). There are several ways to assess validity and reliability that allow different inferences to be made. The reliability of measures used in evaluation research can be assessed by examining consistency across time (test-retest-reliability) and consistency across observers (inter-rater reliability). Internal consistency is a useful measure of reliability for multi-item scales. This indicates the extent to which all items measure a single construct.

Several kinds of validity can be used for the measures used in programme evaluation. Face validity refers to whether the contents of specific questions or measures reflect what the indicator is supposed to measure. Assessments of validity of a measure are usually undertaken in three ways:

- coverage of all the salient features of the intended construct (content validity);
- by ensuring that the instrument does not include phenomena germane to other constructs (construct validity); and
- there is a strong positive correlation between scores on the instrument and other related instruments designed to measure the same construct (concurrent validity).

Ideally, a chosen measure must also be free from response bias and be non-reactive (insensitive) to extraneous factors (including physical setting, client expectation, staff behaviour and accountability pressures). Measures must also be sensitive to changes in the phenomena studied. A measure may be valid yet not sufficiently sensitive to measure subtle but important changes. Generally, the sensitivity of a measure increases with the range of possible scores it has. Thus a dichotomous response item – (e.g. yes or no) – will probably be less sensitive than a scale made up of several items.

Reliability of self-report data

Self-report is the most commonly used data source in evaluation studies. There is sometimes concern about the reliability and validity of such reports. Often these concerns are overstated. The validity of self-report information cannot be considered independently from the method used to obtain the report and will therefore vary across measurement contexts. The extent to which self-report is reliable, depends on a set of dynamic factors (see Babor, Brown and Del Boca, 1990). These include the following:

- the characteristics of the client (e.g. physical and psychological state; motivation and fatigue);
- contextual and task variables (e.g. context and setting for the data collection; confidentiality assurances; nature of the questions and answer format); and
- cognitive factors (e.g. comprehension and memory and retrieval processes)

In addition to client self-report there may be an opportunity to gather other collateral information including biological data. Biological measures such as urine, blood or even hair can be collected and analysed for drugs. There are advantages and disadvantages with each type of sample and expert advice is therefore recommended in the design, use and interpretation of biological data of this sort (Wolff et al., 1999).

Time periods for measures

An important decision concerns the time period over which outcomes will be assessed. For example, although a follow-up period may be six months in

duration, a decision needs to be reached about the time period over which clients will be asked to recall the key outcome measures. The same time period should be chosen for both baseline assessment and the follow-up interview. For most indicators, the recommended recall periods are 30 days and 90 days.

Assessing non-treatment factors that may affect outcome

When people with substance use disorders engage with treatment they are almost always exposed to other influences in their personal and social lives. These influences can exert either a positive or negative influence on the client's ability and willingness to change behaviour (see for example, Moos, Cronkite and Finney, 1994). For example, a client may have contact with friends who are problematic substance users or experience a sudden social problem (such as homelessness) and this may increase the risk of relapse. These influences are important and evaluators should try to measure the nature and extent of these factors when assessing treatment outcome.

There are also several non-treatment factors that may operate when assessing the impact of a treatment programme on a cohort of clients. Firstly, some clients may not really fully engage with the programme but simply 'mature-out' of drug use. This is particularly likely to be the case with younger substance users. Secondly, some clients who approach treatment services may have periods of relative stability or instability that occur independently of any treatment received. Thirdly, there is a rather confusing statistical artefact called 'regression to the mean' which can influence change scores for groups on measures that are not completely correlated and differ from a theoretical "true" score (a statistician should be consulted to review this issue). Fourthly, the data collection and evaluation process itself may influence clients. They may, for example, be motivated to respond in a positive manner to the follow-up interviewer. A follow-up interviewer who is in contact with clients on several occasions may also motivate them to change by asking questions and showing an interest in their situations. Various research designs can be used to control for the influence of non-treatment factors on client outcome. No one method is universally effective and no

single study is ever definitive. It is the convergence of research evidence from different study designs undertaken by different investigators in different countries that enables conclusions to be drawn about the role of treatment and non-treatment factors in patient outcomes.

Subject recruitment and follow-up

For outcome studies, an unbiased sample of subjects should be recruited at entry to a treatment programme. Baseline evaluation information (pre-treatment assessment) should be collected as early as possible in the intake/assessment process. At entry to treatment some clients may be unable to participate in a study due to physical and psychological health complications. Recall periods should be timed from entry into treatment in these situations (and the time from intake to baseline data collection noted).

