Workbook 1

Planning Evaluations

WHO
World Health Organization

UNDCP
United Nations International Drug Control Programme

EMCDDA
European Monitoring Center on Drugs and Drug Addiction

Acknowledgements

The World Health Organization gratefully acknowledges the contributions of the numerous individuals involved in the preparation of this workbook series, including the experts who provided useful comments throughout its preparation for the Substance Abuse Department, directed by Dr. Mary Jansen. Financial assistance was provided by UNDCP/EMCDDA/ Swiss Federal Office of Public Health. Brian Rush (Canada) and Cam Wild (Canada) wrote the original text for this workbook, and Brian Rush (Canada) edited the workbook series in earlier stages. JoAnne Epping-Jordan (Switzerland) wrote further text modifications and edited the workbook series in later stages. Munira Lalji (WHO, Substance Abuse Department) and Jennifer Hillebrand (WHO, Substance Abuse Department) also edited the workbook series in later stages. The first case example for this workbook was written by Cindy Smythe (Canada) and edited by Ginette Goulet (Canada). The second case study was written by John Marsden, Michael Gossop, Duncan Stewart, David Best, Michael Farrell, Petra Lehmann, Carolyn Edwards and John Strang; and edited by Maristela Monteiro. Maristela Monteiro (WHO, Substance Abuse Department) also provided editorial input throughout the development of this workbook.

Some of the material in this workbook was adapted from a NIDA publication entitled “How Good is Your Drug Abuse Treatment Program? A Guide to Evaluation.” Contributions drawing from this report are gratefully acknowledged.

Table of contents

Overview of workbook series

This workbook is part of a series intended to educate programme planners, managers, staff and other decision-makers about the evaluation of services and systems for the treatment of psychoactive substance use disorders. The objective of this series is to enhance their capacity for carrying out evaluation activities. The broader goal of the workbooks is to enhance treatment efficiency and cost-effectiveness using the information that comes from these evaluation activities.

This workbook (Workbook 1) describes step-by-step methods for planning evaluations. These steps span from deciding who will be involved in the evaluation, to defining your research questions and determining your data collection methods.

Introductory Workbook
Framework Workbook

Foundation Workbooks
Workbook 1: Planning Evaluations
Workbook 2: Implementing Evaluations

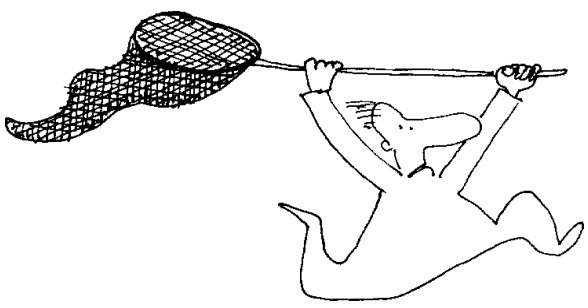
Specialised Workbooks
Workbook 3: Needs Assessment Evaluations
Workbook 4: Process Evaluations
Workbook 5: Cost Evaluations
Workbook 6: Client Satisfaction Evaluations
Workbook 7: Outcome Evaluations
Workbook 8: Economic Evaluations

Introduction

How does one ensure that resources for evaluation are not wasted and that the results are useful? The answer is to develop a detailed plan before proceeding with an evaluation.

In most countries, resources available for the treatment of PSU disorders are scarce. Competition for these resources has increased the need for the evaluation of treatment programmes. Despite the critical need for evaluation, the evaluation process itself takes time and resources. It is important, therefore, that resources available for evaluation are used efficiently, and that results are useful for making decisions about programmes or treatment systems.

How does one ensure that resources for evaluation are not wasted and that the results are useful? The answer is to develop a detailed plan before proceeding with an evaluation. Developing a plan involves discussions with various people who have interest in the programme. In these discussions, specific questions are identified that will be answered by information collected in the evaluation, and be of value to one or more groups in decision-making processes. This workbook discusses the steps involved in planning for evaluation and offers practical strategies to help guide this process. The more carefully you plan at this stage, the more benefits you will reap from your evaluation efforts.



The eight steps of planning evaluations

As described in the framework manual, you need to accomplish eight steps during this evaluation-planning phase:

1. Decide who will be involved in the evaluation.
2. Assess your evaluation resources.
3. Describe your programme for evaluation.
4. Identify and prioritise the evaluation needs.
5. Define your evaluation questions.
6. Determine your evaluation measures.
7. Determine your evaluation design.
8. Ensure that your evaluation resources are sufficient. If not, return to Step 4.

Each of these steps is presented in the following pages. The discussion of these steps will include a fictional example of a group of people planning an evaluating of their opiate detoxification programme and some exercises for you to follow. The fictional example will give you an idea about the key steps undertaken by a group of people evaluating their opiate detoxification programme. The example will show you how the presented steps can be transferred into action. Keep in mind that the example will show you how the steps can be accomplished in an ideal way. The exercises will help you to learn and apply the material to your own situation.

Use the foundation and specialised workbooks together, to help you make the most of the information that is presented.

Workbooks 1 and 2 provide a solid foundation of general information about conducting evaluations, whereas the specialised workbooks (Workbooks 3 through 8) present detailed information for different types of evaluation. If you already know what type of evaluation you are going to conduct, you should consult the workbook that is applicable. Use the foundation and specialised workbooks **together**, to help you make the most of the information that is presented. If you do not know what type of evaluation you are going to conduct, wait until you have developed your evaluation questions (Step 5) to consult a specialised workbook.

If you are conducting a...	Then you should review...
Needs Assessment Evaluation	Workbook 3
Process Evaluation	Workbook 4
Cost Evaluation	Workbook 5
Client Satisfaction Evaluation	Workbook 6
Outcome Evaluation	Workbook 7
Economic Evaluation	Workbook 8

Step 1

Decide who will be involved in the evaluation

An important first step is to identify and meet your evaluation planning "partners," if you have not already done so. Depending on your situation, your partners may include:

- therapists or clinicians
- programme administrators or managers
- researchers
- government representatives
- patients interested in participating

All partners should be closely involved in the evaluation planning. Each partner has unique experiences and perspectives that can contribute to the group's knowledge base and strengthen the overall evaluation. Other benefits include:

- bringing multiple perspectives to the planning
- strengthening everyone's commitment to use the findings
- adding credibility to the process

Into action

Our fictional example concerns a group of people who plan and implement an evaluation in their 25-bedded heroin detoxification service. This in-patient detoxification programme is designed for clients with opiate dependence. Clients stay at the service for 21 days. The needs of each individual are met through individual and group therapy during a three week structured programme that includes relapse prevention, stress management, health and AIDS education programmes. The partners who are involved in planning and implementing the evaluation are:

Sue R.: Psychiatrist and sees patients regularly, she knows what kind of problems experience clients during their detoxification and has a good sense of what kinds of evaluation will/will not work in the programme setting, get along well with others.

Adam S.: Scientific researcher: Knows about statistics, computers and evaluation, can do data analysis, can provide about the "scientific" quality of our evaluation plan, good sense of humour.

Chris C.: Drug worker: Sees clients and does pre-treatment assessments, individual counselling and group session; can co-ordinate and conduct interviews with clients.

Maria M.: Programme secretary: Very organised. She is responsible for entering intake and termination forms of clients into the central database.

Other people who will get involved in their evaluation are representatives of funding bodies, other drug workers and clients attending the 21-day treatment service.

It's your turn (1 A)

1 First, list the names of your partners who are involved in your evaluation planning project. Then, list the unique talents that each person brings to the group.

2 Discuss how you can take advantage of each other's talents to maximise the evaluation planning process.

3 Working individually, list three hopes about the evaluation project, and then list three fears. When everyone has finished, share and discuss the lists. Make plans for how to attain your hopes and overcome your fears.

In any given situation, many other people can be involved potentially in evaluation. These might include:

- other therapists or clinicians
- supervisors or programme managers
- senior programme administrators or Board members
- other researchers
- representatives of government or other funding bodies

You and your partners might want to develop a larger evaluation committee that has representation from these groups, including patients/clients. Another idea is to collect information from them during this planning phase to identify the kinds of questions they would like to have answered, and the decisions they anticipate having to make.

It's your turn (1 B)

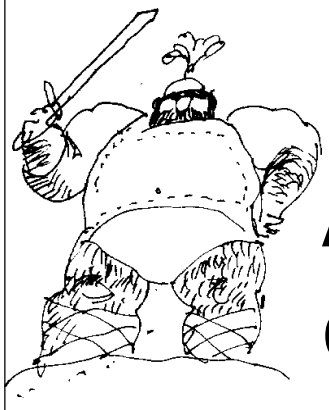
1 List ALL the groups of people involved in your treatment programme, or network of programmes. The list will be unique to your situation, but may include:

– patients/clients
– family members
– senior managers
– Board members
– staff (therapists/clinicians)
– researchers
– funder(s)
– representatives of other programmes or service systems in your community
– people living in your community

2 Decide how to best involve representatives from each of the different groups. Very likely, some groups will be involved closely throughout treatment evaluation, whereas others may participate only in a limited way. You might want to develop an evaluation committee that has representation from these groups, including patients/clients. Or, you could collect information from them during this planning phase to identify the kinds of questions they would like to have answered, and the decisions they anticipate having to make.

Write down how you will involve each of these groups.

Step 2



Assess your evaluation resources

Evaluation requires resources. Specifically, evaluation requires financial/material resources, expertise resources, and time resources. It is important for you to evaluate your level of "evaluation readiness" along each of these dimensions, and then to balance your resources with the type of project that you undertake.

Remember that high quality evaluations CAN be conducted with very few resources. The key is to evaluate your situation realistically, and then choose a project that is practical with the resources that you can devote to it.

In this workbook's Timmins Detoxification Service evaluation case example, planners evaluated their resources and decided subsequently to hire a programme evaluation consultant. While this type of added expense may not be possible for all evaluations, it is noteworthy that it was helpful in this evaluation.

Into action

The fictional evaluation team discussed their resources along the following dimensions:

Financial/Material Resources

Sue R. informs the others, that no internal funding can be devoted to the evaluation project, however the local health purchaser agreed to fund their project with 300,000 US Dollars. This money enables the service to employ one person (Adam S.) part-time for the evaluation project during the planing phase of the evaluation and after the data collection, when the data is to be analysed.

One single office, provided with a computer will be used for data storage, entry and analysis. Other facilities, such as photocopier and cabinet files are available in the office of the programme secretary.

Expertise Resources

Sue R. attended a one-day workshop about evaluation and will communicate her knowledge to other members of the team. Chris C. has been involved in other research studies. He is familiar with data collection, e.g. conducting standardised interviews. Adam S. will assist in planing and analysing the data for the evaluation.

As part of her duties, Maria M. has been entering intake and termination forms for all clients for the last two years. She is familiar with data entry and data entry software.

Time Resources

Existing staff agreed to assist with the data collection. They will have to devote an additional 2 hours per client for the collection of baseline data and follow up. Adam S. will work part time (20 hours/per week) for approximately one month and approximately 3 month after the data collection. Sue R. will devote an additional 3 hours per week for the project. Her time for the project will depend on the demands of the evaluation. Maria M. will not have to devote more time for the project, because entry of records are already part of her duties at the service.

It's your turn

Evaluate your resources along the following dimensions:

Financial/material resources

1 Is there internal funding that can be devoted to this evaluation project? If yes, what amount?

2 Are there external funding agencies that can provide funding for this project? If yes, what amount?

3 Can you afford to hire evaluation staff, or will evaluation be conducted by existing staff members (e.g., treatment agency staff)?

4 Is there a computer (and software) available for data entry and data analysis? (Note: computers are not necessary for all evaluations)

5 Is there a photocopier available for copying survey materials, etc.?

Expertise resources

1 Has anyone involved on the project conducted treatment evaluations before? If yes, in what capacity?

2 Has anyone involved on the project worked on a computer before, specifically doing data entry or data analysis?

3 Do you have access to expert "consultants," who can provide advice on conducting the evaluation and/or analysing the data?

Time resources

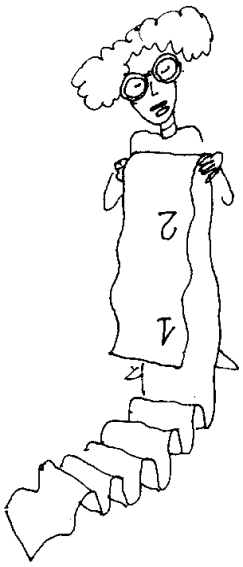
1 How much time will each person have available to devote to the evaluation project each week?

2 If you will rely on existing staff to collect data and/or analyse data, how much time will they have to devote to this each week?

Your answers to these questions should help determine the type of evaluation to undertake. For example, if you do not have access to or expertise with computers, you probably will want to collect data that can be tallied and evaluated by hand. Later, you will have a chance to make sure that the project you have chosen fits well with your resources. But for now, it is sufficient just to know the answers to these questions.

Step 3

Describe your programme for evaluation



Every health or social service can be described according to the following structure:

- resources that are organised for a common purpose
- activities that are planned and undertaken
- immediate outcomes that are to be achieved
- longer term outcomes that are intended
- other positive or negative consequences or side-effects

Into action

In our fictional example, the **resources** devoted to operating a detoxification unit include a 25-bed facility, with sixteen staff and an annual budget of $480,000 (U.S.). Programme **activities** include withdrawal management, medical screening, referral, individual and group therapy, health and AIDS education and general support to help to stabilise the person, physically and emotionally. The **immediate outcomes** are safe withdrawal from opiates and increased motivation for seeking further treatment. **Longer-term outcomes** are increased participation in treatment, decreased PSU and high-risk behaviour, improved health and social well being, decreased use of hospital inpatient facilities and reduction in hospital costs. Reduced criminality rates and unemployment rates may be a positive, unanticipated **side effect** of introducing an opiate detoxification unit. The information about the treatment unit can be represented graphically in the following manner:

Programme Resources

- $ 480,000 annually
- 16 staff members

Programme Activities

- withdrawal management
- individual and group therapy
- motivational interviewing
- referral to long term treatment
- health and AIDS education

Immediate Programme Outcomes

- safe withdrawal
- increase knowledge about AIDS
- increased motivation for further treatment

Longer -Term Programme Outcomes

- increased participation in treatment
- decreased use of illicit substances
- decrease in high-risk behaviour
- increased health

➔ **(Side Effects)**

It is useful to outline your own programme in this manner, so that you can see the underlying logic or rationale of your programme, or the links between your programme activities with shorter and longer-term outcomes. Developing a "programme logic model" like this will help you in planning your evaluation. Why? In addition to challenging you to defend, and perhaps refine the logical basis underlying your programme, the process of stating your objectives will make you aware of the need for clarity and specificity in these objectives. The clearer you make your objectives, the easier the next phase of evaluation planning will be when you decide how to measure your success in achieving those objectives.

In addition, experience has shown that having an agreed upon logic model among key partners serves as a strong foundation for planning and prioritising your evaluation activities. It is one of the most concrete steps you can take towards a successful evaluation plan.

This workbook's case report of the Timmins Detoxification evaluation describes how all partners were involved in creating their programme logic model. Their joint effort helped identify specific evaluation questions.

It's your turn (3 A)

Using the example outlined earlier as a guide, describe your own treatment programme. To help identify your programme activities, ask yourself, "what are the main things we do to accomplish our goals?"

- **Programme Resources**
- **Programme Activities**
- **Immediate Programme Outcomes**
- **Longer - Term Programme Outcomes**
- **(Side Effects)**

Sometimes, it is helpful to describe your programme further by distinguishing between process objectives and outcome objectives.

- **Process objectives:** planned activities or services
- **Outcome objectives:** the expected changes that will occur

Into action

The fictional heroin detoxification programme has the following examples of process and outcome objectives:

Process objectives

- **to administer** standardised assessment instruments
- **to provide** individual and group therapy
- **to assist** the clients in entering long-term treatment
- **to provide** information on consequences of PSU and AIDS

Example of outcome objective

- **to help** the client to withdraw from opiates
- **to increase** health
- **to reduce** PSU
- **to increase** life-skill and knowledge about AIDS and consequences of PSU
- **to decrease** high risk behaviour

You can usually distinguish process and outcome objectives by the verb contained in the sentence. Process objectives use verbs that reflect activities while outcome objectives use verbs that reflect changing something. Examples of process and outcome objectives (used in an overall programme logic model) are shown bellow.

Inpatient opiate detoxification service - programme logic model

Main components

Intake & Assessment
- to gather basic information on clients
- to administer standardised assessment instruments
- to develop a withdrawal management

Counselling
- to provide individual counselling
- to provide group counselling (e.g. relapse prevention)
- to implement clients' treatment plans

Education
- to provide information on consequences of substance use
- to teach AIDS education

Stabilisation
- to provide crisis intervention
- to provide safe withdrawal from opiates
- to assist clients in finding long term treatment
- to teach basic life skills

Short-term outcomes

- to increase clients motivation for further treatment
- to increase health
- to withdraw clients from opiates

- to increase clients knowledge of AIDS and consequences of substance use

- to stabilise clients and their situation
- to improve life skills
- to refer clients to long term treatment

Long-term outcomes

- to increase clients participation in treatment

- to reduce PSU

- to reduce high-risk behaviour

- to improve the general well-being of the client

It's your turn (3 B)

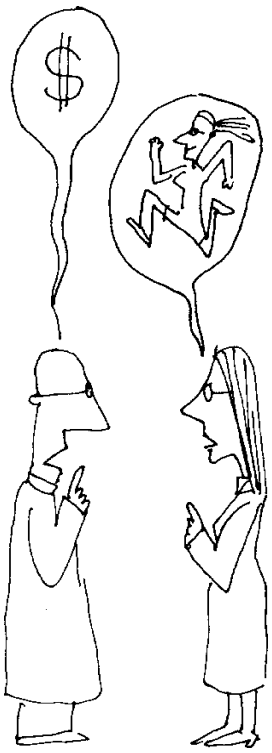
For each of the programme activities that you listed in the last exercise, now list the various process objectives that go with that component. There is usually more than one process objective for each activity.

Do the same with your outcome objectives. Some may be more attainable in the short-term, while others are long-term and only occur after other preceding objectives have been met. There should be an ordering or rationale linking the outcome objectives. Draw these logical connections.

Last, draw a line from each set of process objectives to the short-term outcomes expected as a result of successfully implementing these activities and services. Here you are drawing the causal linkages that not only show what goes on in your programme, but also why it should work.

Main components (e.g. assessment, detoxification, follow-up) **Implementation objectives** (e.g. to determine correct diagnosis, to provide substance resistance skills, to monitor health status.)			
Planned short-term outcome objectives (e.g., to increase motivation for further treatment, to decrease the likelihood of relapse).			
Planned long-term outcome objectives (e.g. to decrease substance abuse, to improve quality of life)			

Step 4



Identify and prioritise evaluation needs

Different groups will have different ideas about what to evaluate. For example, managers or administrators may be most interested in costs and office efficiency, while clinicians may be interested in the benefits of a particular new treatment. It is important for you to understand the views and priorities of your partners, and all key groups, then to come to a decision about the overall needs for evaluation.

Into action

The evaluation team of the opiate treatment unit discussed their ideas of what is important to evaluate within their programme and why. They agreed upon that they are interested in the effectiveness of their AIDS education programme which was incorporated into the treatment programme a few month ago. They are further interested if the programme accomplished what was expected to accomplish in terms of completion rates, referrals to long-term treatment, number of referrals to the service and possible changes in clients' characteristics attending the service. The latter would provide information if the service needs to adjust his management to changes in client characteristics. Information about completion rates, number of referrals to the service and from the service to long term treatment could provide evidence to the sponsors of the treatment that the funds are being used as expected.

It's your turn

1 Working individually, write down your ideas for what is important to evaluate within your programme or treatment network, and why. Next, brainstorm a list of evaluation questions for which you would like to get feedback.

2 Discuss your ideas with your partners. Take time to understand each partner's perspective. Discuss similarities and differences in your ideas.

3 Get as much input as you can from other key people. You can do this with face-to-face discussions, telephone interviews or brief self-completed questionnaires.