An "intent-to-treat" follow-up strategy is also recommended for all outcome evaluations. This means that all clients recruited are sought for follow-up and not only those who completed treatment. There is a general consensus that at least 70% of subjects recruited to an outcome study should be followed-up for the results to be reliable. The timing of follow-up points is also important and published studies report on information obtained at many different points in time. Some evaluations report on client changes at the time of discharge or after a particular period of contact. Other evaluations have followed samples of clients for several years. Most common are reports of outcomes assessed at three, six, or 12 months after treatment has ended. The timing of follow-up and the intervals considered will have a significant impact on the results and conclusions drawn from a study. Short-term follow-up studies are likely to show better results than longer term ones. This is because between 60% and 80% of 'relapses' to substance use occurs in the first three to four months following discharge. Results will depend on the type of outcome being assessed. For example, changes in drinking and drug use may not be stable, but long-term overall improvements in family functioning may occur. Overall, the selection of follow-up periods and intervals must be consistent with the objectives of the evaluation if it is to produce useful results.

Sampling and sample size

A sampling strategy refers to the methods used to select the representative “units” (such as treatment agencies or treatment participants) to be included in a study. A sampling procedure that assures that each element in the population has an equal chance of being selected is referred to as “simple random sampling”. Sometimes a systematic random sampling method can be used where a starting point is chosen (e.g. sampling every 20th school in a population or every 20th class member). An alternative to simple random sampling is stratified random sampling. This method is normally used where:

- It is not possible to generate a list of all the members of a population;
- where the members of a population are too geographically dispersed to be contacted; and
- where the population contains important subgroups that are either of variable sizes and/or geographically dispersed.

In contrast to simple random sampling and stratified sampling, (where single subjects are selected from the population), in “cluster sampling” the subjects are selected in groups or clusters. Cluster sampling is commonly used when the population is very geographically dispersed and may already form natural groups.

When considering the size of a sample that is required, this is to some extent determined by the nature of the study, its objectives and the measures to be used. In outcome studies that aim at comparing outcomes for two or more groups of clients, the sample size depends on the magnitude of the difference on a specified outcome measure that is required to detect change between the groups. A statistician can assist in calculating the required sample size based on the following information:

- the degree of confidence one wants in the results obtained;
- the relative sizes of groups to be compared;
- the expected frequency of the behaviour in one group; and
- the magnitude of the difference that is to be detected between the groups to be compared

When determining the number of people to be followed-up in an outcome evaluation, allowance should be made for clients who cannot be contacted and for whom outcome information will be missing. It is reasonable to expect that up to 30% of cases chosen for follow-up cannot be traced. The sample selected for follow-up should therefore be increased by 30%.

Sample size is also important in needs assessment surveys. In general, the required sample size in such cases is smaller than commonly believed. However, it depends on the number of subgroups that may be of interest and the expected accuracy of results. Consultation with a statistician is recommended for the design and analysis of studies involving survey samples.

Statistical testing and confidence intervals

Almost all outcome evaluations involve an element of statistical testing. In a randomised controlled trial, for example, researchers will ask about the “statistical significance” of differences in response measures between the control and experimental group(s) at intake and follow-up and between these two points. The conventional way of tackling this question is to use statistical hypothesis testing. This usually starts with the assumption that there is no difference between groups (the so-called null hypothesis). A statistical test is then used to see if any observed difference can be considered genuine (at the population level) or due to chance. It is quite common for researchers to present the values of a statistical test and to indicate the probability that the observed difference is due to chance. However, this only indirectly tells the reader how much confidence to have in the results. It is therefore recommended that evaluation studies also report the magnitude of differences over time and between groups in absolute value where the measure is meaningful (e.g. a mortality rate) and in standard deviation units when the outcome is measured in an essentially arbitrary way (e.g. a score on an attitude scale). The upper and lower limits of a confidence interval around this estimate should also be provided (the 95% bound is commonly used).

Section 4

Reporting Results

In this section, we consider the reporting of the results of an evaluation and how to apply the results. Each of the types of evaluation discussed in this document will lead to a report with some unique features. As an example, only reports for an outcome evaluation will be considered here. Box 10 shows one possible structure for an outcome evaluation report. This has been informed by the professional consensus statement published in the Journal of the American Medical Association on key information to report from clinical trials (see Begg, et al., 1996). The box presents a more generic overview of the sorts of material to include in a formal report. An executive summary of the evaluation should be prepared and this is a valuable means of describing the key findings and significance of a study to a wide audience.