– patients/clients

– family members

– senior managers

– Board members

– Staff (therapists/clinicians)

– volunteers

– funder(s)

– representatives of other programmes or service systems

– people living in the community

These initial exercises will help you to see the range of possibilities for your evaluation. They also may help you to see some common trends in ideas.

Step 5

Define your evaluation questions

At this stage in your evaluation planning, you have clearly identified your partners in the evaluation process and involved others in the planning in a meaningful way. You also have evaluated your resources, and started to identify and prioritise your evaluation needs. Finally, you have defined and assessed the structure and logic of the programme, or network of programmes, to be evaluated. It is now necessary to narrow down the precise questions to be answered.

Questions can be addressed about any part of the programme structure, logic or process. The programme logic model is a useful tool for helping to generate these questions and narrowing the focus of the evaluation. This section provides a format for going from your logic model to this more detailed planning of the evaluation.

It is essential that the expected user(s) of the evaluation be involved in the process of identifying evaluation questions. In working through the completion of the next exercises, you may end up with more evaluation questions than can be answered with the time and resources available. If this happens, it is essential that the user(s) of the evaluation information be involved in discussions to narrow the questions to a meaningful number.

Into action

The following are some examples of possible evaluation questions generated by the evaluation team at the heroin detoxification service.

Questions on client characteristics

- Has the number of referrals increased from the previous year?
- Has the number of self-discharges decrease from the previous year?
- Are the characteristics of clients similar to the type of clients reflected in the programme objectives and similar to the characteristics of clients in the previous year?

Questions on clients satisfaction

- Did clients find the types of content in the programme useful?
- How satisfied were clients who attended the detoxification?

Questions on programme delivery

- How many hours of counselling were provided to each client individually and in-groups?
- What proportion of these hours involved direct contact with the client compared to the proportion of hours involving contact with others about the client?

Questions on outcome

- Did attitudes in favour of HIV low-risk practices improve among clients?
- Did clients awareness and knowledge about HIV high-risk situations increase?
- Did clients who completed the programme decrease their PSU?
- Did counselling increase the proportion of clients who accepted long-term treatment?

Remember, these are just EXAMPLES of evaluation questions. Your question(s) may be different.

The case example of the Timmins Detoxification service evaluation, located at the end of this workbook, identified questions from a variety of domains: service awareness, assessment, crisis intervention, withdrawal management, follow up care, and volunteers. In this case, domain represented a different function of the detoxification programme.

It's your turn

1 Using the programme logic model that you created during Step 4 as a guide, write down possible evaluation questions below. At this stage, write down as many questions as possible. Ensure that each partner has an opportunity to contribute questions.

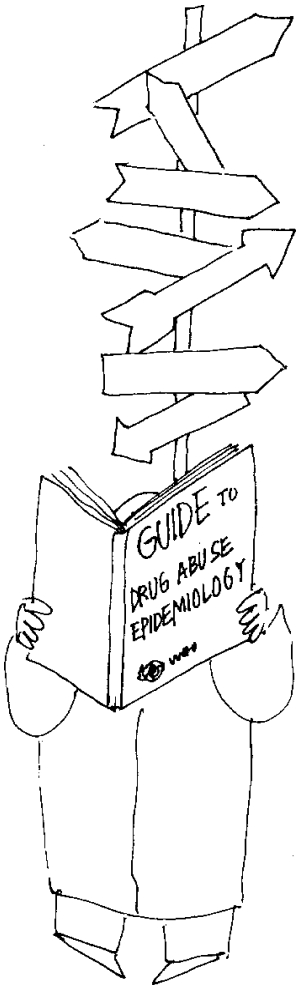
2 For each question that you generated above, list advantages and/or disadvantages of studying THIS evaluation question. Think about your resources (Step 2) while considering advantages and disadvantages.

3 Discuss the options with your partners. As a group, choose the best evaluation question(s).

4 Discuss your choices with the expected user(s) of the evaluation results. Determine if they agree with your choices. Change your questions as appropriate.

Step 6

Determine your research measures



Why should you care about research measures?

What happens if you don't take research measurement issues seriously during programme evaluation? Simply put, any claims you make about the programme will always be open to criticism. Someone could always say: "This sounds nice, but how do you actually know your programme accomplishes that?" By using good measures and data collection techniques, you can provide a better response to this question. For this reason, **systematic measurement in programme evaluation is the best tool you have for convincing people about what your programme does, how it functions, what outcomes are achieved and what has been done to improve it.**

This section (Step 6) provides necessary background information to ensure that you make good decisions about your research measures. If you have a researcher on your evaluation planning team, he/she should review the material in this section, then explain key concepts to the rest of the group. Although some of this information may seem "abstract," it is important for all partners to have a basic understanding, and for at least one partner to have a complete understanding of what is presented. By doing this, your team will be in a better position to conduct high quality evaluations.

If you need detailed information or training on how to gather valid, reliable and timely data on the prevalence, trends, and patterns of substance use, qualitative and quantitative methods for data collection and analysis, there is a variety of publications available in substance use epidemiology which you may wish to consult. The WHO Programme on Substance Abuse has also prepared a comprehensive Guide to Drug Abuse Epidemiology which is available upon request (WHO/MSA/PSA/97.14).

It's your turn (6 A)

Choose the best-suited person from your evaluation planning team to read the following section, **A Primer on Measurement Concepts**, then explain key concepts to the rest of your team.

A primer on measurement concepts

The next pages present crucial information for you to understand before choosing a specific measure. So make yourself comfortable and take some time to carefully read the following information. Throughout this section, you will have the opportunity to test your knowledge about the information that has been presented, and put it into action concerning your own evaluation project.

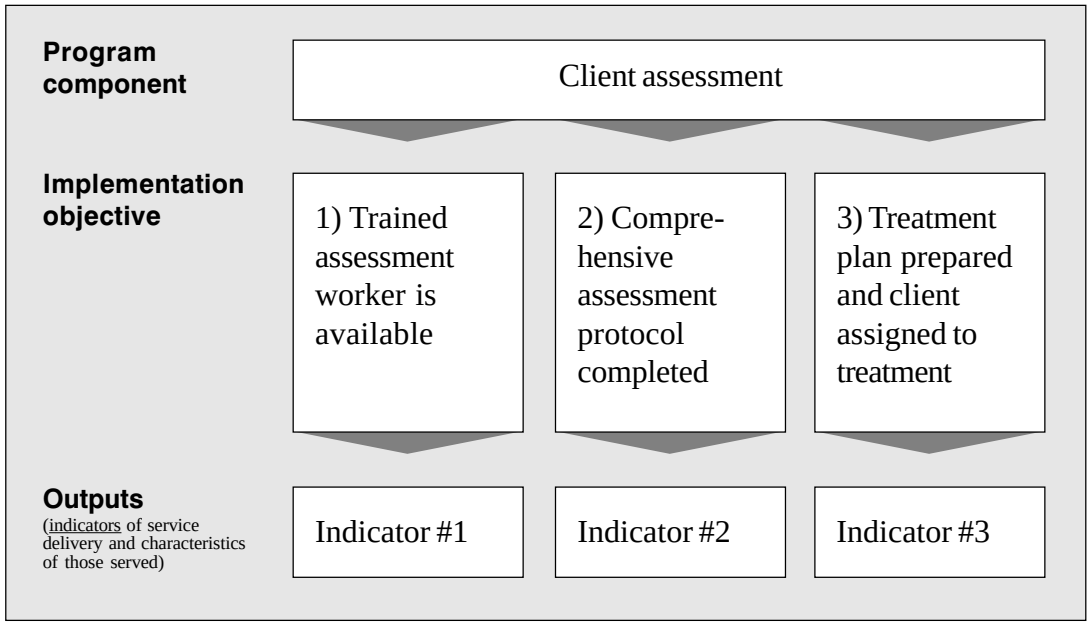
1. What are "indicators?"

Indicator variables are measurable pieces of information that indicate whether a programme is achieving an objective. Earlier, you learned about programme logic models. A logic model outlines the expected activities and services to be provided (i.e., process objectives) and the changes expected as a result of these activities and services (i.e., outcome objectives). In your evaluation planning, you use this logic model to develop the indicators to assess the achievement of the objective. Each objective of your treatment service or system has corresponding outputs and there are one or more indicators of these outputs. This relationship is illustrated below.

How does measurement fit into this situation?

Each of the implementation objectives in a programme has an output, and each output has an *indicator* of service delivery and/or characteristics of those served. Each indicator is a *measure* of the outputs that a programme produces at each stage of the treatment process.

Every programme should have clearly defined components and implementation objectives. *For each objective, you should try to come up with an indicator or measure that tells you whether the objective has been met.* By keeping a systematic record of all the indicators associated with each objective or outcome, you are systematically measuring how well the programme is doing.



2. Why bother with indicators?

By systematic collection of programme indicators, we can more precisely measure (and not guess at) whether the objectives are being met.

Why bother to use indicators at all? Can't we just look at a service or treatment system and tell whether objectives are being met? The simple answer is: NO! In this area, people are notorious for making poor "intuitive judgements." For example, individuals tend to remember events that confirm what should happen. They tend to forget events that contradict what should happen. People also tend to overgeneralise from one or two memorable cases to the overall situation. The best way to avoid these biases is to systematically collect information on a regular basis. By systematically collecting programme indicators we can more precisely measure (and not guess at) whether the objectives are being met. More importantly, having good indicators for our programme **enhances our confidence** in claims made about the programme. Consider the example on this page.

In this example, the claim being made by a programme manager is the same: "Our PS awareness programme is effective." *But our confidence level in believing this claim varies directly with the **quality** of the indicators (measures) used to support it*.

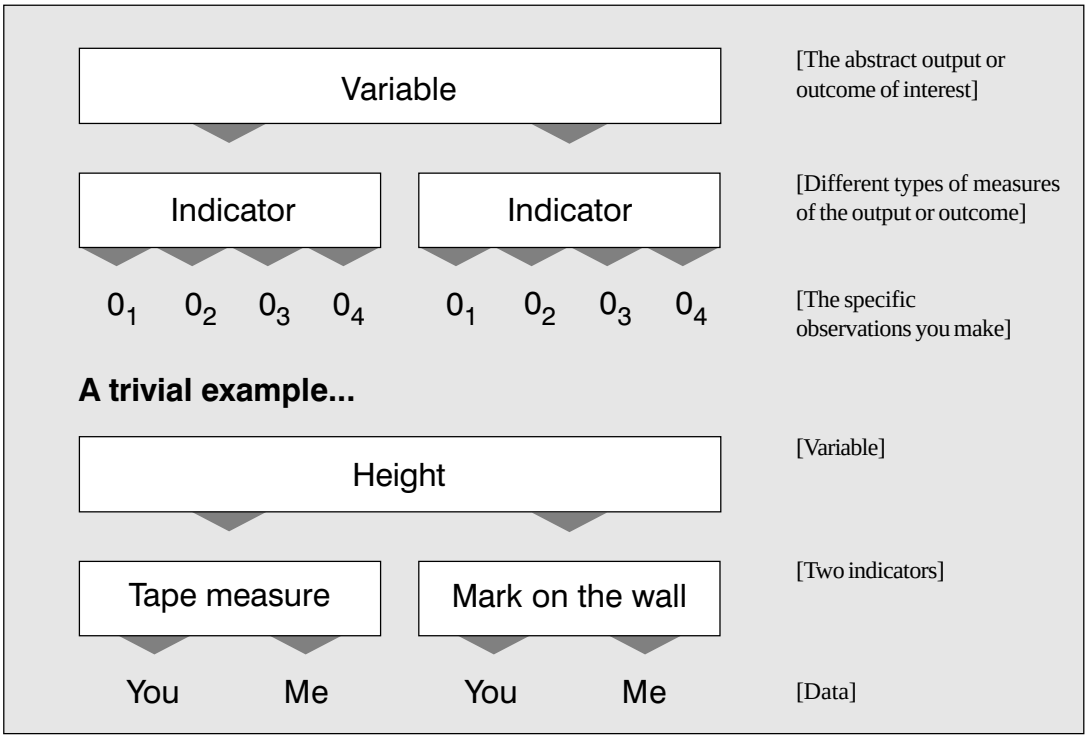
Claim	Confidence Level	Indicator
"Our PS awareness programme is effective"	Very Low	I heard it through the grapevine
"Our PS awareness programme is effective"	Low	I talked to John, who participated, and he said it worked for him
"Our PS awareness programme is effective"	Medium	40 out of 50 people in the programme checked off "effective" or "highly effective" in a satisfaction questionnaire completed at the end of the programme.
"Our PS awareness programme is effective"	High	Compared to pre-test scores, scores on an alcohol knowledge test went up by an average of two points
"Our PS awareness programme is effective"	Highest	Compared to a control group of people who weren't exposed to the programme at all, participants in the programme increased scores on seven out of eight measures of PS awareness

For instance, we'd be very suspicious of this claim if all the manager could say was that he or she 'heard it through the grapevine'. At the other end of the spectrum, we'd have a lot of confidence in this claim if the manager could show us that, in comparison to a group of people who weren't exposed to the programme, participants increased their scores on seven out of eight PS awareness measures. In the first case, the indicator is poor. In the second case, the indicator is very good. Thus, having quality indicators for your programme evaluation greatly increases the confidence you can have in claims being made on behalf of the programme.

Once you have chosen the basic indicators for your research questions, you must select a specific method for measuring them. Because different methods of collecting data provide different sources of information, you must be clear about what you want to know before you choose a method. As we have seen above, most indicators can be measured in more than one way.

3. The language of measurement

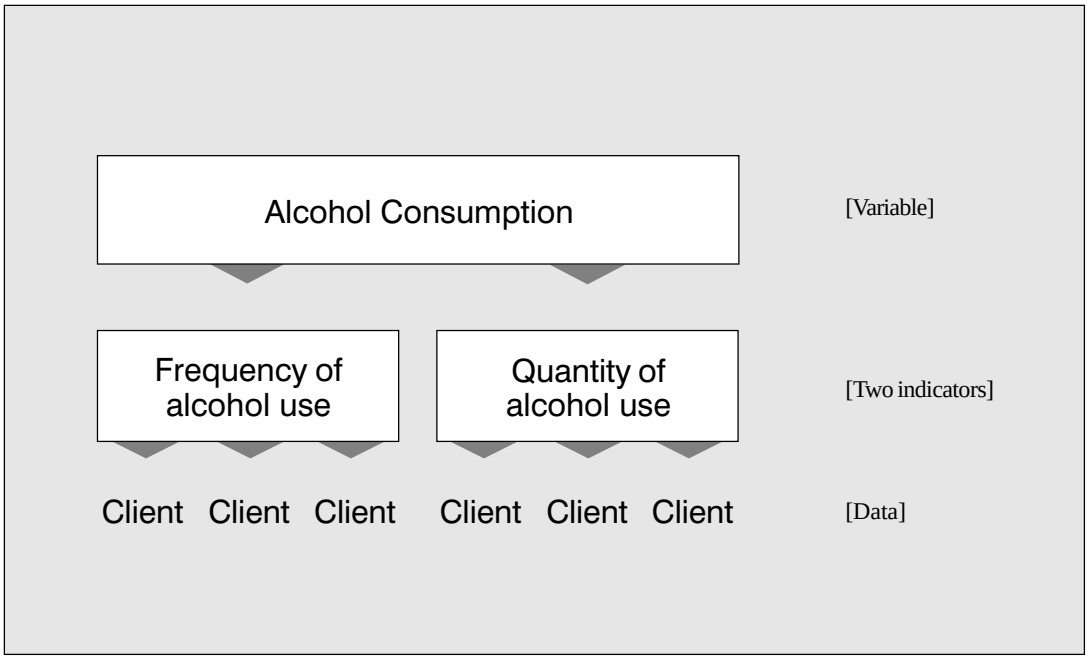
Let's define some basic terms. **Data collection** is about the process of finding or creating good indicators for programme objectives and systematically collecting information on these indicators. There are three basic concepts involved in data collection. *Variables* refer to the abstract output or outcome that interests you, such as PSU problems or depression. *Indicators* refer to the specific way you measure variables of interest, such as a self-reported measure of number of drinks per day, or a standardised questionnaire of depressive symptoms. *Data* refers to the concrete observations that you make with respect to each indicator, in other words, the specific responses from participants. The relationship among variables, indicators, and data is shown below.



A simple example of relations between variables, indicators, and data can be given with respect to measuring **height**. As a variable, height is an abstract concept that can be measured by at least two different indicators: a tape measure, and a mark on the wall. For each indicator, we could collect data on your height and my height. The outcomes we obtain for you and I (inches, using the tape measure indicator; marks on the wall, using the pencil indicator) constitute our data for the height variable.

Treatment services and systems are full of variables. The trick will be to find or construct good indicators for each of the variables that are specified in your programme logic model and evaluation questions. You must then systematically collect data (observations) for each indicator. Unfortunately, that is not as easy as it sounds. For many variables in programmes, indicators are not as obvious as in the 'height' example.

For instance, the figure below presents a more complex variable – alcohol use. What indicators could we use to see whether people actually change alcohol use after participating in a treatment programme? Two indicators that could be used are questionnaire items assessing frequency of alcohol use and another test measuring quantity of alcohol use. In this example, we have one variable (alcohol consumption) and two indicators of it (frequency and quantity). If we wanted to collect data to see whether our programme accomplished its goal of decreasing alcohol consumption, we could administer the two indicators to clients before and after treatment. These observations constitute data that we collect on the alcohol consumption variable, using the two indicators.



Into action

After generating possible evaluation questions, the evaluation team decides to concentrate their evaluation on coverage at the activity level and effectiveness of their AIDS education. An evaluation about basic coverage at the activity level provides answers whether their programme accomplished what was expected to accomplish. For example if clients were referred to long term treatment after their detoxification. An evaluation of the effectiveness of their AIDS education provides answers, if this intervention achieves its objectives. For example if clients' knowledge about AIDS increased after the intervention. The intervention consists of five 1 hour group sessions in which clients learn about the human immunodeficiency virus (HIV), transmission, prevention, symptoms, HIV testing and focuses on personal susceptibility and situational analysis.

Main research questions and their indicators

	Questions	Indicators
1	Has the number of referrals increased from the previous year?	Review of records
2	Did the characteristics of the clients change in comparison to last years clients?	Review of intake form records
3	Has the number of self-discharges decreased from the previous year?	Review of termination form records
4	Has the number of clients who were referred to long-term treatment increased from the previous year?	Review of termination form records
5	Did attitudes in favour of HIV low-risk practices improve among clients?	Before/after intervention questionnaire asking about attitudes towards AIDS.
6	Did client's HIV-risk behaviour change in favour of low risk practices after treatment?	Before/after intervention questionnaire asking about risk behaviours.
7	Did clients knowledge about AIDS increased after the AIDS education intervention?	Before/after questionnaire assessing knowledge about AIDS.
8	Did clients self-efficacy regarding their ability to use skills maintaining AIDS harm reduction behaviours increase?	Before/after questionnaire assessing self/efficacy.

It's your turn (6 B)

1 Test your knowledge. Describe the following terms to your evaluation partners:

- Variable
- Indicator
- Data

Provide an example of each from your PSU programme or treatment network.

2 Select indicators for each of the research questions that your group has chosen (see Step 5). At this point, do not concern yourself with exact measurement tools; selection of exact tools will be addressed below. For now, just write down using simple language some reasonable indicators for your research questions.

4. Quantitative and qualitative measurement

You are using a quantitative approach to indicators whenever the measurement of a variable is conducted through numbers.

You are using a qualitative approach to indicators whenever you measure a variable with words.

The preceding section distinguishes among variables, indicators, and data. We see that for many aspects of PSU services and systems (unlike for physical properties, such as height and weight), the variables are complex. Variables such as PSU awareness, reduction in drinking, consequences of PSU, self-esteem, motivation, well-being, and other objectives of programmes can be measured in several ways. The next step is to critically think about the different **types of indicators** that are available to help you measure such complex variables in your evaluation.