Increasing the impact of evaluation

It is unfortunately the case that evaluations of treatment for substance abuse do not always have an impact on decisions about what treatment programmes will be supported and how treatment will be delivered (Miller, 1987; Ogborne, 1988). There are many reasons for this including the vested interests of some treatment providers, the poor quality of some evaluations and, in some instances, the lack of clear-cut results. However, the impact of evaluation may also be limited by poor communication between evaluators and decision-makers.

The importance of involving key decision makers in the earliest discussions of evaluation will increase the chances that evaluations will generate timely and useful information. Good communication goes well beyond simply providing funding bodies copies of evaluation reports. These reports are often lengthy and sometimes quite technical and may fail to hold readers' attention. Also reports in themselves may not be sufficient to convince decision makers that particular programmes should be supported. This is especially so when evaluations suggest that new programmes would have to be funded at the expense of existing programmes or that existing programs need to be radically altered. In these cases there is likely to be opposition to change from those with a vested interest in the status quo. Resistance to change can also occur when there is a

lack of skills and resources in existing delivery systems or lack of understanding of the need for change and a lack of appreciation of the benefits of change. In some cases prevailing belief and value systems can impede the adoption of new treatments and in others cases resistance to change reflects fears for loss of jobs or status among service providers.

Several strategies have been proposed for increasing the impact of evaluation and there is also a large literature on the diffusion of innovation that has some important lessons for those seeking to promote evaluation and the adoption of treatments for substance abuse (Davis, 1978; Eveland, 1986; Backer, 1991). It is recommended that, where possible, priority should be given to ways that clearly demonstrate the advantages of evidence-based treatments. This might include the implementation of special demonstration projects to be visited by those interested in using the same methods. The demonstration programmes can also provide on-site training opportunities. As well it would be useful to develop clear guidelines or "how to" manuals for using the new methods.

Change agents – people who aim to bridge the gap between knowledge and practice by actively promoting evidence – based practices – have proved to be helpful in many situations. These should be people who have, or are able to earn the respect of those they seek to influence and who have a good understanding of new methods. They need strong interpersonal skills and the ability to find creative ways to promote change in different contexts.

Structure and content of a formal report
from an outcome evaluation

Box 10

HEADING	SUBHEADING	DESCRIPTION
Title		Description of study and the research team and source of any funding support
Abstract		Structured format (e.g. aims and objectives; study design; treatment(s) studied and their settings; participants; measurements; methods and procedures; findings; conclusions)
Introduction		Background to the study and summary of relevant previous research State prospectively defined hypothesis, clinical objectives, and planned subgroup or covariate analysis
Methods	Protocol	Subject population studied Participant inclusion/exclusion criteria Interventions studied (modality; intensity and duration) Primary and secondary outcome measure(s) Describe how sample size was calculated (power calculation) Plan for statistical analysis (main and subsidiary analyses)
	Procedure	Subject recruitment methods Unit of randomisation (e.g. individual, cluster, geographic) – if used Randomisation methods Number and timing of follow-up points Subject tracking procedures and interview methods (e.g. face-to-face; telephone; self-completion)
Results	Descriptive	Study profile (figure) summarising number of participants through the study Numbers of subjects in treatment over time and numbers interviewed and lost to the study at follow-up Baseline subject characteristics
	Inferential	Effect of programme on primary and secondary outcome measures
Discussion/ commentary		Summary of main findings Extent of bias and/or design weaknesses and study limitations Implications of the findings Recommendations

Summary and Conclusions

There is still much to be learned about the most effective, efficient and humane ways to help people with substance use disorders and to reduce the associated harms to users themselves and to others. There are still many treatments that have not been rigorously evaluated and some treatments, found effective in some context or cultures, have not also been tested in others. The emergence of new drugs, new types of drug users and new patterns of use may also require the development of new treatments that need to be evaluated. Evaluation should therefore be a feature of treatment systems and treatment policy. Evaluation will help to (a) identify treatment needs (b) plan needs-based interventions, (c) show if these interventions are consistent with needs and plans and (d) show if these interventions are effective and efficient.