The world of measurement generally uses one of two different strategies, depending on whether indicators are numbers or words. You are using a **quantitative** approach to indicators whenever the measurement of a variable is conducted through **numbers**. Going back to the example used earlier, if one objective of our programme is to ensure that an assessment worker is available, we could construct a quantitative indicator of this objective by computing the percentage of work days per month in which the worker was available. This percentage would ideally total 100, and provide a number that summarises information for the variable. Another quantitative indicator may be the number of days a person must wait for their assessment. Quantitative indicators are useful because they:

- conveniently summarise a large amount of data reflecting key objectives of a programme
- can be easily translated into graphs which portray the results of an evaluation
- can be analysed using statistical techniques

Qualitative indicators also can be used in programme evaluation. You are using a **qualitative** approach to indicators whenever you measure a variable with words. One distinct advantage of the qualitative approach is that it can preserve the unique point of view of the people being studied. For example, if we are interested in the variable "well being" as an outcome of

our programme, we may not have a good quantitative indicator of well being. A qualitative option would be to ask clients about their well being at the end of a programme. We could do this with questions about their physical and mental health and their ability to cope with the stresses of daily living. We might tape-record their answers and identify themes that reflect "well being" and perceptions that it has changed as a result of programme participation. In this case, language (not numbers) provides indicators for the variable(s) of interest. Because the qualitative approach relies heavily on words it incorporates the distinct point of view of the person(s) talking. Well being may be discussed in various ways by different people and cultural groups.

Review this example: The Timmins detoxification service evaluation, described in the back of this workbook, used both quantitative and qualitative indicators. The planners' rationale for why they chose qualitative and/or quantitative indicators for each variable is well described in the case example.

It's your turn (6 C)

Test your knowledge. Choose whether each of the following variables is quantitative or qualitative:

1 Number of alcoholic drinks consumed in past 7 days.

2 Total depression score on a self-reported questionnaire of depressive symptoms.

3 Patients' descriptions of the circumstances surrounding their first use of PS.

4 Age of patients.

5 Gender of patients.

6 Number of different types of PS that patient has used in past 90 days.

7 Patients' descriptions of their experiences in your PSU treatment programme.

8 Patients' ratings, on 0 – 10 scale, of their satisfaction with your PSU treatment programme.

9 Clinicians' reports of the extent of patients' improvement while in the programme.

10 Clinicians' ratings, on a 1 - 5 scale, of patients' improvement while in the programme.

Answers: Quantitative - 1, 2, 4, 5, 6, 8, 10 / Qualitative - 3, 7, 9

5. Quantitative indicators: determine what measurement scales you need

Depending on your specific evaluation question, quantitative data are available from a variety of sources, including observation, questionnaires, interviews, and record reviews (reviewed in point 8 below). Regardless of how you obtain quantitative indicators, if your evaluation is attempting to provide **numbers** for key variables in your program, you need to determine the kind of quantitative scales that should be used to collect the data. A scale is simply a classification scheme for describing the nature of your observations. There are four basic types of quantitative scales. You will probably use a combination of them in quantitative programme evaluation:

Nominal scales provide labels for the data you collect. A simple example would be to assign different numbers for hair colours: red, black, blond, brown, and so on. Many variables in a programme evaluation can be labelled using a nominal scale. For instance, if you want to track the number of men and women in your programme in a given time period, you could assign the number 1 for males and 2 for females. The important thing to note about nominal scales is that they are **mutually exclusive**. That is, you can't be **both** male and female, **both** blond and red-haired, and so on.

Ordinal scales are arranged according to a ranking (e.g., few —— many; bad —— good). One commonly used ordinal scale in programme evaluation is socio-economic status. One scale might assign the value 1 to people at a low socio-economic class, 2 = middle class, and 3 = upp er class. You cannot add or subtract the difference among values on the scale.

Interval scales provide numerical labels reflecting a **magnitude** of differences; however, you can only add or subtract the differences among values in the scale. Often attitudes and beliefs are measured on interval scales. For instance, you might ask your staff to indicate their degree of agreement with the statement: "Should people convicted of driving while impaired be forced to go to treatment?" on a scale ranging from 1 (strongly disagree) to 9 (strongly agree). Another common example is clients= rating of satisfaction with the programme: (1) very unsatisfied; (2) unsatisfied; (3) neutral; (4) satisfied; and (5) very satisfied. Another example might be to count the number of areas of the client's life affected negatively by their PSU. A client reporting three problem areas (e.g., physical health, social relationships and work) has two problem areas more than another client reporting consequences in only one area (e.g., mental health).

Finally, **ratio scales** provide numerical labels that can handle any mathematical equation, and include a true zero point. For example, weight and distance are arranged on ratio scales. We know that 10 kilometres is twice as far as 5 kilometres. Figure 8 summarises the four levels of quantitative measurement. The figure contains useful reminders of the types of information that can be derived from using these types of indicators.

Four types of scales

Level of measurement	Type of information
Nominal	**Two or more unordered categories** e.g., Did your spouse influence your decision to enter treatment? ☐ YES ☐ NO
Ordinal	**Two or more ordered categories** e.g., Rank of sources of motivation to enter treatment (e.g. family, employer, courts...) ☐ Rank 1 ☐ Rank 2 ☐ Rank 3
Interval	**Numerical labels reflecting magnitude of differences** e.g., 1 = strongly disagree / 9 = strongly agree 1 2 3 4 5 6 7 8 9
Ratio	**Continuous scale, true zero point** e.g., How may drinks did you have in the past week? ☐ drinks

6. When should you use qualitative and quantitative measurement?

One issue that arises in programme evaluation is determining when each measurement strategy is appropriate. The question is: when do you use numbers (quantitative approach) and when do you use language (qualitative approach)? One good rule of thumb is to use quantitative indicators whenever you already have a good idea about how to measure the output or outcome of interest in your programme evaluation. There are several ways to do this. Perhaps the best way is to take advantage of someone else's previous work and use their scale or indicator for the output or outcome. The workbooks that are part of this series offer advice on quantitative measures that can be used to measure client characteristics, treatment process, client satisfaction and outcomes. If you do enough background research and library work, you may be able to find a standardised measure or scale that can be used in your evaluation. However, it is important to be sensitive to the fact that measures that are deemed to be useful in a certain culture may not be valid or useful in other cultures. If you can't find an appropriate measure for your setting, then you may have to create something new. This situation is discussed later in this workbook.

What if you don't have a good idea about how to measure the output or outcome quantitatively? Qualitative methods can help. For example, if your search of the literature did not result in finding a culturally appropriate, quantitative measure of "PSU," you might want to approach the measurement issue from a qualitative point of view. This could be done by interviewing programme participants about this variable. This approach has the advantage of not presuming what 'PS consumption' is for the respondents. It allows them to introduce their own unique interpretation of the variable, which may better reflect their unique situation and culture.

Other factors affecting whether you use quantitative or qualitative methods in your programme evaluation are (a) the degree to which you have expertise and/or training in each method, and (b) the resources (time, staff) available to you.

7. Reliability and validity

Reliability refers to whether an indicator is consistent across time and observers.

As mentioned earlier, your confidence in making claims about the programme increases as the quality of the indicators and data used to measure programme variables increases. But how do we determine our confidence in the measures used? Regardless of whether the measurement strategy adopted in your programme evaluation is quantitative or qualitative, measures that you use in collecting data about your evaluation questions must be both **reliable** and **valid**.

Reliability

This refers to whether an indicator is **consistent across time and observers**. A simple example would be to consider a thermometer. A reliable thermometer, when plunged into boiling water, should give the value of 100 degrees centigrade. Further, if the instrument is truly reliable, it should give the same measurement every time it is plunged into boiling water, no matter who is viewing the temperature.

Clearly, our confidence in the indicator (the thermometer) decreases if different results are obtained each time we take it in and out of this same water. The same situation is true for indicators designed to answer your specific evaluation questions. For some indicators (e.g., gender, income, education), reliability of the measures may not be compromised. That is, unless respondents are lying or misinterpreting the question, you can be relatively confident that data collected on these indicators is reliable. Reliability is not so easily established, however, for more complex variables, such as reduction in alcohol use, self-esteem, PS awareness and quality of life. One advantage of using existing indicators (e.g., quantitative scales, qualitative interview schedules), is that it is likely that they have been tested previously for reliability.

For most subjective and objective outcome measures, an individual should answer questions in the same way if the question is asked more than once within a short period of time. This is called **test-retest reliability**. If this pattern does not occur for most observations, the indicator is probably unreliable. As part of a pilot test of your evaluation, you might want to conduct a test-retest study using a small number of people who are similar to those who respond to your final evaluation. This approach will determine the consistency of their responses.

In the case of qualitative data, you should use two different people to analyse the information to see if they identify the same themes. For example, a client may have been asked to describe her perceptions of how treatment has affected her life. If two people analyse the transcript of the interview and both identify a theme of "losing friends after treatment" then the result would be reliable. This concept of reliability is called **inter-rater reliability**, and refers to the ability of independent raters to agree on measurements provided by indicators.

You may also want to build reliability/consistency checks into a questionnaire. You can do this by including several questions which people would be expected to answer the same way. For example, a measure of client satisfaction could ask respondents to indicate how pleased they were with the programme, how much they enjoyed the programme, and areas of the programme that were problematic. A high proportion of inconsistent ratings on these dimensions would indicate that the items are unreliable. This concept of reliability is called **internal consistency**.

Validity

Validity concerns the extent to which you are actually measuring what you intend to measure in your programme evaluation.

An indicator may be reliable and yet not be at all valid. Returning to the thermometer example, if the thermometer consistently reports a temperature of 110 degrees centigrade every time it is put into boiling water, it is reliable, yes, but invalid (since water boils at 100 degrees, not 110 degrees). Validity concerns **the extent to which you are actually measuring what you intend to measure in your programme evaluation**. As another example, your interview (or qualitative) indicator of self-esteem would have low validity if it just measured how much the respondent liked the interviewer. As in reliability, the more valid your indicators, the more confidence can be placed in the data collected on those variables.

Several kinds of validity can be determined for indicators used in a programme evaluation. **Face validity** refers to whether the content of the specific questions or measures reflect what the indicator is *supposed* to measure. For example, asking a question such as "How well did you like the other clients in the programme?" would have poor face validity if the indicator is really supposed to measure client self-esteem.

An indicator can also be validated by comparing it to a known measure of the same variable. The result is called **concurrent validity**.

The validity of qualitative indicators can be established by (a) providing concrete examples of themes identified in written materials, (b) demonstrating that the people responding to the interviews were not attempting to lie or deceive, and (c) relating themes identified to **other indicators** that enhance confidence in your interpretation of the materials.

In summary, paying serious attention to the reliability and validity of your indicators can greatly enhance confidence in any claims made during the course of a programme evaluation. When you use reliable and valid indicators, people will put much more weight on your conclusions than if you used indicators with low reliability and validity. Often, you can use existing instruments that have been tested and fine-tuned to increase reliability and validity. When this option is available to you, it is preferable to use existing indicators (e.g., scales, interview schedules, as opposed to developing new tools).

It's your turn (6 D)

1 Describe reliability and validity to your evaluation partners. Explain why each is important in your choice of measures.

2 As a group, complete the following exercise:

A driver is stopped by the police for suspicion of driving under the influence of alcohol. Which of the following measures is a reliable and/or valid indicator of the driver's level of intoxication?

Measure	Reliable?	Valid?
Driver's report of number of drinks he/she had that night		
Performance on tests of motor co-ordination		
Blood alcohol level		
Eye colour		

3 Think of at least two other reliable and valid measures of PS intoxication.

Answers to question 2

Measure	Reliable?	Valid?
Driver's report of number of drinks he/she had that night	Yes	No
Performance on tests of motor co-ordination	Yes	Yes
Blood alcohol level	Yes	Yes
Eye colour	Yes	No

8. Types of measures

A final measurement concept concerns the different types of measures that you can use in your research. Each method has advantages and disadvantages (explained below). Different measures may be more or less appropriate for different studies.

• Observation

Observation is a procedure for gathering information by carefully watching and writing down the behaviour of individuals or events.

Observation is a procedure for gathering information by carefully watching and writing down the behaviour of individuals or events. Researchers choose the time and place of the observation so that they will have a good chance of seeing the people or behaviour that they wish to observe. Usually people or groups are observed repeatedly to make certain that the observation is complete and accurate. The observers may record everything that they see, or they may only record certain variables such as number of alcoholic drinks consumed.

For the observations to be useful in an evaluation, it is essential that staff use a standard protocol for recording them. Rating forms completed before treatment and at various points during and after treatment can be used. Staff must be carefully trained to ensure that they all use the observation procedures the same way. That is, if several staff members interview the same client on the same day and fill out an observation from, you would want to be confident that their observations or ratings have high inter-rater reliability. Otherwise, your findings will be unreliable. For example, some counsellors consider "engagement or involvement in the treatment process" to be a useful indicator of client progress and a predictor of relapse; an instrument could be designed to rate client involvement over time. If the instrument were reliable and valid, the results could be subsequently correlated with information about relapse or reduced drinking to test the hypothesis that client involvement predicts outcome.

Observation is a good way of checking the validity of information learned through questionnaires or interviews. Observing behaviour allows you to confirm these statements.

On the other hand, observation is time-intensive, and does not allow you to understand participants' thoughts and feelings.

• Questionnaires

Questionnaires are useful when you want to collect a small number of clearly-defined facts from a large number of people.

A questionnaire is a written set of questions that a participant answers by writing in the spaces provided on the same sheet. Questionnaires are often self-administered. This means that the research participant reads the instructions and completes the questionnaire with little help from the investigator.

Questionnaires can include *closed-ended questions*, in which participants must choose from a list of possible answers, and/or *open-ended questions*, in which no specific answers are provided: people can answer any way they choose. Closed-ended questions are usually quantitative, while open-ended questions are usually qualitative.

Questionnaires are useful when you want to collect a small number of clearly-defined facts from a large number of people. They may be a good choice when you want to gather information about sensitive topics, such as sexual behaviour. Some people will feel more comfortable answering a written question than talking to an investigator face to face.

A major disadvantage of questionnaires is that only people who read and write can use them. If you know or suspect that many clients have reading problems, do not use self-administered questionnaires. Be alert for signs that a client may be having difficulty understanding the instrument. Be prepared to help such clients by having a staff member read the questions. An additional disadvantage is that participants may get caught in a *response set* — that is, they may answer a series of questions in the same way without thinking through each one.

• Interviews

Personal interviews offer several benefits.

Interviewing involves meeting face to face with an interviewer, who asks the person specific questions and records the person's answers. Questions may be open-ended or closed-ended.

Personal interviews offer several benefits. They avoid difficulties associated with reading problems. They permit you to probe for in-depth information and to continue a particular line of questioning when appropriate. If an item is failing to work as expected, an interviewer may recognise a problem that would not be obvious from self-administered instruments. An interviewer can make sure that identifying information that links one instrument to another — for example, to link a client's satisfaction survey to a record of the client's time in the programme — is filled in completely and accurately.

Interviews can be subject to the same respondent biases as self-administered questionnaires. For example, tendencies to a response set and giving socially desirable answers can occur in interviews as well as in self-administered questionnaires. They can also be subject to interviewer bias and interview error. For these reasons, interviewers must be trained. Interview time involves cost considerations. Costs are highest for personal interviews, less for telephone interviews, and comparatively modest for group interviews.

Interviews done in person usually produce a slightly higher response rate than telephone interviews, but they are more costly. Telephone interviews, while less costly, have two potential problems: not all clients may have telephones, and privacy and confidentiality may be difficult to ensure. Clients may also be reluctant to talk or give honest answers if family members or coworkers are present during their telephone interviews. Also, the typical telephone interview is shorter than a personal interview.

Group interviews can be conducted by having an interviewer read questions from a self-administered form while clients record their responses on their own form. This is a practical approach to improving response rates and addressing low literacy, without the time and expense of personal interviews.

If your programme staff is conducting interviews, they should be trained carefully. They should know such interviewing techniques as ensuring confidentiality to respondents, maintaining objectivity, asking questions exactly as stated, recording responses properly and legibly, probing, establishing rapport, and being sensitive to the cultural values of respondents.

• Focus group discussions

Focus groups are best seen as a way to identify issues and clarify concepts.

Another interview format that can be used is a focus group, which is a general discussion between 7-8 individuals on a selected topic. Focus groups usually last for one to two hours. They are run by a trained moderator whose function is to guide the discussion. The

overall aim of focus group interviews is to provide an understanding of the thoughts and feelings of participants as they consider an issue (e.g., an aspect of your programme). Focus groups are best seen as a way to identify issues and clarify concepts. Transcripts of focus groups can be used to supplement programme evaluations by providing in-depth accounts of participant reactions to an issue. It would be unwise to rely on focus groups as the sole source of observations for your programme evaluation. This is because focus groups are open-ended and can provide information about many different variables not all of which are relevant for your evaluation. It is also difficult to establish reliability and validity of focus group data. For these reasons, focus groups are more useful for certain types of evaluation activities (e.g., planning an evaluation, needs assessment) than others (e.g., outcome evaluation and/or economics).

Specific instructions for using the focus group method are outlined in Appendix 1 of this workbook.

• Examining Routine Records

Existing data from other agencies is less expensive to use in evaluation than data obtained directly from the client in a self-report or interview format

You may obtain useful records for measuring programme processes and outcomes from your programme files. Increasingly, this information is available through computer information systems. Relevant information may also be accessible from other treatment programmes that may treat some of your clients, such as from school health, social service records, or police or court records. Information routinely submitted to government from all treatment agencies in your jurisdiction may also be accessible.

Three types of instruments commonly used in treatment programmes for PSU disorders are used for programme evaluation intake records, case notes, and termination forms. The *intake data* may contain a brief client-flow or census form, a general background and behavioural history assessment. It may also contain a more extensive diagnostic assessment battery. The data are typically collected by an intake worker, treatment counsellor, or records clerk in one or more interview sessions. *Case notes* may contain information designed to document how a client has adjusted and made progress in changing behaviour. The notes can also include information obtained on schedule to assess domains of treatment process and client progress over time. The *treatment termination form* shows the time of, and reasons for, a client's discharge.

Existing data from other agencies is less expensive to use in evaluation than data obtained directly from the client in a self-report or interview format. This is because the information has already been collected. However, since the data were collected for other purposes, they may contain biases that are difficult to discover. If you ask other agencies for access to their records, you must be very specific about what information is needed, what analyses will be done, and how the confidentiality of records will be protected. You must also indicate to whom the results will be reported.

Formal request for access to records is only part of the process. You may want to meet informally with the agencies' managers before you submit a written request. Offer to meet with the organisation's staff to work out procedures to get the data in the least disruptive way possible. Get the name of a contact person who can answer future questions about how to interpret the data, and develop a working relationship with the contact person as soon as possible. You will need the organisation's cooperation. Collecting data often takes several months, and learning how to use it may take time. Request access to records or archival information ahead of when you actually need it.

Into action

The evaluation team thought about the following types of measurement for their research questions.

1 Has the number of referrals increased from the previous year?

2 Did the characteristics of the clients change in comparison to last years clients?

3 Has the number of self-discharges decreased from the previous year?

4 Has the number of clients who were referred to long-term treatment increased from the previous year?

Type of measurement: To answer questions 1-4, the team would use their *intake records* and *termination forms*. Their intake data contains information about each client relating to age, gender, marital status, education, source of referral, involvement in health and justice system, general mental and physical health and their individual drug history including adverse consequence of their PSU use and risk behaviours. Their treatment termination form includes the time of, reasons for client's discharge and if the client was referred to any long-term treatment.

5 Did attitudes in favour of HIV low-risk practices improve among clients?

Type of measurement: The team agreed upon using attitude scale (questionnaire) assessing perceived susceptibility to AIDS, benefits of risk reduction, attitudes towards slips in drug-use and sexual risk behaviour. The scales chosen are interval scales. The client has to indicate his/her degree of agreement with statements about AIDS on a scale ranging from 1 (strongly disagree) to 4 (strongly agree).