These guidelines summarise the main features of evaluation and we hope that they will motivate the reader to support and champion the evaluation of substance abuse treatment within their countries and communities. It is also hoped that readers will be motivated to learn more about evaluation and be interested in consulting the WHO/UNDCP/EMCDDA workbooks.

The guidelines have not attempted to conceal the fact that some types of evaluation are quite complex and require special skills. This would be a disservice to the cause of evaluation and to the many people who have worked hard to develop the evaluation methodologies that are helping to improve treatment in some contexts. We have, however, also attempted to show that evaluation does not have to be daunting. All readers are strongly encouraged to consider supporting or undertaking basic evaluations even if more complex alternatives are not at present feasible. These initial efforts can later be extended as more skills are acquired and more resources become available. Evaluations and related activities that do not require highly special skills or extensive resources are:

- Need assessments that use existing quantitative data and/or key informant interviews;
- Use of logic models to plan new services and to clarify the assumptions and objectives of existing programmes or systems;
- Development and implementation of basic, routine or periodic monitoring systems to show (1) what services are being provided, to what clients and at what cost (2) amount of treatment received by different clients (3) modes of discharge (completed program, drop out, discharged for cause etc); and
- Periodic use of simple self-completion questionnaires or confidential interviews to assess clients satisfaction

Examples of evaluations of these types are given in the WHO/UNDCP/EMCDDA workbooks and in an evaluation casebook published by the Addiction Research Foundation of Ontario (Graham et al., 1994).

References

- Adams E.H, Kopstein A.N The non-medical use of *prescription* drugs in the United States. In: Cooper J.R., Czechowicz D.J, Molinari S.P., Petersen R.C. (Eds.) *Impact of Prescription Drug Diversion Control Systems on Medical Practice and Patient Care*. NIDA Research Monograph 131. Rockville MD: National Institute on Drug Abuse, 1993.
- Babor, T.F., Brown, J., & Del Boca, F.K. Validity of self-reports in applied research on addictive behaviours: fact or fiction? *Behavioral Assessment*, 1990,12: 5-31.
- Backer, T.E. *Drug abuse technology transfer*. NIDA monograph serious # 1555; Rockville, MD; National Institute on Drug Abuse, 1991.
- Begg, C., Mildred, C., Eastwood, S., Horton, R., Moher, D., Olkin, I., Pitkin, R., Rennie, D., Schulz, K.F., Simel, D. & Stroup, D.F. Improving the quality of reporting of randomised controlled trials. *Journal of the American Medical Association*, 1996, 276:637-639.
- Bradshaw, J.S. A taxonomy of social need. In: McLachlan, G. (Ed.) *Problems and Progress in Medical Care: essays on current research*. 7th series. Oxford: Oxford University Press, 1972.
- Centre for Substance Abuse Treatment *The Costs and Benefits of Substance Abuse Treatment*. Available online at [<http://neds.calib.com/products/pdfs/costben.pdf>], 1999.
- Chen, H.-T. *Theory-driven evaluations*. Newbury Park: Sage, 1990.
- Davis, H.A. Management of change and innovation in mental health services. *Hospital and Community Psychiatry*, 1978, 29: 649-658.
- Dennis, M.L. Ethical and practical randomised field experiments. In: Wholey, J.S., Hatry, H.P.& Newcomer, K.E. (Eds.) *Handbook of Practical Program Evaluation*. San Francisco: Jossey Bass, 1994.
- DeWit, D.J., & Rush, B.R. Assessing the need for substance abuse services: A critical review of needs assessment models. *Evaluation and Program Planning*, 1996, 19: 41-64.
- Drummond, M.F., Stoddart, G.L. & Torrance, G.W. *Methods for the economic evaluation of health care programmes*. Oxford: Oxford University Press, 1987.
- Eveland, J.D. Diffusion, technology transfer and innovation. Thinking and talking about change. *Knowledge: Creation, Diffusion and Utilization*, 1986, 8: 303-322.
- Graham, K, Woo, G.A. Smythe, C. and others *The Evaluation Casebook: Using Evaluation to Enhance Program Quality in Addictions*. Toronto: Addiction Research Foundation. (Now part of the Centre for Addiction and Mental Health), 1994.
- Holder, H., Longabaugh, Miller, Rubonis. The cost effectiveness of treatment for alcoholism first approximation. *Journal of Studies on Alcohol*, 1991, 52: 517-525.
- Holder, H.D. & Blose, J.O. The reduction of health care costs associated with alcoholism treatment: A 14 year longitudinal study. *Journal of Studies on Alcohol*, 1992, 53: 293-302.
- Landry, M. *Overview of Addiction Treatment Effectiveness*, Office of Applied Studies. 116.Rockville, MD: SAMHSA, 1995.
- Lebour, D.L. Similarities and differences between mental health and health care evaluations studies assessing consumer satisfaction. *Evaluation and Programme Planning*, 1983, 2:197-245.
- Marsden, J., Strang, J. with Lavoie, D., Abdulrahim, D., Hickman, M & Scott, S. Epidemiologically-based needs assessment: Drug misuse. In: Stevens, A. & Raftery, J. (Eds.) *Health Care Needs Assessment*. Oxford: Radcliffe, forthcoming.
- McLellan, A.T., Hagan, T.A., Levine, M., Gould, F., Meyers, K., Bencivengo, M., and Durell, J. Supplemental social services improve outcomes in public addiction treatment. *Addiction*, 1998,93: 1489-1499.
- Miller, W.E. and Hester, R.K. *The effectiveness of treatment techniques: what works and what doesn't*, in *Treating addictive behaviours: processes of change*, W.E. Miller and N. Heather, Editors, Plenum Press: New York, 1986.