6 Did client's HIV-risk behaviour change in favour of low risk practices after treatment?

Type of measurement: The team decided to assess HIV-risk behaviour by using a questionnaire asking about risk behaviours associated with drug use and administration and sexual risk behaviours (see sample questionnaire about risk behaviour, appendix 2). The level of measurement would be nominal, e.g. Have you ever shared a needle, syringe/ spoon/ cooker or cotton/ filter with anyone at any time in your life? and interval, e.g. during the past 90 days how many days did you inject any kind of psychoactive substance?

7 Did clients knowledge about AIDS increase after the AIDS education intervention?

Type of measurement: Knowledge scales assessing knowledge of methods whereby HIV is transmitted, methods for prevention of transmission and understanding of the HIV antibody test. The level of measurement of this scale would be nominal. The number of correct responses to Yes/No items (0-15) would be recorded.

8 Did clients self-efficacy regarding their ability to use skills maintaining AIDS harm reduction behaviours increase?

Type of measurement: Scales assessing self-efficacy regarding cleaning injection equipment with bleach and using condoms. The level of measurment would be nominal. Clients would have to indicate their degree of confidence (1 = not at all confident - 4 = extremely confident) regarding five behaviours.

It's your turn (6 E)

With your group, discuss what type(s) of measurement (e.g., questionnaires, record review) you would use for each of the following research questions, and why:

1 Did patients who reported prior sexual abuse finish the programme more often than those who reported no abuse?

2 Did patients who completed the programme decrease their PSU?

3 What are the attitudes among community members about our treatment programme?

Now what? Take a deep breath. You have just completed the primer on measurement concepts! Now it's time to carry forward with the important task of selecting or prepare a data collection instrument.

Select or prepare a data collection instrument. It is important to note that your data collection instrument can contain combinations of qualitative and quantitative measures taken from existing scales, indicators obtained from other sources, newly-created measures, and interview questions. Several PSU-related questionnaires are contained in the back of this workbook, in Appendix 2. Here is a very simple data collection instrument:

A simple data collection instrument (Not validated or reliable)

Quantitative indicators
(for 6 variables)

1	Program	**1**	Detox	**2**	Residential
2	Client age		years		
3	Client education		(highest level attained)		
4	Days employed in past six months				
5	Rosenberg Self-Esteem Score				
6	Alcohol Dependence Scale Score				

Qualitative indicators
(for 2 variables)

7 Aspect of the treatment programme liked best:

8 Aspect of the treatment programme liked least:

As you can see in from this figure, this instrument contains two main sections. One section records data for six quantitative variables (program, client age, client sex, client education, days employed, self-esteem score and alcohol dependence score). Note that

some of the indicators are relatively simple (age, sex), while other indicators in this section consist of reliable and valid measures that have been used in previous research (e.g., Rosenberg self-esteem score; Score on the Alcohol Dependence Scale). A second section of this data collection instrument consists of client responses to questions about aspects of the treatment programme they liked most and least. Space is presented on the form to record the language the client used to describe these variables. This part of the data collection instrument measures qualitative information.

Tips for developing questions for your data collection instruments

Sometimes you cannot rely on existing instruments, scales, forms, or interview schedules to create your data collection instrument. This section provides tips for constructing questions that can be used in either quantitative or qualitative data collection. It is worth repeating that there are many advantages to using existing instruments that have been proven to be valid and reliable. Be sure, however, that they collect the data you need, and that they are applicable for your culture and/or setting. In addition, be aware that any revisions you make to existing instruments such as rewording, eliminating, adding, or reordering items may diminish the validity and reliability of an established instrument. If you must design a new indicator or question, here are some tips for constructing it:

- Be sure that the question collects data on the measures needed for your evaluation questions.

- If pre-test and post-test items are used to measure change over time, measure items that are sensitive to change. Also measure items that different clients are likely to answer differently before treatment than after treatment. For example, ask *How many days last month did you use cocaine?* rather than *Have you ever used cocaine*?

- Divide questions in which two or more different or conflicting concepts are presented at the same time. Otherwise the answers will not be meaningful because some respondents will answer one question and others the second. For example:

  (Avoid) *How do you rate the **comfort** and **convenience** of the group therapy sessions?*

  (Use two questions instead): *How do you rate the comfort of the meeting area for the group therapy? How do you rate the convenience of the time that groups are scheduled?*

- Use an "other" category when it is not practical to provide respondents with an exhaustive list of response categories. For example, if you want to know the job positions of treatment staff, list only commonly held positions and let those who do not fit those categories specify their positions. If enough common answers are noted in analysis of the "other" category, each can be given a code for purposes of analysis. Alternatively, you can use footnotes to give examples of the "other" positions.

- Use open-ended questions sparingly since they require an extensive amount of time to interpret, classify, and code. (An open ended-question does not include response categories such as yes/no, or a list of answers from which the respondent can choose.). Rather, it is a broad question for which an infinite variety of responses might be given — for example: *What has motivated you to come for treatment?*

- Whenever possible avoid skip patterns, because some respondents find them confusing and others will ignore instructions and respond inappropriately. A "skip pattern" refers to instructions to skip certain questions that do not apply to the respondent. For example, an

intake from containing questions about alcohol use and pregnancy would usually contain instructions for male clients and women who had never been pregnant to skip ahead. If the use of skip patterns seems warranted because items will apply only to part of the respondent sample make the instructions as clear as possible.

- Tailor the language used to the reading level of the respondents. If reading levels range considerably, items should be worded so that they can be answered by persons with different skills.

- Make certain that the wording of items is sensitive to gender, age, ethnic, and cultural differences in interpretation.

- Most important, remember to keep your data measurement as simple as possible. Do not collect information you do not need.

The second case example located in the back of this workbook describes the development of a new data collection instrument: The Maudsley Addiction Profile (MAP). The evaluators designed the MAP to be used in treatment outcome evaluations. It assesses several domains, including substance use, risk behaviour, and social functioning.

The figure bellow summarises several key points to remember if you decide to write your own questions as part of a qualitative or quantitative programme evaluation.

Writing good questions: a checklist for quick reference

1 Are the words simple, direct and familiar to all?

2 Is the question as clear and specific as possible?

3 Is it a double question?

4 Does the question have a double negative?

5 Is the question too demanding?

6 Are the questions leading or biased?

7 Is the question applicable to all respondents?

8 Is the question objectionable?

9 Will the answers be influenced by response styles?

10 Have you exhausted the response alternatives?

It's your turn (6 F)

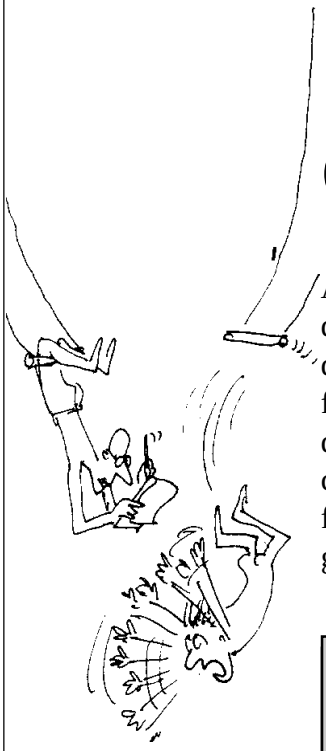
1 With your group, create the data collection instrument using the indicators you need. Use the information from the primer and the checklist above to guide your decision-making. While preparing the instrument, decide what method(s) you will use to collect the data (e.g., interview, questionnaire, record review).

2 Review your data collection instrument in light of your original research questions. Are you asking the right questions?

3 Review your data collection instrument with the expected users of the research. Revise as needed.

Step 7

Prepare a data collection plan



After selecting measures, you also need to develop an evaluation design. An evaluation design is a set of instructions about when and from whom to collect data. The evaluation design should be sound, so that you can be confident in the conclusions that you make from you research results. For example, a good design will increase the confidence that your patients are improving, and/or that the programme itself is producing the results.

Workbooks 3 through 8 present detailed information about how to conduct different types of research. You should consult the workbook that is applicable for your project.

If you are conducting a...	Then you should review...
Needs Assessment Evaluation	Workbook 3
Process Evaluation	Workbook 4
Cost Evaluation	Workbook 5
Client Satisfaction Evaluation	Workbook 6
Outcome Evaluation	Workbook 7
Economic Evaluation	Workbook 8

If you are not sure what type of evaluation you are conducting, review the Framework Workbook to determine your evaluation type.

Into action

After selecting their measures, the team discussed their evaluation design. Adam S., the researcher explained to the other team members that they need to think about a set of instructions about when and from whom to collect the data. He explained to them that according to their research questions, they were planing a process evaluation and an outcome evaluation. Their process evaluation posed questions regarding to coverage at the activity level of the service. For example, they wanted to know what proportion of clients completed the detoxification. Outcome evaluation relates to the questions regarding the effectiveness of their AIDS education. Adam S. explained that in order to evaluate the effectiveness of this intervention, the best design option would be a randomised control trial. One group of clients would be randomly assigned the AIDS education intervention and one group of clients would not receive this intervention. However, in view of the available resources, the team would have to choose a pre-post design, whereby clients who attend the programme would be assessed before and after they completed the intervention. This design would not control for competing explanations for changes in HIV risk behaviours, however it could determine if the objectives of this intervention were achieved.

It's your turn (7 A)

1 Determine what type of evaluation (e.g., needs assessment, client satisfaction) you are doing.

2 If you have not already done so, review the appropriate specialised workbook to assist you with choosing your research measure(s) and developing your data collection plan.

Set a time frame for data collection

It is important to allow enough time to pass for data collection.

As part of your overall evaluation, you must define a period for data collection. It is important to allow enough time to pass for data collection. Exactly how much time depends on several factors. You must consider what type of change you want to measure. For process evaluation, this will often be a 6-12 month period long enough to measure programme operations, but short enough to answer your evaluation questions in a reasonable period of time. For outcome monitoring, at least a six month follow-up should be planned; often, longer time frames are required, because some behavioural changes in clients take a long time. An even longer period is required to determine whether changes in behaviour are more than temporary in nature. Outcome evaluations between two and five years are common. One option is to plan for an initial evaluation period in which you collect data (e.g., 6 or 12 months), and after analysing the results, develop a plan to extend the data collection period.

Into action

The team decided to conduct a process evaluation and an outcome evaluation. The next step would be to define a period for data collection for each evaluation.

For the process evaluation the team estimated a period of 4 months. By the end of the year, the secretary would have enter all intake and termination forms into the central database. In the following 3 months Adam S. would then calculate number of referrals, dropouts, and referrals to long-term treatments and mean values for client characteristics for the present year and the previous year in order to do a comparison.

For the outcome evaluation, the team planned a one-year data collection period. In the course of this year, all clients attending the AIDS education intervention would be interviewed at the second day of their detoxification about their knowledge, high-risk behaviour and attitudes about AIDS (baseline interview, T_0). One follow-up would be conducted, six month after the first contact with the client (T_1).

It's your turn (7 B)

1 Using the appropriate specialised manual as a guide, choose your time frame for data collection.

2 Review your data collection time frame in light of your original research questions. Does your time frame adequately address your questions?

3 Review your time frame with the expected users of the research. Revise as needed.

Determine the sample you will use

Because of limited resources, you will have to base conclusions about the programme on a subset of all the possible outputs and outcomes that can be measured during a data collection.

After you have determined an appropriate period for data collection, the next step is to determine how the data collection instrument will be used. It should be obvious that no data collection can go on forever. In fact, every programme evaluation has limited resources and time. Because of limited resources, you will have to base conclusions about the programme on a **subset** of all the possible outputs and outcomes that can be measured during a data collection. Limitations on resources and other practical considerations also mean that you will only collect information from a subset of the people involved with your programme, including your clients. The following basic statistical concepts will help you here. *Populations and Samples*: A **population** refers to all of the variables of interest for a particular evaluation question. For example, the population of suicides in your community refers to all the suicides that have occurred — every one of them. Another example might be the population relapses after attending your programme — again, the population refers to every single relapse that could be measured. Any summary of **quantitative** information produced for the entire population of

outputs and outcomes is called a **parameter**. Any **qualitative** summary information we produce for the population is called an **ideal type**. For example, if we had access to the self-esteem scores of all the clients who ever participated in our program, we could generate the average self-esteem score among the population of clients served by our program: a parameter. Alternatively, if we summarise a single theme reported by the entire population of clients, this represents an ideal type.

Because we typically don't have time or resources to measure all the observations, a **sample** refers to those individual variables that are **actually examined** in the course of data collection. For example, because we don't have enough money to measure the population of outcomes for your programme (i.e., every client who has passed through the door) we take a sample of clients — for instance, all of the clients who leave the programme in a six-month period. In this case, we must base any conclusions or claims about client outcomes **in the population** on just the six months of data collected **in our sample**. If we have a **representative** sample, a sample of measured outputs and outcomes that closely mirrors the larger population, our confidence in our data increases.

Six different ways to sample a population:

ACCIDENTAL	Man-on-the street method — weakest method with many sources of bias
REPUTATIONAL	Selection depends on someone's judgement of who is and who isn't a "typical" representative of the population (e.g., surveying agency directors about their organisation)
RANDOM	Every output or outcome has an equal chance of being selected and included in the sample. Clients are picked through a random procedure.
STRATIFIED	If you want to make sure certain sub-groups are included then sample within each group (e.g., assessing needs of organisation make sure you have programme staff, clerical staff, secretarial staff, etc.)
CLUSTER	Used if population spread out in order to keep costs down if face-to-face interviews used (e.g., randomly select two programs and survey in only these two)
QUOTA	Sometimes used if you have limited resources but want a "quick reading" of the population (e.g., say client population is 70% males and 30% females. A sample of 100 people could be drawn - 70/30.)

Sample size and power of your study

Few studies discuss how the number of patients who took part in the study was decided. Often this number seems to have been arrived at by some consideration of administrative convenience. However, using too many patients will almost certainly be a waste of time and money. On the other hand, too few will lead to insufficient power of a study. In order to understand the meaning of power, it is important to know the following statistical terms:

In every statistical analyses there are two hypotheses under consideration:

The **null hypotheses** assumes that the variable being investigated is without effect (e.g. treatment intervention will not affect the quantity and frequency of alcohol consumption) and the data you are collecting are given the opportunity to disapprove this assumption. If you can reject the null hypotheses, then your **research hypothesis** which is the reason that the investigators did the study is accepted (e.g. treatment intervention leads to reduction of alcohol consumption). The basic overall principle is that the research hypothesis is considered false until demonstrated beyond reasonable doubt to be true. What is considered a "reasonable " doubt is called the **alpha error** or **significance level**. By convention in scientific research, the alpha error or level of remaining doubt is one below either 5% or 1%.

The error in rejecting the **research hypothesis** when true is called beta error. This value should be as small as possible. The value 1-beta is called power and we wish that value to be as large as possible. **Power** is the probability of obtaining a significant result (i.e. rejecting the null hypothesis) on the assumption that your research hypothesis is true. In other terms, power is the sensitivity of the study to detect clinically important differences or effects.

Researchers are very good at dealing with methodological and procedural issues in their investigation. However, they quite frequently fail to address power and sample size: On the assumption that the research hypothesis is true, it is necessary that the investigators determine the sample size needed to give this hypothesis a reasonable chance of being proven correct. A **power analysis** of one's planned investigation provides the researcher with the number of subjects needed (N) to achieve a desired power of the study or estimates the power of the study if a sample size is already given. For the calculation of N or power you need to know the following components:

1. Sample size of your study
2. Power or the odds that you will observe a treatment effect when it occurs.
3. **Critical Effect size**: This is a measure of how strong a theory must minimally be to be "important to society". The specification of this measure is based on your research design and your population characteristics under investigation. A statistician will assist you in specifying this measure.
4. Alpha error (, or significance level) or the odds that the observed result is due to chance.

Given values for any three of these components, it is possible to compute the value of the fourth. For instance you might want to determine what a reasonable sample size would be for your study. A statistician will be able to calculate the power or sample size of your study, if you provide him with these components, which have to be assumed in advance of your study. They depend on your research design and on the statistical analysis you plan to conduct. The statistician will also assist you with estimating the components (e.g. effect size), if you explain him your design.

To give you an example of what information is necessary in order to calculate the power of a study, imagine the following scenario. You want to find out if a group of methadone maintenance clients consume on average significantly less units of alcohol after receiving a certain treatment (e.g. motivational interviewing) compared to a group of methadone maintenance clients who did not receive motivational interviewing. The statistician needs to know:

- Alpha level = .05. You want to be 95% certain that your observed difference are not due to chance.
- Effect size = .833
- Sample size per group, e.g. N = 5

With this information you statistician will tell you that you have only a 42% chance of detecting a clinical important difference, assuming it exists (= power of your study). However, a rule of thumb in social research is that you want a statistical power of at least 80%. That is, you want to have at least 80 chances out of 100 of finding an effect when there is one. All mentioned factors in this example are interrelated with power. Changes in any one of these factors will lead to a change in power. You can also calculate what sample size you will need in order to have a power of 80%. In our example, a power of 82% would be achieved if the number of clients per group was increased to 20. Again, you should consult a statistician to assist you with calculating the sample size for your study. There are also specialised software programmes which can conduct a power analysis (e.g. nQuery Advisor 2.0).

Into action

The next step was to determine the sample and sample size of the two evaluation projects. The number of clients for the process evaluation would be determined by the number of clients who entered the treatment for the present and previous year.

All clients who enter the treatment an initial assessment form and termination form is completed and this data is entered to the main database of the programme. The number of clients for the outcome evaluation would be determined by the number of clients who complete the AIDS education intervention during the year of data collection and who would be able follow up after a period of 6 months. All clients who enter the treatment have to participate in the AIDS education intervention. Approximately 430 clients attend the service every year. If it is assumed that only 60% of clients will be able to follow-up after 6 months, 270 clients are expected to participate in the outcome evaluation. Adam S. conducted a power analysis, which revealed that the study would have a power of over 80% if this sample size was assumed.

It's your turn (7 C)

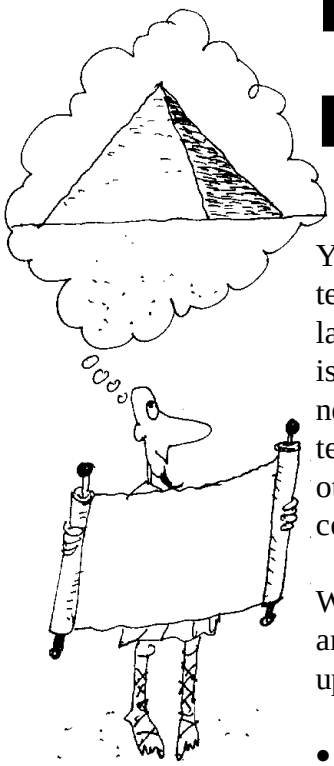
1 Using the appropriate specialised workbook as a guide, determine the sampling strategy that you will use to collect your data.

2 Review your sampling strategy in light of your original research questions. Are you addressing your questions adequately?

3 Review your sampling strategy with the expected users of the research. Revise as needed.

Step 8

Ensure that your research resources are sufficient



You have already considered issues of material, expertise, and time resources. In this last task, it is essential to summarise the issues and make final decisions. It is also necessary to record the decisions in a written evaluation plan that also summarises other key elements of your planning process.

With respect to data collection, analysis and reporting, it is necessary to decide upon:

- the personnel needed, including the use of consultants, and the skills and expertise needed
- fees, salaries, etc. for consultants and extra staff
- non-salary costs such as printing questionnaires or standard forms, data entry, editing and production of reports
- time and expertise required

Ask yourself: is your research project realistic? Do you have the necessary resources to be successful? If not, it may be necessary to re-evaluate your plans.

Into action

As a last step, the team discussed their research resources. With respect to data collection, analysis and reporting they needed the following resources:

1 Personnel needed

Planning the evaluations (Evaluation team)
Sue R.: Psychiatrist
Adam S.: Scientific researcher
Chris C.: Drug Worker

Data collection
Programme secretary for entering intake and termination records
Baseline interview for evaluating AIDS education intervention
Conducted by clinical staff (drug workers)

Follow-up interview
Conducted by clinical staff (drug workers)

Supervision of interviews
Chris C.: Drug Worker

Data entry for outcome evaluation
Adam S., scientific researcher

Data analysis
Adam S.: scientific researcher

Supervision of implementing the outcome evaluation
Sue R.: Psychiatrist

2 Salaries

Adam S. would be employed to assist the evaluation.