Miller, W.R., *Behavioral alcohol treatment research advances: barriers to utilization*. Advances in Behavior Research and Therapy, 1987, 9: 145-164.

Moos, R.H., Finney, J.W. & Cronkite, R.C. Alcoholism treatment. Context, Process and Outcome. New York: Oxford University, 1990.

National Institute on Drug Abuse, *Principles of Drug Addiction Treatment: A research-based guide*, National Institutes of Health: Washington D.C., 1999.

Ogborne, A.C. Bridging the gap between the two cultures of alcoholism research and treatment. *British Journal of Addiction*, 1988, 83: 729-733.

Roberts, G. and Ogborne, A.C., *Best Practices: Substance Abuse Treatment and Rehabilitation*, Office of Alcohol, Drugs and Dependency Issues, Health Canada: Ottawa, 1999.

Rossi, P.H., & Freeman, H.E. *Evaluation: A systematic approach* (2nd Ed.). Beverly Hills, California: Sage Publications, 1982.

Rush, B.R., & Ogborne, A.C. Program logic models: Expanding their role and structure for planning and evaluation. *Canadian Journal of Program Evaluation*, 1991, 6(2): 95-106.

Rydell, C.P. and S.S. Everingham, *Controlling Cocaine: Supply versus Demand Programs*, Rand: Santa Monica, CA, 1994.

Single, E., Robson, L., Xie, X., Rehm, J. *The costs of substance abuse in Canada*. Ottawa. The Canadian Centre on Substance Abuse, 1996.

Stevens, A. & Raftery, J. *Health Care Needs Assessment. The Epidemiologically based needs assessment reviews*. Oxford: Radcliffe Medical Press, 1994.

Tims, F. *Drug Abuse Treatment Effectiveness and Cost-Effectiveness*, Rockville, Maryland: National Institute on Drug Abuse, 1995.

United Nations International Drug Control Programme *World Drug Report*. Oxford University Press, New York, 1997.

Wholey, J.S. Assessing the feasibility and likely usefulness of evaluation. In: J.S. Wholey, J.S., Hatry,

H.P. & Newcomer, K.E.(Eds.) *Handbook of Practical Program Evaluation*. San Francisco: Jossey Bass, 1994.

Wolff, K., Farrell, M., Marsden, J., Monteiro, M.G., Ali, R., Welch, S. & Strang, J. A Review of Biological Indicators of illicit drug use: practical considerations and clinical usefulness. *Addiction*, 1999, 94: 1279-1298.

World Health Organization, United Nations Drug Control Program and European Monitoring Centre on Drugs and Drug Addiction *Evaluation of Psychoactive Substance Use Disorder Treatment. Facilitators Manual and Workbooks 1-8*. WHO Geneva, 2000.

WHO Expert Committee on Drug Dependence, *30th Report*, World Health Organization: Geneva, 1996.



WHO



UNDCP



EMCDDA

E.M.C.D.D.A.
European Monitoring Centre
for Drugs and Drug Addiction