3 Non-salary costs

Cost for printing questionnaires, follow-up telephone interviews, editing and printing reports

4 Time

In order to plan and implement both process and outcome evaluation, a total of 18 months would be required.

It's your turn

1 Retrieve the following "It's Your Turn" exercises from this workbook:

Step 2
Step 3
Step 4 (A)
Step 4 (B)
Step 5
Step 6 (B)
Step 6 (F)
Step 7 (B)
Step 7 (C)

These exercises constitute the bulk of the planning work you have completed so far. Get all the exercises together, and then review them step-by-step. Ask yourself: do they create a reasonable, well-organised plan?

2 Pay particular attention to the expected time, costs, and expertise needed to complete your evaluation. Are your resources (Step 2) sufficient? If not, return to Step 4 and reformulate your plans.

Appendix 1

Focus group method

The focus group is a method for collecting data on a specific group or population. It is useful for:

- assessing a situation
- determining the needs and attitudes of that population
- planning appropriate interventions and responses. It is easy to organise and can provide quick and relevant answers to specific questions.

A focus group consists of a small group of participants, preferably 6-10 people. It is co-ordinated by a facilitator who can propose open-ended questions on a chosen topic (like tobacco use, drug injection — see some suggestions below in relation to PSU). These discussions then can serve as the basis for developing a storyline on the specific topic that is appropriate for the target group (for example, adolescent PS users) or to deciding on a health message.

The storyline should include some of the problems and possible solutions raised by the group. It can then be used for the production of educational materials, using a variety of techniques, as described in the various sections of this guide.

Depending on the activity being planned (for example a TV commercial or a play), it will be necessary to hold a number of focus groups to obtain the desired information, sometimes using the same group of participants to discuss other issues on the same topic, and sometimes different groups of participants to discuss different views on the same topic.

Decide what you want to know

Before convening a focus group, the organisation (for example, the school, community centre, non-governmental organisation, or health care facility) in association with the facilitator, must determine the nature and extent of the information that they require.

Lists should be prepared of the questions and issues that need to be addressed. The list will remind the facilitator during the discussion of all the aspects that need to be discussed.

At this point it is a good idea to consult with a few people from the target population you will be working with, whom you already know. Show them your plan and the list of questions, and ask them whether they think the questions are relevant and appropriate.

Identify the participants

Participation in the focus group discussion should be voluntary. One good way to get a mixture of participants to attend a focus group is to use a technique called "snowballing." Ask two to three people whom you already know to introduce you to some of their friends. Then ask these new participants to introduce you to other children whom the first group does not know.

Depending on the objective of your activity and the issues you will be discussing, there should be separate focus groups for girls and boys.

The type of participant you select will determine many other aspects of the focus group. For example, the facilitator many need to be more active and to reduce the number of questions if the invited participants include people who regularly use PS.

In order to attract individuals to attend the discussion, you could consider offering an incentive. For example, you could provide food during, or at the end of, the session. This may, or may not, be a good idea, depending on the group expectations and any future activities you are planning.

The facilitator

The facilitator should have some training and experience in group activities and also be familiar with the problems related to PSU, particularly with regards to the target group or community with whom he or she is working. The facilitator must be familiar with the dialect or slang of the group.

It is also important that the members of the focus group feel that the facilitator cares about them and their problems. The main tasks of the facilitator are:

- to create an atmosphere of safety
- to help the group focus its energy on the task by suggesting methods and procedures
- to make sure all the members of the group have an opportunity to participate
- to help participants present their ideas to the group
- to protect members of the group from personal attacks, put-downs, and criticism

Depending on the activity you are planning and the characteristics of the group, you may consider having two facilitators if the group has eight or more participants.

Select a documenter

You need a documenter to record the discussion of the group. This person should always:

- ask the permission of the group to record the discussion, and any other information
- record the discussion accurately (writing and/or recording, photographing, filming, etc.) and keep it confidential
- take care not to influence the record with her or his own opinions
- be familiar with the dialects and slang used by the group
- observe and record non-verbal information about the group discussion (for example, the emotional tone of the discussion, important hand gestures, unusual behaviour)

Confirm attendance and follow-up meetings

The facilitator should:

- keep in contact with the participants until the time comes when he or she can confirm their attendance
- select an appropriate meeting place, date, time
- arrange all the materials necessary for the discussion, including those specific for the creative workshop
- inform the participants about the time and place of any follow-up meetings

Conducting a focus group discussion

The facilitator should:

- arrive at the location ahead of time, check that everything is in order, including tape recorders, batteries, etc., as required
- welcome the participants
- introduce himself/herself (and others assisting) and explain what will be happening, and who is recording the proceedings (asking their permission to do so)
- arrange an introductory activity to help participants get to know each other and relax before they start discussing the actual topic; it could be a song, a prayer, or a brief game, including an opportunity for people to introduce themselves
- start the discussion by stating the general purpose of the activity and the primary topic of the focus group
- explain the procedure of the discussion
- ask the group for questions, suggestions, and their expectations
- begin the discussion with a general, open-ended question about the topic

Facilitate the dialogue

The facilitator must pay attention to the process as well as the content of the discussion. The process includes issues such as:

1. who speaks and who does not,
2. what topics are avoided,
3. what issues upset the group,
4. whether the pace of the discussion is slow or quick, and
5. how the participants interact with the facilitator.

The facilitator should also:

- encourage participants to share as much information and as many insights as possible
- try to maintain an atmosphere in which people take each other seriously, but humour is still welcome, and help to make it safe for people to share the feelings behind their opinions
- show genuine interest in everything that is said, and comment on special contributions of members and on accomplishments of the group

- deal politely with irrelevant information
- encourage the expression of different viewpoints; the more important the decision, the more important it is to have all relevant facts, feelings, and opinions
- take care not to judge responses nor make long comments
- control over-talkative members and ask for comments from quiet ones
- use different kinds of questions to increase participation and interest; search for all possible answers to a problem by changing the perspective of the discussion
- keep the discussion focused on the subject
- summarise the discussion at frequent intervals
- try to hold the participants attention by, taking a break, stretching, switching seats, saying something humorous, or playing a brief game
- postpone a scheduled break if the group is absorbed in its work
- defuse personal arguments between members
- keep the emotional atmosphere of the discussion at a level that can be tolerated by all the participants; if any of the members becomes too distressed, consider addressing his or her feelings immediately, or letting the whole group take a break
- as the discussion continues, check that all the issues you listed have been properly covered

Conclude the focus group discussion

Towards the end of the session, the facilitator should restate the objectives of the discussion, and try to pull together the main points made by the participants.

The facilitator should also:

- ask the participants whether the discussion has missed any important issues or questions
- express sincere appreciation for the participants' attention, time, and contributions
- inform the participants of subsequent activities, if any
- end the focus group discussion with a feeling of togetherness — sing a song, shake hands, or do a similar activity that affirms the group and puts a sense of closure to the time spent together

Appendix 2
Sample questionnaires

The following pages contain a small sampling of data collection instruments that can be used in PSU evaluation, as well as the World Health Organization's ICD-10 diagnostic criteria for substance use disorders.

The EuropASI is reprinted here with permission of Dr. Anna Kokkevi of the University Mental Health Research Institute (UHMRI), Athens, Greece. For more information about this instrument, see Kokkevi, A. and Hartgers, C. (1995). EuropASI: European adaptation of a multidimensional assessment for drug and alcohol dependence. European Addiction Research, 1:208-210.

The AUDIT was developed by the World Health Organization and is used widely to assess alcohol use patterns.

The brief instruments measuring psychoactive substance use, risk behaviour and health and correctional service utilisation and client motivation are adapted from a data collection protocol for treatment process and outcome monitoring being developed by the Addiction Research Foundation, Ontario, Canada. Information about these instruments can be obtained from: Addiction Research Foundation, 100 Collip Circle, Suite 200, London, Ontario, Canada, N6G 4X8. These measures are presented as examples only. Reliability and validity data are not available. You must decide on their appropriateness and availablility for your clients and your culture. In addition to considering these instruments, a review of the ARF Outcome Measures Directory (undated) is highly recommended. This Directory contains many potentially useful instruments for process evaluation and discusses reliability, validity, and practical issues in administration.

In this appendix you will also find a sample consent form, copies of the EuroQol EQ-5D, the WHOQOL-BREF and scoring instruction as well as three examples of questionnaires that can be used to assess client satisfaction. Remember, it is up to you to select the instrument(s) that are most appropriate to your evaluation.

ICD-10 Diagnostic criteria for substance use disorders

The consumption of psychoactive substances can lead to a range of problems that affect the health or social status of the user. Many of these problems are defined in the **International Statistical Classification of Diseases and Related Health Problems, Tenth Revision** (ICD-10). ICD-10 disorders are not ranked in terms of severity and are not mutually exclusive from one another. In other words, a substance user may qualify for more than one ICD-10 disorder, and Dependence is not necessarily the most severe diagnosis.

Acute intoxication

Clear evidence of recent use of a psychoactive substance at sufficiently high dose levels to be consistent with intoxication.

Symptoms or signs of intoxication compatible with the known actions of the particular substance and of sufficient severity to produce disturbance in the level of consciousness, cognition, perception, affect, or behaviour that are of clinical importance.

Not accounted for by a medical disorder unrelated to the substance use, and not better accounted for by another mental or behavioural disorder.

Harmful use

Clear evidence that substance use is responsible for (or substantially contributed to) physical or psychological harm, including impaired judgement or dysfunctional behaviour, which may lead to disability or have adverse consequences for interpersonal relationships.

The nature of the harm is clearly identifiable.

The pattern of substance use has persisted for at least 1 month or has occurred repeatedly within a 12-month period.

The disorder does not meet the criteria for any other mental or behavioural disorder related to the same substance in the same time period (except for acute intoxication).

Dependence

At least 3 symptoms from the following:

- A strong desire or sense of compulsion to take the substance.
- Impaired capacity to control substance-taking behaviour in terms of onset, termination, or levels of use.
- A physiological withdrawal state when substance use is reduced or ceased.
- Evidence of tolerance to the effects of the substance.
- Preoccupation with substance use.
- Persistent substance use despite clear evidence of harmful consequences.
- The pattern of substance use has persisted for at least 1 month or has occurred repeatedly within a 12-month period.
- The symptoms have occurred within the same period of time.

Psychoactive substance use

Substance	Average quantity per day of use in past 90 days*	Used in past 12 months (1 = Yes / 2 = No)	Number of days used in past 90 (Days)	Use currently a problem? (1 = Yes / 2 = No)
Alcohol (beer, liquor, wine)				
Cocaine/ crack/ coke				
Amphetamines/ other stimulants				
Cannabis (hash, weed, grass, pot, marijuana)				
Benzodiazepines				
Barbiturates				
Heroin/ opium				
Prescription opioids				
Over-the-counter codeine preparations				
Hallucinogens				
Glue/ other inhalants				
Tobacco				
Other psychoactive substances				

* It may be difficult to quantify the exact amount for certain substances. Indirect estimates can be made from the number of times per day a substance is injected, inhaled, snorted, or smoked.

Risk behaviour

1 Thinking about your use of psychoactive substances, have you:

☐ Never injected ☐ Injected prior to one year ago

☐ Injected in the last 12 months ☐ Unknown

If ever injected, answer the following questions:

i) During the past 90 days, on how many days did you inject any kind of psychoactive substance? ☐ days

ii) Have you ever shared a needle, syringe, cooker/spoon or cotton/filter with anyone at any time in your life?

☐ Yes ☐ No ☐ No response

If **Yes**, during the past 90 days, on how many days did you share a needle, syringe, cooker/spoon or cotton/filter with anyone? ☐ days

During the past 90 days, with how many people have you shared? ☐ people

2 How often do you use condoms with your sexual partner or partners?

☐ Never ☐ Sometimes ☐ Always

During the past 90 days, how many times have you had unprotected sex? ☐ times

3 During the past 90 days, on how many days have you driven a motor vehicle or used a machine at the workplace while under the influence of alcohol or other psychoactive substances? ☐ days

Health and correctional service utilisation

1 Thinking about physical health problems, during the past 90 days, how many:

- times have you had to go to the emergency room ☐ times
- nights total did you spend in the hospital ☐ nights
- times did you have an outpatient surgical procedure ☐ times
- times did you see a doctor in an office or outpatient clinic ☐ times

2 a) Thinking about mental health problems, during the past 90 days, how many:

- times have you had to go to the emergency room ☐ times
- nights total did you spend in the hospital ☐ nights
- times did you see a doctor in an office or outpatient clinic ☐ times

b)Are you currently in any type of treatment or counselling for mental or emotional problems?

☐ Yes ☐ No ☐ No response

3 Over the last 90 days, how many days have you received alcohol or substance use treatment at the following places?

- a hospital overnight for withdrawal or related problems ☐ days
- an inpatient substance use treatment facility (3 -90 days) ☐ days
- a long-term (3 to 12 months) residential program or therapeutic community for substance use disorder treatment ☐ days
- a methadone or other opioid treatment program ☐ days
- an assessment or outpatient substance use treatment facility ☐ sessions
- a mental health centre or facility as an outpatient ☐ sessions
- an employee assistance program ☐ sessions
- a family and/or marital counselling service ☐ sessions
- an emergency room ☐ days
- a private doctor's office ☐ visits
- a prison or jail ☐ days
- some other place (please describe _____) ☐ days

4 a) How many self-help meetings, (e.g., AA, NA, ACOA) have you attended for your substance use problem in the past 90 days? ☐ meetings

b) How many self-help meetings have you attended for issues other than substance use problems in the past 90 days? ☐ meetings

5 a) During the past 90 days, how many days have you been on probation or parole or been in jail or custody?

- Probation ☐ days
- Parole ☐ days
- Jail/prison/closed custody ☐ days
- Open custody ☐ days

b) During the past 90 days, how many times have you been charged for breaking the law (please do not count minor traffic violations)? ☐ times

Were you charged with:	Please check if Yes	# of charges in the last 90 days
• driving while impaired	☐	☐
• drunkenness or other liquor law violation	☐	☐
• possession, distribution, or sale of illegal substances	☐	☐
• sexual assault	☐	☐
• theft (including B&E, theft over and theft under)	☐	☐
• violence against family or others	☐	☐
• major crime	☐	☐
• Other (please describe _____)	☐	☐

Client Motivation

INSTRUCTIONS: Please indicate whether you agree or disagree with each of the following statements by placing the number that best reflects your own personal opinion in the blank provided. Remember, there are no right or wrong answers, and your responses are completely confidential.

Treatment entry questionnaire:
Use the following scale to make your ratings

		Strongly Disagree 1 2 3 4 5 6 7 Strongly Agree
1	If I remain in treatment it will probably be because I feel that it's the best way to help myself.	1 2 3 4 5 6 7
2	I plan to go through with a treatment program because I'll hate myself if I don't get my habit under control.	1 2 3 4 5 6 7
3	I have agreed to follow a treatment program because I was referred for treatment by the legal system.	1 2 3 4 5 6 7
4	I plan to go through with a treatment program because it's a challenge to learn how to live without misusing psychoactive substances.	1 2 3 4 5 6 7
5	I plan to go through with a treatment program because my friends and family won't approve of me unless I do.	1 2 3 4 5 6 7
6	Being in a program is a way for me to avoid getting punished for my behaviours.	1 2 3 4 5 6 7
7	I decided to enter a program because I was interested in getting help.	1 2 3 4 5 6 7
8	I decided to enter a program because I won't like myself very much unless my substance use problem is under control.	1 2 3 4 5 6 7
9	I had no choice about coming into a treatment program.	1 2 3 4 5 6 7
10	I plan to go through with the treatment program because having a substance use problem makes it hard for me to do things I want to do.	1 2 3 4 5 6 7

11 My family made sure that I entered a program. 1 2 3 4 5 6 7

12 If I remain in treatment it will probably be because others will be angry with me if I don't. 1 2 3 4 5 6 7

13 I decided to enter a program because I really want to make some changes in my life. 1 2 3 4 5 6 7

14 I have agreed to follow a program because I want others to see that I am really trying deal with my habit. 1 2 3 4 5 6 7

15 I plan to go through with treatment because I'll be ashamed of myself if I don't. 1 2 3 4 5 6 7

16 I decided to enter this program because no one other than myself can change the way I am. 1 2 3 4 5 6 7

17 The reason I am in treatment is because other people have pressured me into being here. 1 2 3 4 5 6 7

18 If I remain in treatment it will probably be because I'll feel like a failure if I don't. 1 2 3 4 5 6 7

19 I plan to go through with a treatment program because I'll get into trouble with the law if I don't remain in treatment. 1 2 3 4 5 6 7

20 I plan to go through with a treatment program because I have freely chosen to be here. 1 2 3 4 5 6 7

21 If I remain in treatment it will probably be because people will think I'm a weak person if I don't. 1 2 3 4 5 6 7

22 I decided to enter a program because it feels important for me personally to deal with my substance use problem. 1 2 3 4 5 6 7

23 I have agreed to follow a treatment program because I'll get in trouble with my friends and family if I don't follow all the guidelines. 1 2 3 4 5 6 7

24 I plan to go through with a treatment program because not having problems due to substances is a choice I really want to make. 1 2 3 4 5 6 7

25 My friends strongly pressured me to come into a program. 1 2 3 4 5 6 7

26 If I remain in treatment it will probably be because I'll feel very bad about myself if I don't. 1 2 3 4 5 6 7

27 I have agreed to follow the procedures of the treatment program because it's a personal challenge for me to deal with my problem. 1 2 3 4 5 6 7

28 I have agreed to follow a treatment program because I was pressured to come. 1 2 3 4 5 6 7

29 I decided to enter a program because people will like me better when I have dealt with my habit. 1 2 3 4 5 6 7

30 I was basically forced into a treatment program. 1 2 3 4 5 6 7

Sample Consent Form

PLEASE READ THIS CAREFULLY AND RETURN A SIGNED COPY TO YOUR COUNSELLOR. PLEASE KEEP THE SECOND COPY FOR YOUR OWN RECORDS.

This form deals with your consent to take part in a follow-up study conducted by _____.

The purpose of this study is to help evaluate the services provided by the program.

IF YOU ARE 16 YEARS OF AGE OR YOUNGER you may also wish to have your parent(s) or guardian(s) read this form and provide their written consent. If they have any questions regarding this study they should feel free to contact the staff of programme at... telephone no... during regular business hours.

In consenting to participate in this study I understand:

1 I will be contacted by mail or telephone in about 6 months by a follow-up worker to arrange a personal interview;

2 that at the interview I will be asked questions about my psychoactive substance use and other behaviours during the last six months;

3 that in the event the follow-up worker is unable to reach me at the telephone number or address given below, he/she may contact the following people to determine my whereabouts upon the condition the he/she does not reveal any details about my participation in the study or why he/she wishes to contact me;

Name of contact person	Area Code & Telephone No.	Relation
1. _____	_____	_____
2. _____	_____	_____

4 that the information given to the follow-up worker will be treated as confidential. It will not be shared with my assessment worker, any persons at the program, or any other agencies;

5 I will not be identified in any reports and all published reports based on this study will only refer to grouped data;

6 I reserve the right to decline the interview, or if I agree to the interview, I may refuse to answer specific questions or terminate the interview at any time.

7 also understand that my participating in the study does not promise any therapeutic benefit. If I decline to participate in the study or withdraw later, this will not affect the services I receive from the staff of the program.

I, (signature), (date) hereby consent to take part in the follow-up study as outlined above.

Please print:

_____	_____	
Name of Client	Address	
_____	_____	_____
Name of Witness	Date	Signature

PARENT OR GUARDIAN:

My signature, (date) will serve to acknowledge my having read this form and agree that my child/ward may take part in the follow-up study subject to the conditions described above.

_____	_____	_____
Name of Witness	Date	Signature

EXAMPLES 1-3: From Addiction Research Foundation

EXAMPLE 1

Examples 2 - 4 are adapted from a data collection protocol for treatment process and outcome monitoring being developed by the Addiction Research Foundation, Ontario, Canada. Information about the instruments can be obtained from: Addiction Research Foundation, 100 Collip Circle, Suite 200, London, Ontario, Canada, N6G 4X8.

There are no reliability or validity data of these instruments. However, they may be helpful to stimulate ideas for the development of a questionnaire unique to your needs. You must decide on their appropriateness and availability for your clients and your culture.

In addition to considering these instruments, a review of the ARF Outcome Measures Directory (undated) is highly recommended. This Directory contains many potentially useful instruments for process evaluation and discusses reliability, validity, and practical issues in administration.

Please check the box at the end of each question that best reflects your impression of detoxification:

	Never	Rarely	Sometimes	Often	Always	Not Applicable
1 The staff tried to understand my problems.	☐	☐	☐	☐	☐	☐
2 The information and advice that staff gave me were helpful.	☐	☐	☐	☐	☐	☐
3 The staff helped me feel better physically while drying out.	☐	☐	☐	☐	☐	☐
4 Information about me was kept confidential.	☐	☐	☐	☐	☐	☐
5 I was pleased with the way staff treated me.	☐	☐	☐	☐	☐	☐
6 During the day, I could see staff when I needed to.	☐	☐	☐	☐	☐	☐
7 At night and on weekends, I was able to see staff when I needed to.	☐	☐	☐	☐	☐	☐
8 The meals were satisfactory.	☐	☐	☐	☐	☐	☐
9 The detox was comfortable.	☐	☐	☐	☐	☐	☐
10 The detox was clean.	☐	☐	☐	☐	☐	☐
11 I felt safe in the detox.	☐	☐	☐	☐	☐	☐
12 The staff treated me with dignity and respect.	☐	☐	☐	☐	☐	☐
13 The staff were friendly and supportive.	☐	☐	☐	☐	☐	☐
14 There is enough space to be alone or to talk privately with staff.	☐	☐	☐	☐	☐	☐
15 The type of help I was looking for was provided.	☐	☐	☐	☐	☐	☐
16 The staff handled it well when someone had to be sent to hospital.	☐	☐	☐	☐	☐	☐
17 The staff handled it well when there fights or trouble.	☐	☐	☐	☐	☐	☐
18 The staff made me feel welcome.	☐	☐	☐	☐	☐	☐

EXAMPLE 2

Questions about whether client's needs were met. Questions about whether client's needs were met in a detoxification centre

1 What were your needs upon admission? Were these needs met?

	No	Yes		No	Partly	Yes
• food	☐	☐	If Yes	☐	☐	☐
• shelter	☐	☐	If Yes	☐	☐	☐
• personal safety	☐	☐	If Yes	☐	☐	☐
• support	☐	☐	If Yes	☐	☐	☐
• medical attention	☐	☐	If Yes	☐	☐	☐
• shower/bath	☐	☐	If Yes	☐	☐	☐
• other _____	☐	☐	If Yes	☐	☐	☐

2 Please check one or more answers to fill the space beside the question.

The most helpful information that I got was about:

☐ my feelings _____

☐ family or sexual problems _____

☐ my substance use problem _____

☐ housing _____

☐ treatment choices _____

☐ social agencies _____

☐ counselling _____

☐ health problems _____

☐ none of the above _____

Retrospective rating of usefulness of programme components

We recognise that many things could have happened to people over the two years of follow-up that had nothing to do with the treatment. The following questions largely concern the effects that specific features of the treatment had on you. Please answer so as to describe the effects the treatment had on you apart from other events that may have happened in your life.

EXAMPLE 3

Rate the value of each of the following features of treatment using the following 4-point scale:

		Not helpful	No opinion	Helpful	Very helpful
1	Assessment	1	2	3	4
2	Therapist	1	2	3	4
3	The readings (two handouts)	1	2	3	4
4	Homework assignments	1	2	3	4
5	Problem-solving approach (identifying triggers and consequences, evaluating options, using action plans)	1	2	3	4
6	Emphasis on doing it on my own (using my own strengths and resources)	1	2	3	4
7	Self-selection of substance use goals	1	2	3	4
8	Availability of further treatment if desired	1	2	3	4
9	Follow-up contacts	1	2	3	4
10	Follow-up worker	1	2	3	4

EUROQOL EQ-5D Scoring System

	DEDUCTION	
→	0.081	Constant deduction (for any dysfunctional state)
	0.269	N3 (level 3 occurs within at least one dimension)
		Mobility
→	0	1. No problems
	0.069	2. Some problems walking about
	0.314	3. Confined to bed
		Self-care
→	0	1. No problems
	0.104	2. Some problems walking or dressing self
	0.214	3. Unable to wash or dress self
		Usual activities
	0	1. No problems
→	0.036	2. Some problems with performing usual activities
	0.094	3. Unable to perform usual activities
		Pain/discomfort
	0	1. No pain or discomfort
→	0.123	2. Moderate pain or discomfort
	0.386	3. Extreme pain or discomfort
		Anxiety/depression
	0	1. Not anxious or depressed
	0.071	2. Moderately anxious or depressed
→	0.236	3. Extremely anxious or depressed

EuroQol scores are calculated by subtracting the relevant coefficients from 1.000. The **constant** term is used if there is any dysfunction at all. The **N3** term is used if there is any dimension at level 3.

For example, consider the person who responds 1,1,2,2,3 to each of the dimensions, respectively.

1.000	**Full health**
- 0.081	Constant term (for any dysfunction)
- 0	Mobility (level 1)
- 0	Self-care (level 1)
- 0.036	Usual activities (level 2)
- 0.123	Pain or discomfort (level 2)
- 0.236	Anxiety or depression (level 3)
- 0.269	N3 (level 3 within at least 1 dimension)
= 0.255	**Total Quality Adjusted Life Year (QALY)**

WHOQOL-BREF

Introduction, administration, scoring and generic version of the assessment

Field Trial Version December 1996

This manual was drafted by Alison Harper on behalf of the WHOQOL group. The WHOQOL group comprises a coordinating group, collaborating investigators in each of the field centres and a panel of consultants. Dr J. Orley directs the project. He has been assisted in this by Professor M. Power, Dr W. Kuyken. Professor N. Sartorius, Dr M. Bullinger and Dr A. Harper. The field centres involved in initial piloting of the WHOQOL were: Professor H. Herrman, Dr H. Schofield and Ms B. Murphy. University of Melbourne, Australia; Professor Z. Metelko, Professor S. Szabo and Mrs M. Pibernik-Okanovic, Institute of Diabetes, Endocrinology and Metabolic Diseases and Department of Psychology, Faculty of Philosophy, University of Zagreb, Croatia; Dr N. Quemada and Dr A. Carla, INSERM, Paris, France; Dr S. Rajkumar and Mrs Shuba Kumar. Madras Medical College, India; Dr S. Saxena and Dr K. Chandiramani, All India Institute of Medical Sciences, New Delhi, India; Dr M. Amir and Dr D. Bar-On, Ben-Gurion University of the Negev, BeerSheeva, Israel; Dr Miyako Tazaki, Department of Science, Science University of Tokyo, Japan and Dr Ariko Noji, Department of Community Health Nursing, St Luke's College of Nursing, Japan; Dr G. van Heck and Mrs J. De Vries, Tilburg University, The Netherlands; Professor J. Arroyo Sucre and Professor L. Picard-Ami, University of Panama, Panama; Professor M. Kabanov, Dr A. Lomachenkov and Dr G. Burkovsky, Bekhterev Psychoneurological Research Institute. St. Petersburg, Russia; Dr R. Lucas Carrasco, University of Barcelona. Spain; Dr Yooth Bodharamik and Mr Kitikorn Meesapya, I ltitute of Mental Health, Bangkok, Thailand; Dr S. Skevington, University of Bath, United Kingdom; Professor D. Patrick. Ms M. Martinand, Ms D. Wild, University of Washington, Seattle. USA and; Professor W. Acuda and Dr J. Mutambirwa, University of Zimbabwe, Harare. Zimbabwe.

New centres using the field version of the WHOQOL-100 are: Dr S. Bonicato, FUNDONAR, Fundacion Oncologica Argentina, Argentina; Dr A.E. Molzahn, University of Victoria, Canada; Dr G. Yongping, St Vincent's Hospital, Victoria, Australia; Dr G. Page. University of Quebec at Rimouski, Canada; Professor J. Fang, Sun Yat-Sen University of Medical Sciences, People's Republic of China; Dr M. Fleck, University of the State of Rio Grande do Sul, Brazil; Professor M.C. Angermeyer, Dr R. Kilian, Universitätsklinikum Klinik and Poliklinik fur Psychiatrie. Leipzig. Germany; Mr L. Kwok-fai, Queen Elizabeth Hospital. Kowloon, Hong Kong; Dr B.R. Hanestad. University of Bergen, Norway; Dr M.H. Mubbashar, Rawalpindi General Hospital, Pakistan; Dr J. Harangozo, Semelweis University of Medicine, Budapest & Dr L. Kullman, National Institute of Mental Rehabilitation, Budapest, Hungary; Professor I. Wiklund. Health Economics & Quality of Life, Astra Hassle AB. Sweden; Dr C. Fidaner, Dr Behçet Uz Paediatric Hospital, Balçova/lzmir, Turkey; Dr G. de Girolamo. Servizio Salute Mentale USL 27, Italy; Professor P. Bech, Frederiksborg General Hospital, Denmark; Dr R.S. Pippalla, Howard University, College of Pharmacy and Pharmaceutical Sciences, Washington, DC, USA and Dr H. Che Ismail, School of Medical Sciences, Kelantan, Malaysia.

Further information can be obtained from:
Dr John Orley
Programme on Mental Health
World Health Organization
CH-1211 Geneva 27, Switzerland

Introduction

The WHOQOL-100 quality of life assessment was developed by the WHOQOL Group with fifteen international field centres, simultaneously, in an attempt to develop a quality of life assessment that would be applicable cross-culturally. The development of the WHOQOL-100, has been detailed elsewhere (i.e. Orley & Kuyken, 1994; Szabo, 1996; WHOQOL Group 1994a, 1994b, 1995). This document gives a conceptual background to the WHOQOL definition of quality of life and describes the development of the WHOQOL-BREF, an abbreviated version of the WHOQOL-100. It also includes a generic English language version of the WHOQOL-BREF, instructions for administering and scoring, and proposed uses for this short form of the WHOQOL.

Rationale for the development of the WHOQOL-100

WHO's initiative to develop a quality of life assessment arose for a number of reasons. In recent years there has been a broadening in focus in the measurement of. health, beyond traditional health indicators such as mortality and morbidity (e.g. World Bank, 1993; WHO, 1991), to include measures of the impact of disease and impairment on daily activities and behaviour (e.g. Sickness Impact Profile; Bergner, Bobbitt, Carter et al, 1981), perceived health measures (e.g. Nottingham Health Profile; Hunt, McKenna and McEwan, 1989) and disability / functional status measures (e.g. the MOS SF-36, Ware et al, 1993). These measures, whilst beginning to provide a measure of the impact of disease, do not assess quality of life per se, which has been aptly described as "the missing measurement in health" (Fallowfield, 1990). Second, most measures of health status have been developed in North America and the UK, and the translation of these measures for use in other settings is time-consuming, and unsatisfactory for a number of reasons (Sartorius and Kuyken, 1994; Kuyken, Orley, Hudelson and Sartorius, 1994). Third, the increasingly mechanistic model of medicine, concerned only with the eradication of disease and symptoms, reinforces the need for the introduction of a humanistic element into health care. By calling for quality of life assessments in health care, attention is focused on this aspect of health, and resulting interventions will pay increased attention to this aspect of patients' well-being. WHO's initiative to develop a quality of life assessment arises from a need for a genuinely international measure of quality of life and a commitment to the continued promotion of an holistic approach to health and health care.

Steps in the development of the WHOQOL-100

The WHOQOL-100 development process consisted of several stages. These are explained in brief within this document. For a detailed description, the reader is referred to the WHOQOL Group (1994a, 1994b, in preparation). In the first stage, concept clarification involved establishing an agreed upon definition of quality of life and an approach to international quality of life assessment.

Quality of life is defined as individuals' perceptions of their position in life in the context of the culture and value systems in which they live and in relation to their goals, expectations, standards and concerns.

This definition reflects the view that quality of life refers to a subjective evaluation which is embedded in a cultural, social and environmental context. Because this definition of quality of life focuses upon respondents' "perceived" quality of life, it is not expected to provide a means of measuring in any detailed fashion symptoms, diseases or conditions, but rather the effects of disease and health interventions on quality of life. As such, quality of life cannot be equated simply with the terms "health status", "life style", "life satisfaction", "mental state" or "well-being". The recognition of the multi-dimensional nature of quality of life is reflected in the WHOQOL-100 structure.

In the second stage of development, exploration of the quality of life construct within 15 culturally diverse field centres was carried out to establish a list of areas/facets that participating centres considered relevant to the assessment of quality of life. This involved a series in meetings of focus groups which included health professionals, patients and well subjects. A maximum of six specific items for exploring each proposed facet were generated by each centre's focus group. To enable the

collaboration to be genuinely international the 15 field centres were selected world-wide to provide differences in level of industrialisation, available health services, and other markers relevant to the measurement of quality of life (e.g. role of the family, perception of time, perception of self, dominant religion).

In the third stage of development, questions from each centre were assembled into a global pool. After clustering semantically equivalent questions, 236 items covering 29 facets were included in a final assessment. Pilot work involved administration of this standardised assessment to at least 300 respondents within each centre.

Following field testing in these 15 centres, 100 items were selected for inclusion in the WHOQOL-100 Field Trial Version. These included four items for each of 24 facets of quality of life, and four items relating to the 'overall quality of life and general health' facet (see Table 1). The method by which these 100 items were selected is fully documented elsewhere (The WHOQOL Group, in preparation). The WHOQOL-100 Field Trial Version is currently being tested in new centres worldwide (these centres are outlined on page 6 of this document). The initial conceptual framework for the WHOQOL-100 proposed that the 24 facets relating to quality of life should be grouped into 6 domains. Recent analysis of available data, using structural equation modelling, has shown a four domain solution to be more appropriate. For a more detailed explanation of this, the reader is referred to The WHOQOL Group (in preparation). The WHOQOL-BREF is therefore based on a four domain structure (see Table 1).

Table 1: WHOQOL-BREF domains

Domain	Facets incorporated within domains
1 Physical health	Activities of daily living Dependence on medicinal substances and medical aids Energy and fatigue Mobility Pain and discomfort Sleep and rest Work Capacity
2 Psychological	Bodily image and appearance Negative feelings Positive feelings Self-esteem Spirituality / Religion / Personal beliefs Thinking, learning, memory and concentration
3 Social relationships	Personal relationships Social support Sexual activity
4 Environment	Financial resources Freedom, physical safety and security Health and social care: accessibility and quality Home environment Opportunities for acquiring new information and skills Participation in and opportunities for recreation / leisure activities Physical environment (pollution / noise / traffic / climate) Transport

Development of the WHOQOL-BREF

The WHOQOL-100 allows detailed assessment of each individual facet relating to quality of life. In certain instances however, the WHOQOL-100 may be too lengthy for practical use. The WHOQOL-BREF Field Trial Version has therefore been developed to provide a short form quality of life assessment that looks at Domain level profiles, using data from the pilot WHOQOL assessment and all available data from the Field Trial Version of the WHOQOL-100. Twenty field centres situated within eighteen countries have included data for these purposes (see Table 2). The WHOQOL-BREF contains a total of 26 questions. To provide a broad and comprehensive assessment, one item from each of the 24 facets contained in the WHOQOL-100 has been included. In addition, two items from the Overall quality of Life and General Health facet have been included.

Table 2: Centres included in development of the WHOWL-BREF

Centres In the pilot version of the WHOQOL	Centres In the field trial of the WHOQOL-100
Bangkok, Thailand	Bangkok, Thailand
Beer Sheva, Israel	Beer Sheva, Israel
Madras, India	Madras, India
Melbourne, Australia	Melbourne, Australia
New Delhi, India	New Delhi, India
Panama City, Panama	Panama City, Panama
Seattle, USA	Seattle, USA
Tilburg, The Netherlands	Tilburg, The Netherlands
Zagreb, Croatia	Zagreb, Croatia
Tokyo, Japan	Tokyo, Japan
Harare, Zimbabwe	Harare, Zimbabwe
Barcelona, Spain	Barcelona, Spain
Bath, UK	Bath, UK
St Petersburg, Russia	Kowloon, Hong Kong
Paris, France	Leipzig, Germany
	Mannheim, Germany
	La Plata, Argentina
	Porto Alegre, Brazil

The WHOQOL-BREF is available in 19 different languages. The appropriate language version, and permission for using it, can be obtained from The WHOQOL Group, Programme on Mental Health, World Health Organisation, CH-1211 Geneva 27, Switzerland. Under no circumstances should the WHOQOL-BREF be used without consultation with The WHOQOL Group. A methodology has been developed for new centres wishing to develop a further language version of the WHOQOL-100 or the WHOQOL-BREF. This can be obtained from The WHOQOL Group, Programme on Mental Health, World Health Organisation, CH-1211, Geneva 27, Switzerland.

Questions should appear in the order in which they appear in the example WHOQOL-BREF provided within this document, with instructions and headers unchanged. Questions are grouped by response format. The equivalent numbering of questions between the WHOQOL-BREF and the WHOQOL-100 is given in the example version of the WHOQOL-BREF to enable easy comparison between responses to items on the two versions. The WHOQOL-100 field test permit-

ted centres to include national items or facets that were thought to be important in assessing quality of life. Where centres wish to include additional national items or modules to the WHOQOL-BREF, these should be included on a separate sheet of paper and not scattered amongst the existing 26 items. There are three reasons for this:

1) To control for item order effects which could occur and change item meaning.

2) The WHOQOL-BREF represents an agreed upon core set of international items.

3) The WHOQOL-BREF is likely to be used where quality of life is amongst one of several parameters being assessed. Therefore additional national information can be obtained by including additional modules and measures

Administration of the WHOQOL-BREF

For any new centre not previously involved in either the development or field testing of the WHOQOL-100, the procedure being followed to field test the WHOQOL-BREF should be identical to that used to field test the WHOQOL-100. The instrument should be piloted on at least 300 people. This figure is based on the required numbers of respondents needed for analysis of pilot data. The sample of respondents to whom the assessment should be administered ought to be adults, with 'adult' being culturally defined. While stratified samples are not essential, a sampling quota should apply with regard to:

- Age (50% = <45 years, 50% = 45 + years)
- Sex (50% = male, 50% = female)
- Health status (250 persons with disease or impairment; 50 well persons)

With respect to persons with disease or impairment, this group should contain a cross-section of people with varied levels of quality of life. One way of attempting this would be to include some people with quite severe and disabling chronic diseases, some people in contact with health facilities for more transient conditions, possibly some attending a family practitioner, and others who are in contact with the health service for reasons that are not likely to impinge upon their quality of life to any great extent. By sampling patients from a cross-section of primary care settings, hospitals and community care settings this could most likely be achieved.

The WHOQOL-BREF should be self-administered if respondents have sufficient ability: otherwise, interviewer-assisted or interview-administered forms should be used. Standardised instructions, given on the second page of the WHOQOL-BREF example assessment, should be read out to respondents in instances where the assessment is interviewer-administered.

For centres who have already participated in the development and field testing of the WHOQOL-100, the above option of testing the WHOQOL-BREF is preferred, but not imperative where specific studies of patient groups are planned.

Frame of reference and time frame

A time frame of two weeks is indicated in the assessment. It is recognised that different time frames may be necessary for particular uses of the instrument in subsequent stages of work. For example, in the assessment of quality of life in chronic conditions, such as arthritis, a longer time frame such as four weeks may be preferable. Furthermore, the perception of time is different within different cultural settings and therefore changing the time scale may be appropriate.

Proposed uses of the WHOQOL-100 and the WHOQOL-BREF

It is anticipated that the WHOQOL assessments will be used in broad-ranging ways. They will be of considerable use in clinical trials, in establishing baseline scores in a range of areas, and looking at changes in quality of life over the course of interventions. It is expected that the WHOQOL assessments will also be of value where disease prognosis is likely to involve only partial recovery or remission, and in which treatment may be more palliative than curative.

For epidemiological research, the WHOQOL assessments will allow detailed quality of life data to be gathered on a particular population, facilitating the understanding of diseases, and the development of treatment methods. The international epidemiological studies that would be enabled by instruments such as the WHOQOL-100 and the WHOQOL-BREF will make it possible to carry out multi-centre quality of life research, and to compare results obtained in different centres. Such research has important benefits, permitting questions to be addressed which would not be possible in single site studies (Sartorius and Helmchen, 1981). For example, a comparative study in two or more countries on the relationship between health care delivery and quality of life requires an assessment yielding cross-culturally comparable scores. Sometimes accumulation of cases in quality of life studies, particularly when studying rare disorders, is helped by gathering data in several settings. Multicentre collaborative studies can also provide simultaneous multiple replications of a finding, adding considerably to the confidence with which findings can be accepted.

In clinical practice the WHOQOL assessments will assist clinicians in making judgements about the areas in which a patient is most affected by disease, and in making treatment decisions. In some developing countries, where resources for health care may be limited, treatments aimed at improving quality of life through palliation, for example, can be both effective and inexpensive (Olweny, 1992). Together with other measures, the WHOQOL-BREF will enable health professionals to assess changes in quality of life over the course of treatment.

It is anticipated that in the future the WHOQOL-100 and the WHOQOL-BREF will prove useful in health policy research and will make up an important aspect of the routine auditing of health and social services. Because the instrument was developed. cross-culturally, health care providers, administrators and legislators in countries where no validated quality of life measures currently exist can be confident that data yielded by work involving the WHOQOL assessments will be genuinely sensitive to their setting.

Scoring the WHOQOL-BREF

The WHOQOL-BREF (Field Trial Version) produces a quality of life profile. It is possible to derive four domain scores. There are also two items that are examined separately: question 1 asks about an individual's overall perception of quality of life and question 2 asks about an individual's overall perception of their health. The four domain scores denote an individual's perception of quality of life in each particular domain. Domain scores are scaled in a positive direction (i.e. higher scores denote higher quality of life). The mean score of items within each domain is used to calculate the domain score. Mean scores are then multiplied by 4 in order to make domain scores comparable with the scores used in the WHOQOL-100. Explicit instructions for checking and cleaning data, and for computing domain scores, are given in Table 3. A method for the manual calculation of individual scores is given on page 1 of the WHOQOL-BREF assessment form. The method for converting raw scores to transformed scores when using this method is given in Table 4, on page 11 of these instructions. The first transformation method converts scores to range between 4-20, comparable with the WHOQOL-100. The second transformation method converts domain scores to a 0-100 scale.

Where more than 20% of data is missing from a assessment, the assessment should be discarded (see Step 4 in Table 3). Where an item is missing, the mean of other items in the domain is substituted. Where more than two items are missing from the domain, the domain score should not be calculated (with the exception of domain 3, where the domain should only be calculated if < 1 item is missing).

Any national items should be scored separately from the core 26 item of the BREF. During the analysis the performance of any national items will be examined for possible use in alter national studies. At this stage of field testing national and core items must not be mixed in administration or scoring of the BREF.

An SPSS syntax file that automatically checks, recodes data and computes domain scores may be obtained from Professor Mick Power, Department of Psychiatry, Royal Edinburgh Hospital, Morningside Park, Edinburgh, EH10 5HF (email: mj@srv2.med.ed.ac.uk; fax: + 131 447 6860)

Table 3: Steps for checking and cleaning data and computing domain scores
Steps

SPSS syntax	for carrying out data checking, cleaning and computing total scores
1 Check all 26 items from assessment have a range of 1-5	RECODE Q1 Q2 Q3 Q4 Q5 Q6 Q7 Q8 Q9 Q10 Q11 Q12 Q13 Q14 Q15 Q16 Q17 Q18 Q19 Q20 Q21 Q22 Q23 Q24 Q25 Q26 (1=1) (2=2) (3=3) (4=4) (5=5) (ELSE=SYSMIS). (This recodes all data outwith the range 1-5 to system missing).
2 Reverse 3 negatively phrased items	RECODE Q3 Q4 Q26 (1=5) (2=4) (3=3) (4=2) (5=1). (This transforms negatively framed questions to positively framed questions)
3 Compute domain scores	COMPUTE DOM1=MEAN.6(Q3,Q4,Q10,Q15,Q16,Q17,Q18)*4. COMPUTE DOM2=MEAN.5(Q5,Q6,Q7,Q11,Q19,Q26)*4. COMPUTE DOM3=MEAN.2(Q20,Q21,Q22)*4. COMPUTE DOM4=MEAN.6(Q8,Q9,Q12,Q13,Q14,Q23,Q24,Q25)*4 (These equations calculate the domain scores. All scores are multiplied by 4 so as to be directly comparable with scores derived from the WHOQOL-100. The '.6 in 'mean.6' specifies that 6 items must be endorsed for the domain score to be calculated).
4 Delete cases with >20% missing data	COUNT TOTAL=Q1 TO Q26 (1 THRU 5). (This command creates a new column 'total'. 'Total' contains a count of the WHOQOL-100 items with the values 1-5 that have been endorsed by each subject. The 'Q1 TO Q26' means that consecutive columns from 'Q1', the first item, to 'Q26', the last item, are included in the count. It therefore assumes that data is entered in the order given in the assessment). FILTER OFF. USE ALL. SELECT IF (TOTAL>=21). EXECUTE. (This second command selects only those cases where 'total', the total number of items completed, is greater or equal to 80%. It deletes the remaining cases from the data set).
5 Check domain scores	DESCRIPTIVES VARIABLES=DOM1 DOM2 DOW DOW /STATISTICS=MEAN STDDEV MIN MAX. (Running descriptives should display values of all domain scores within the range 4-20).
6 Save data set	Save data set with a new file name so that the original remains intact.

Table 4: Method for converting raw scores to transformed scores

DOMAIN 1		
Raw score	Transformed scores	
	4-20	0-100
7	4	0
8	5	6
9	5	6
10	6	13
11	6	13
12	7	19
13	7	19
14	8	25
15	9	31
16	9	31
17	10	38
18	10	38
19	11	44
20	11	44
21	12	50
22	13	56
23	13	56
24	14	63
25	14	63
26	15	69
27	15	69
28	16	75
29	17	81
30	17	81
31	18	88
32	18	88
33	19	94
34	19	94
35	20	100

DOMAIN 2		
Raw score	Transformed scores	
	4-20	0-100
6	4	0
7	5	6
8	5	6
9	6	13
10	7	19
11	7	19
12	8	25
13	9	31
14	9	31
15	10	38
16	11	44
17	11	44
18	12	50
19	13	56
20	13	56
21	14	63
22	15	69
23	15	69
24	16	75
25	17	81
26	17	81
27	18	88
28	19	94
29	19	94
30	20	100

DOMAIN 3		
Raw score	Transformed scores	
	4-20	0-100
3	4	0
4	5	6
5	7	19
6	8	25
7	9	31
8	11	44
9	12	50
10	13	56
11	15	69
12	16	75
13	17	81
14	19	94
15	20	100

DOMAIN 4		
Raw score	Transformed scores	
	4-20	0-100
8	4	0
9	5	6
10	5	6
11	6	13
12	6	13
13	7	19
14	7	19
15	8	25
16	8	25
17	9	31
18	9	31
19	10	38
20	10	38
21	11	44
22	11	44
23	12	50
24	12	50
25	13	56
26	13	56
27	14	63
28	14	63
29	15	69
30	15	69
31	16	75
32	16	75
33	17	81
34	17	81
35	18	88
36	18	88
37	19	94
38	19	94
39	20	100
40	20	100

References

Bergner, M., Bobbitt, R.A., Carter, W.B. et al. (1981). The Sickness Impact Profile: Development and final revision of a health status measure. Medical Care, 19, 787-805.

Fallowfield, L. (1990). The Quality of Life: The Missing Measurement in Health Care. Souvenir Press.

Hunt, S.M., McKenna, S.P. and McEwan, J. (1989). The Nottingham Health Profile. Users Manual. Revised edition.

Kuyken, W., Orley, J., Hudelson, P. and Sartorius, N. (1994). Quality of life assessment across cultures. International Journal of Mental Health, 23 (2), 5-27.

Olweny, C. L. M. (1992). Quality of life in developing countries. Journal of Palliative Care, 8, 25-30.

Sartorius, N. and Helmchen, H. (1981). Aims and implementation of multi-centre studies. Modern Problems of Pharmacopsychiatry, 16, 1-8.

Sartorius, N. and Kuyken, W. (1994). Translation of health status instruments. In J. Orley and W. Kuyken (Eds). Quality of Life Assessment: International Perspectives. Heidelberg: Springer Verlag.

Szabo, S. (1996). The World Health Organisation Quality of Life (WHOQOL) Assessment Instrument. In Quality of Life and Pharmaeconomics in Clinical Trials (2nd edition, Edited by Spilker B.). Lippincott-Raven Publishers, Philadelphia, New York.

The WHOQOL Group. The World Health Organization Quality of Life assessment (WHOQOL): position paper from the World Health Organization. Soc. Sci. Med., 41, 1403, 1995.

Ware, J. E., Snow, K., K., Kosinski, M. and Gandek, B. (1993). SF-36 Health Survey: Manual and Interpretation Guide. New England Medical Center, MA, USA.

World Bank. (1993). World Development Report: Investing in Health. New York: Oxford University Press.

World Health Organization. (1991). World Health Statistics Annual. Geneva: WHO.

The WHOQOL Group. (1994a). Development of the WHOQOL: Rationale and current status. International Journal of Mental Health, 23 (3), 24-56.

The WHOQOL Group. (1994b). The development of the World Health Organization quality of life assessment instrument (the WHOQOL). In J. Orley and W. Kuyken (Eds) Quality of Life Assessment: International Perspectives. Heidelberg: Springer Verlag.

The WHOQOL Group. (In preparation). The World Health Organisation Quality of Life Assessment (WHOQOL): Development and General Psychometric Properties.

WHOQOL-BREF

For office use only

	Equations for computing domain scores	Raw score	Transformed scores*	
			4-20	0-100
Domain 1	(6-Q3) + (6-Q4) + Q10 + Q15 + Q16 + Q17 + Q18 ☐ + ☐ + ☐ + ☐ + ☐ + ☐ + ☐	=		
Domain 2	Q5 + Q6 + Q7 + Q11 + Q19 + (6-Q26) ☐ + ☐ + ☐ + ☐ + ☐ + ☐	=		
Domain 3	Q20 + Q21 + Q22 ☐ + ☐ + ☐	=		
Domain 4	Q8 + Q9 + Q12 + Q13 + Q14 + Q23 + Q24 + Q25 ☐ + ☐ + ☐ + ☐ + ☐ + ☐ + ☐ + ☐	=		

* Please see Table 4 on page 72 of the manual, for converting raw scores to transformed scores.

ABOUT YOU

I.D. number ☐ ☐ ☐ ☐

Before you begin we would like to ask you to answer a few general questions about yourself: by circling the correct answer or by filling in the space provided.

What is your gender?

☐ Male ☐ Female

What is you date of birth?

☐ / ☐ / ☐ Day / Month / Year

What is the highest education you received?

☐ None at all ☐ Primary school ☐ Secondary school ☐ Tertiary

What is your marital status?

☐ Single ☐ Separated ☐ Married

☐ Divorced ☐ Widowed ☐ Living as married

Are you currently ill?

☐ Yes ☐ No

If something is wrong with your health what do you think it is?

☐ illness / ☐ problem

Instructions

This assessment asks how you feel about your quality of life, health, or other areas of your life. *Please answer all the questions.* If you are unsure about which response to give to a question, please *choose the one* that appears most appropriate. This can often be your first response.

Please keep in mind your standards, hopes, pleasures and concerns. We ask that you think about your life **in the last two weeks**. For example, thinking about the last two weeks, a question might ask:

	Not at all	Not much	Moder-ately	A great deal	Com-pletely
Do you get the kind of support from others that you need?	1	2	3	4	5

You should circle the number that best fits how much support you got from others over the last two weeks. So you would circle the number 4 if you got a great deal of support from others as follows.

	Not at all	Not much	Moder-ately	A great deal	Com-pletely
Do you get the kind of support from others that you need?	1	2	3	4	5

You would circle number 1 if you did not get any of the support that you needed from others in the last two weeks.

Please read each question, assess your feelings, and circle the number on the scale for each question that gives the best answer for you.

		Very poor	Poor	Neither poor nor good	Good	Very good
1 (G1)	How would you rate your quality of life?	1	2	3	4	5

		Very dissatisfied	Dissat-isfied	Neither satisfied nor dissatisfied	Satisfied	Very satisfied
2 (G4)	How satisfied are you with your health?	1	2	3	4	5

The following questions ask about how much you have experienced certain things in the last two weeks.

		Not at all	A little	A moderate amount	Very much	An extreme amount
3 (F1.4)	To what extent do you feel that physical pain prevents you from doing what you need to do?	1	2	3	4	5
4 (F11.3)	How much do you need any medical treatment to function in your daily life?	1	2	3	4	5
5 (F4.1)	How much do you enjoy life?	1	2	3	4	5
6 (F24.2)	To what extent do you feel your life to be meaningful?	1	2	3	4	5

		Not at all	A little	A moderate amount	Very much	Extremely
7 (F5.3)	How well are you able to concentrate?	1	2	3	4	5
8 (F16.1)	How safe do you feel in your daily life?	1	2	3	4	5
9 (F22.1)	How healthy is your physical environment?	1	2	3	4	5

The following questions ask about how completely you experience or were able to do certain things in the last two weeks.

		Not at all	A little	Moderately	Mostly	Completely
10 (F2.1)	Do you have enough energy for everyday life?	1	2	3	4	5
11 (F7.1)	Are you able to accept your bodily appearance?	1	2	3	4	5
12 (F18.1)	Have you enough money to meet your needs?	1	2	3	4	5
13 (F20.1)	How available to you is the information that you need in your day-to-day life?	1	2	3	4	5
14 (F21.1)	To what extent do you have the opportunity for leisure activities?	1	2	3	4	5

		Very poor	Poor	Neither poor nor good	Good	Very good
15 (F9.1)	How well are you able to get around?	1	2	3	4	5

The following questions ask you to say how good or satisfied you have felt about various aspects of your life over the last two weeks.

		Very dissatisfied	Dissatisfied	Neither satisfied nor dissatisfied	Satisfied	Very Satisfied
16 (F3.3)	How satisfied are you with your sleep?	1	2	3	4	5
17 (F10.3)	How satisfied are you with your ability to perform your daily living activities?	1	2	3	4	5
18 (F12.4)	How satisfied are you with your capacity for work?	1	2	3	4	5

	Very dissatisfied	Dissatis-fied	Neither satisfied nor dissatisfied	Satisfied	Very Satisfied
19 (F6.3) How satisfied are you with yourself?	1	2	3	4	5
20 (F13.3) How satisfied are you with your personal relationships?	1	2	3	4	5
21 (F15.3) How satisfied are you with your sex life?	1	2	3	4	5
22 (F14.4) How satisfied are you with the support you get from your friends?	1	2	3	4	5
23 (F17.3) How satisfied are you with the conditions of your living place?	1	2	3	4	5
24 (F19.3) How satisfied are you with your access to health services?	1	2	3	4	5
25 (F23.3) How satisfied are you with your transport?	1	2	3	4	5

The following question refers to how often you have felt or experienced certain things in the last two weeks.

	Never	Seldom	Quite often	Very often	Always
26 (F8.1) How often do you have negative feelings such as blue mood, despair, anxiety, depression?	1	2	3	4	5

Did someone help you to fill out this form?

How long did it take to fill this form out?

Do you have any comments about the assessment?

THANK YOU FOR YOUR HELP

Comments about case examples

The two case examples in this workbook are divergent, yet each demonstrates how evaluation planning can (and should) be tailored to address the unique needs of a given situation.

The first case describes the evaluation planning process for a new mobile crisis intervention and withdrawal management service located in rural Northern Canada. Evaluation was needed to determine the exact nature of the services that were being provided, and to evaluate how well the service was reaching its objectives. In this case report, the author describes the steps that evaluation members took to plan their evaluation: assembling their team, creating a programme logic model, outlining evaluation questions, and choosing evaluation measures and other data collection strategies. The overall evaluation was large in scope, encompassing all components of the service and including follow up data. However, the evaluators eased their resource burden somewhat by using information from standard clinical forms as much as possible. Other services with fewer resources might choose to design an evaluation that is narrower in scope. Regardless, the same basic principles of evaluation planning would still apply.

The second case describes the development of a new data collection instrument: the Maudsley Addiction Profile (MAP). In this situation, the evaluators decided that they needed an instrument that assessed a wide range of substance use, yet was brief to administer and simple to score. Once developed, the MAP could be used for future treatment evaluations. Similar steps could be used by other evaluators interested in developing instruments specific to their needs.

Case example of evaluation planning

The Timmins Home Detox Service An Implementation Evaluation

The author alone is responsible for the views expressed in this case example.

by
Cindy Smythe
Evaluation Consultant

Addiction Research Foundation
Social Evaluation and Research Department
The Godron J. Mogenson Bldg.
100 Collip Circle, Suite 200
U.W.O. Research Park
London, Ontario N6G 4X8 CANADA

Who was asking the question(s) and why did they want the information?

The Timmins Home Detox Service (THDS) is a mobile crisis intervention and withdrawal management service located in Timmins Ontario, Canada, a city in Northern Ontario with a population of approximately 30,000 people. The service was developed in response to the Addictions Services Multi-Year Plan for the Cochrane District Health Council. The District Health Council had conducted a need assessment that identified accessibility to district detoxification services by the Timmins population as problematic. At the time of the report, Timmins residents needing detoxification services could only be treated at the Cochrane District Detoxification Centre in Smooth Rock Falls, a 110 kilometre trip. Not only was it less likely that a client from Timmins would seek treatment, but it was also expensive to transport the client to and from the Detox Centre.

The THDS provides assistance to people in Timmins who are experiencing problems related to substance use. Through mobile teams of trained professionals and volunteers, the service offers crisis intervention and supportive guidance, withdrawal management, and advice related to addiction issues to persons experiencing problems, their families, and other supporters, including employers. Depending on client need, assistance is provided in the client's home, in hospital, at the Cochrane District Detox Centre, and in other safe places such as the home of a volunteer or friend.

The THDS was implemented by the Cochrane District Detox Centre, but was planned and developed with a community coalition coordinated by the Timmins Chamber of Commerce. Subcommittees prepare and implement program design, recruit volunteers, fundraise, and promote the service. For a more detailed description of the model this service used to guide its development, please see *A Guide for Planning Withdrawal Management Services in Rural and Remote Areas and Small Urban Centres of Ontario* (Addiction Research Foundation, 1994).

The Addiction Research Foundation has been a resource partner in the program design, and the integration of research and planning for community development. The director of the Cochrane District Detox Centre who is also the director of the THDS has worked closely throughout the project and in all phases of the project with the local Addiction Research Foundation (ARF) program consultant and the chair of the service's planning coalition.

The evaluation had two major purposes. First, the THDS was a new way of handling withdrawal management in Ontario and was serving as a model for other communities who might wish to establish such services. (At the time of the evaluation, most detoxification in the province was provided in one of 29 residential, non-medical detoxification services funded by the Ministry of Health.) Therefore, it was necessary to monitor and document how the THDS did its work in order to develop information and training products for other communities. Second was the interest of the various stakeholders of the program in knowing how well the service was doing in achieving its objectives. Stakeholders included the manager of the THDS, the head of the volunteer association, the clinical manager of the service, the volunteers, the Addiction Research Foundation, the community coalition, and the government funder (the Ministry of Health). Since the service was new and still evolving, we planned to put the emphasis on process or implementation evaluation rather than on client outcomes.

In the province of Ontario, at the same time as this new way of providing detoxification services was being implemented, the entire addictions treatment system was being rationalized. In other words, the Ministry of Health was looking for ways to provide good or better service for clients more efficiently. One of its guiding recommendations was that non-residential detox be considered in any regional treatment plan. Although there has been research done in other countries such as Australia, England, Scotland and the United States (see, for example Stockwell et. al, 1991; Stinnet, 1982; Hayashida, 1989) demonstrating the safety and effectiveness of outpatient and home detox, home detox in the Ontario setting had not been evaluated. For that reason, the ministry was very interested in the data and results of this evaluation as it wanted some data to back up its recommendation.

Component 1 - Service Awareness

One of the reasons for developing the Timmins Withdrawal Management Service was in response to the needs assessment done in the community. The community had identified the need for better access for potential target groups than was being provided in the traditional residential social-setting detox model. Four target populations in particular were identified as being more likely to access an at-home service: aboriginals, females, the elderly and youth. It was also felt that certain groups of referrers might increase their use of the service if there were a home component available to their clients. These were from medical, mental health, law enforcement and aboriginal services. Therefore, the work in this component was to ensure that any person or agency who may have need of the Timmins Home Detox Service was

aware of its existence and how to access the service. This was attempted by providing educational materials and presentations to potential referrers. For the evaluation, we wanted to know which kinds of awareness materials worked best with which referrers. For example, did the service need to develop different materials to educate physicians compared to educating other service providers such as mental health workers?

Questions: What awareness materials were developed? How were materials disseminated? How satisfactory were materials for the intended audience of referrers? How many clients were referred to each detox option (residential or home) from the targeted groups, i.e., aboriginals, females, elderly, youth? How many referrals to each detox option came from the targeted sectors, i.e., medical, mental health, law enforcement, aboriginal services?

Component 2 - Assessment

The main work in this component was assessing and identifying clients needing either immediate crisis intervention or referral to withdrawal management services and arranging transportation to the least intrusive location consistent with client needs. Every client who is deemed appropriate and lives in the city of Timmins was offered the non-traditional option, that is, some form of home detox. For the evaluation we wanted to know how many people were being assessed, how many were being referred to specialized care, that is how many clients have particular medical or other needs that first must be dealt with, how many were being referred to Home Detox, and how many to the Detox Centre.

Questions: How many clients were being assessed? What were the demographic characteristics of clients being assessed? How many clients were referred to (1) crisis intervention, (2) withdrawal management, (3) elsewhere? What were the criteria that determined location of client referral?

Component 3 - Crisis Intervention

The work in this component was to intervene with clients needing immediate service because they were in crisis. (Normally clients who call when not in crisis who wish to undergo withdrawal management, decide on a date for the process to begin.) The developers of this component hoped that through the crisis intervention service, clients who do not normally access the addictions treatment system would be introduced to it and would follow through with withdrawal management post crisis. However, they acknowledged that some clients will only stay with the service a short time — while they are in crisis, usually about 24 hours. For those clients, the THDS was able to provide crisis intervention in the clients' home community thus saving the cost of transporting clients 110 kilometres for a short stay.

For the evaluation, we were interested in who was accessing this service, that is, what were the demographics of those clients, and whether clients who received crisis intervention would otherwise have accessed the system. We were also interested in whether clients were following through on referrals to managing withdrawal or other forms of continuing care.

Questions: If crisis intervention were not offered, where would clients go for help? Do clients who are referred to specialized assistance come back to the program for withdrawal management? What are the demographic characteristics of clients who receive crisis intervention? What are the demographic characteristics of clients who stay in the program for withdrawal management following crisis? Are clients satisfied with the service?

Component 4 - Managing Withdrawal

The work in this component was the actual withdrawal and stabilization of the client. Trained staff monitored the client in his or her home throughout the process and arranged for medical services as needed. It is important to note that clients could come directly into the managing withdrawal component following assessment or could come into this component following crisis intervention. At the first visit to a client's home, a Home Environment Assessment Form (Cooper, 1994) was completed to be sure that home detox was appropriate for the client and to determine the amount of support available to the client. The amount of support available by family or friends helps to determine the number and frequency of contacts made by the service. Clients also complete a Symptom Severity Checklist (SSC) (Cooper, 1994) on the first visit and a shorter form of the SSC each day. The Severity of Alcohol Dependence Questionnaire (SADQ) (Stockwell, Murphy, and Hodgson, 1983) was also completed by the client on the second or third day depending on the client's physical and mental state. Results from these scales were monitored by service staff to determine whether the client needed any medical intervention. The service also provided clients and supporters with a telephone number for immediate contact with the service.

Questions:

1. The first set of questions we wanted to ask were about the demographic differences between clients who chose the non-traditional or home option for detoxification and those who preferred the regular Detox Centre: What were the demographic characteristics of clients in the different withdrawal program options? Why did clients choose each option?
2. The next set of questions were about the withdrawal management process itself and its effects on the client and family members: What is the frequency, duration and nature of in-person and telephone client contacts by agency staff/volunteers? Is the amount and nature of contact adequate? Can a client who lives alone complete withdrawal satisfactorily? What is the effect on family members of home withdrawal management? What roles do family members play? In which setting are clients more likely to complete withdrawal? Do clients who withdraw at home miss the companionship of others found at a detox centre?

Component 5 - Facilitating Continuing Care

The work in this component was to ensure that clients continue care following withdrawal and stabilization. All clients in both the Home Detox program and at the Detox Centre are referred to one of the two Addictions Assessment and Referral Services in the district (for assessment and referral to continuing care) and to mutual aid (e.g., AA, NA) following withdrawal. The detox program had made agreements with the Addictions Assessment and Referral Services to the effect that the detox would be given notification of whether each client referred followed through on the appointment to the service. Part of the client consent form asked for permission from the client to document this information. Because of this relationship between the detox service and the Assessment and Referral services, we will be able to look at this important outcome measure even though the evaluation is essentially concerned with implementation.

Questions: What are the characteristics of clients who do and who do not continue care? Does the type of detox experienced, that is home or residential, make a difference in intention to continue treatment?

Component 6 - Volunteers

Because volunteer recruitment, training, motivation and supervision were extremely important to this service, we decided to include the component as part of this evaluation. Trained volunteers work in pairs, under the supervision of professional staff, to monitor clients who are withdrawing in their homes. We were particularly interested in finding out how to meet the demand for service by using an appropriate mix of professionals and volunteers. It is important that the volunteer pool not be bigger than necessary because administering the volunteer program takes staff time that could be used for service delivery.

Questions: How many volunteers per professional staff are necessary to meet demand? How many volunteers are necessary in the pool vis à vis potential demand? How long should a volunteers term be? What skills should volunteers have?

What resources were needed to collect and interpret the information?

Because this evaluation was seen as rather large and complex, an evaluation consultant from the Addiction Research Foundation participated in all phases of the planning and implementation of the evaluation and was responsible for the coordination of all data collection. However, were the service to undertake the evaluation itself, it would not be necessary to undertake such a large evaluation all at once. The evaluation could be implemented in stages, looking at specific questions over a period of time. In that case, an evaluation consultant may not be necessary or one could be used sparingly over time.

For our evaluation, we relied on the fact that the Detox Centre already collected client information that was stored in a database on the service's computer. The Centre was able to provide statistics about service utilization and client profiles. Therefore we planned that the service itself would take the responsibility for collecting all the data necessary for the evaluation. We planned that through its normal record keeping computer database, the service would generate most of the statistics for the evaluation report. For most of the statistics we would gather, however, it would not be absolutely necessary to use a computer though a computer would make the job easier. Many of the service utilization statistics could be summarized by hand with the help of a calculator.

How were the data collected?

In order to begin the evaluation, we developed a Program Logic Model (see Chart on next page). The first version of the logic model for this program was developed by the director of the THDS in collaboration with the program consultant from the Timmins local office of the Addiction Research Foundation. This model was further developed and refined with input from the program evaluator. All versions of the model were shared with members of the community coalition. Once the final version of the model was approved by all the stakeholders, a similar procedure was used to generate and refine the evaluation questions. The questions were generated following a full day meeting of the detox director, the ARF program consultant and the evaluator. The questions were then shared with the coalition and were also sent to the chair of the Ontario Detox Directors Association for more feedback. Attempts were made to incorporate all suggestions if appropriate.

Working with a program logic model was very useful in gaining consensus about the workings of the service and how the services provided were expected to have an impact on clients. It should be noted that most of the stakeholders had no real knowledge about program evaluation. They were unclear about the relationship between objectives as stated in the logic model and the eventual evaluation questions we had generated. For that reason, time had to be scheduled at the feedback meetings mentioned above for education of the stakeholders. The evaluator prepared teaching materials and led the group in a simple logic model exercise so that everyone was familiar with the concepts.

The planned length of the evaluation was twelve months. During that time we hoped to monitor all clients who entered both the Home Detox and the Detox Centre programs to document client demographics, referrer demographics, services received, and client satisfaction. We also planned to evaluate the volunteer component of the service.

Component 1 - Service Awareness

When we considered how we would collect the information necessary to answer our evaluation questions, we made every effort to use as many of the normal record keeping forms and materials that the service was already using as we could. This would make the data collection for the evaluation less of a burden on the agency whose main function is to provide service to its clients.

In order to answer our questions about the awareness materials, we decided records should be kept of how the materials were disseminated, that is, how were brochures distributed, were information sessions held, etc. A simple form was developed to record this information. We then planned to survey referrers to assess their satisfaction with the materials and any presentations they may have attended. Because the THDS has a volunteer component, we planned to have volunteers telephone referrers following their receipt of the materials. The volunteers would use a short questionnaire we developed.

In order to answer our questions about service utilization and referrals from targeted referrers, we decided that normal record keeping from client face sheets would be used to record this information. In Ontario, every government funded addictions treatment agency is part of The Drug and Alcohol Treatment Information System (DATIS). The face sheet form completed for each client as part of DATIS contains the name and profession/agency of the referrer, the age, sex and ethnic background of each client as well as other demographic and substance use information. We planned to compare client demographic data from the DATIS face sheets to existing information from the needs assessment to discover whether client profiles matched those identified as needing better access to detox services. We also would be able to compare client and referrer demographics related to use of and referral to home and residential detox.

All agencies and persons making referrals, including self-referrals would also be asked as part of the normal Intake Assessment Form, completed for each client, how they heard about the service and what they would have done if the service were not available.

Finally, we planned to survey referrers to assess their satisfaction with the service and their reasons for referring particular clients to the service. Again, we planned to use volunteers to either phone referrers or to mail out a short questionnaire.

Quantitative and Qualitative Approaches: As always when collecting data, it is important to consider the form the information should be in, that is, do

we want numbers (quantitative data) or do we want words and description (qualitative data)? We knew the information on service utilization, in this case, would be quantitative, that is, we were interested in knowing the number of people in different demographic categories who use the service. We also wanted quantitative information about the categories the referrers represent. However, when we came to evaluating the materials and satisfaction with the service, we decided to use a mixture of quantitative and qualitative approaches. Qualitative approaches would be particularly useful during the process of developing the promotional materials. For example, we planned to distribute a prototype of a brochure to a group of people representing potential referrers and hold a focus group in which people would react to the brochure and suggest improvements. This information would then be used to redesign the brochure, if necessary.

When assessing referrer satisfaction with the service, a mix of both qualitative and quantitative methods would be used. We planned to develop some five-point scales or a checklist on which referrers could mark their satisfaction with aspects of the service like ease of access, promptness of service, etc. However, we also wanted to ask some open-ended questions about what referrers liked most and least about the service.

In general, it is more time consuming to collate and interpret qualitative data than quantitative. Having access to a computer or calculator and some basic computing skills makes organizing and analysing quantitative data fairly efficient. On the other hand, there is no fast way to organize qualitative data, but it can provide some very rich information and is particularly useful in the development phase of a product or program.

Component 2 - Assessment

Again, we were able to use normal record keeping forms to document most of this information. We planned to use one form for each client assessed containing demographic information as described in the first component (DATIS) and location of the referral (Intake Assessment Form). Information about the clients' condition from the Intake Assessment Form (that is, substances used and last use, client's level of intoxication, any suicidal ideation, symptoms of withdrawal, and withdrawal history) would be compared to client referral information to determine how decisions were made when referring a client.

Component 3 - Crisis Intervention

Again, we were able use normal record keeping forms to document most of this information. There was one form for each client assessed containing demographic information as described in the first component (DATIS) as well as referral destination and whether clients referred to other services recontacted the Home Detox following specialized treatment.

In order to determine client satisfaction with the service, we decided to use the eight-item version of the Client Satisfaction Questionnaire (CSQ-8) (Larsen et al., 1979) to be completed by clients before they left the program. Information about the CSQ-8 as well as information about other client outcome measures can be found in the *Directory of client outcome measures* (Graham et al., 1993). Added to that questionnaire would be a question in which clients were asked what services they would have used had they not been aware of the THDS. Options they could check off were: hospital, physician, jail/police, on the street, stayed home, and other.

CHART: Programme Logic Model for

Component	Service Awareness	Assessment	Crisis Intervention
Implementation Objective	• To develop and update profiles of potential and actual client population • To develop awareness materials and educate sources on the service and how to access it. • To negotiate agreements and protocols with referal sources and partner agencies, especially mental health and district assessment and referral agencies	• To provide professional assessment of clients in person or by telephone • To refer clients as appropriate to - immediate specialised assistance (e.g. medical) -crisis intervention or managing withdrawal • To arrange transportation (if necessary) to the least intensive location appropriate to client's needs.	• To provide crisis intervention as appropriate to client's immediate needs • To refer clients to continuing care consistent with care plan goals.
Short-term Outcome Objectives	• To increase overall utilisation of the service • To increase referrals from the following populations: aboriginals, females, elderly, youth • To increase referrals from the following referral sources: medical, mental health, law enforcement, and aboriginal services	• To increase clients understanding of their needs through development of care plans • To increase appropriate referrals to care based on clients' needs	• To increase the number of clients receiving appropriate crisis support • To increase the number of clients who follow through on care plan goals post crisis • To increase access for those clients who do not normally enter the addiction treatment system
Long-term Outcome Objectives	• To increase per person cost of withdrawal management services • To increase access to withdrawal management services for traditional and non-traditional clients		

Timmis Withdrawal Management Service

Managing Withdrawal	Faciliating Continuing Care	Volunteers
• To provide or arrange for a safe withdrawal on a positive non-threatening place • To provide trained staff to monitor client during withdrawal for signs that alternate care is required and/ or withdrawal is completed. • To arrange for an access to emergency medical services should complication arise. • To arrange for ongoing medical management as indicated in care plan.	• To help make contact with A/R and other services and advise professional care providers regarding relevant care issues • To obtain confirmation from formal referral destinations regarding client entry and completion	• To develop and utilise materials for recruitment of volunteers • To design and implement a plan for volunteer development and retention • To design and implement a training programme for service volunteers
• To increase the number of clients who complete withdrawal management	• To increase the number of clients who enter formal and informal continuing care	• To ensure numbers and proportionals of volunteers to professionals are necessary to meet service demands • To ensure an adequate skill base among volunteers appropriate to client needs.

- To improve continuity of care for clients requiring withdrawal management services
- To increase the number of low-cost, innovative withdrawal management services in Ontario

Component 4 - Managing Withdrawal

Again, we were able to use normal record keeping forms to document this information as described in the above components. In addition, clients were asked for information about why they chose the withdrawal management sites and options they chose and if they thought their choices were appropriate. If clients chose the non-traditional option, we were interested in whether they would have called if the Detox Centre option were the only choice. We added these additional questions to the client satisfaction questionnaire described below.

We were also able to use normal record keeping forms to document all of the client contact information, including number, type and length of contacts. This was accomplished by using the Service Log which is part of the DATIS system mentioned earlier, and used by all Ontario addiction treatment programs. This Service Log function allowed us to enter every contact the client had either in person or by telephone with the program staff. We were also able to code what happened in the contact, for example, 'the client's wife wanted to discuss her concerns'. If such a sophisticated system were not available, the same information could be hand coded and looked at for a sample of clients. Because of this system, we were able to track frequency and duration of contacts for clients over the year of the study.

In order to assess client and family satisfaction, we used adaptations of the forms found in David Cooper's book, *Alcohol Home Detoxification and Assessment* (1994). In order to use these forms we received permission from the book's publisher. To the satisfaction forms we added questions about the roles family members and/or friends played. We also asked clients if they missed the companionship of others in a similar situation they would have found in a Detox Centre.

Component 5 - Facilitating Continuing Care

There are few addictions treatment programs in northern Ontario, consequently, the managers of the services were very collegial and already helped each other track their clients. However, we wanted to ensure that clients knew they were part of a study; and so they were told about the evaluation and were asked to complete the client consent form giving us permission for their records to be included. We assured them that they would never be identified by name but only by an identification number. We were then able to add the information about whether or not clients kept their appointments at the Addictions Assessment and Referral Service. It was beyond the scope of this evaluation to follow up clients beyond the Assessment and Referral Service to find out if they had then gone to subsequent treatment.

Component 6 - Volunteers

As was mentioned above, we were able to document the time spent by professionals and volunteers for each part of the service through our DATIS Service Logs. This quantitative information could then be paired with qualitative information from professional staff and volunteers about their subjective experience and with client and supporters' satisfaction with the service they received. We decided to interview staff and volunteers to find out their perceptions of their work load. From this we hoped to be able to comment on the appropriate mix of staff.

The question about the number of volunteers that should be available at any one time was important because administering a large volunteer pool is time consuming and takes away from other service. We

decided we would monitor use of the volunteer pool to document demand over the year. We were also interested in interviewing volunteers to find out how they felt about actively working in the program as opposed to being available but called infrequently.

How were the data analysed?

The detox service director and the evaluator decided that all data needed to evaluate the program would be collected by the service and input into the service's database in the usual manner by the service support staff. Once the data were in the service's computer, the service would generate basic descriptive statistics as usual. All analyses for the evaluation would be descriptive in nature, that is, we would be documenting the operation of the service as described above. We planned to make some comparisons between the clients in the detox centre and those in the non-traditional environment using cross tabulations. Clients would be monitored for one year following the beginning of data collection. Each component of the model would be described separately.

All data would be linked by the client's unique DATIS identifier. Therefore, client and family or supporter satisfaction questionnaires would be precoded by the detox service. In order to ensure client confidentiality, detox workers would distribute the questionnaires with envelopes and ask that the completed forms be sealed in the envelopes before they are returned. Only the support staff who enters the data would see the questionnaires and that person would not know the client or supporter's name.

At the end of the evaluation, we would be able to describe the clients who enter both detox services and crisis management; the referrers and to which service they referred; length of stay in the services and number and duration of contacts in the home detox; satisfaction with services; and number and description of clients who keep their appointments at the Assessment and Referral services. We would also be able to comment on the use of volunteers in the service.

Evaluation Plan

At this point we needed to develop a written evaluation plan. The written plan was important because in it we could set down the various issues that had been discussed in the course of this case study, for example, what led up to the evaluation or why was it being done; who the stakeholders were; the program logic model; the evaluation questions; what the organizing principles were (that is, what key principles and processes would guide the evaluation? In our case we are trying to make the process participatory and focus the evaluation on process; how the results would be reported and used; and finally, costs.

We decided that the evaluator, in consultation with the service's director, the ARF consultant and the chair of the community coalition would write the plan. We would discuss all components of the plan and the final draft of the plan would be circulated among the service's staff and the members of the coalition. This seemed to be a fair division of work since the service was taking responsibility for all of the data entry and generating the statistical reports. The evaluator would also be responsible for writing the final report of the evaluation. Interim reports would be the responsibility of the service's director and the ARF program consultant in consultation with the evaluator. The evaluation results would be available to all of the stakeholders mentioned previously. We hoped that the detox service staff would use the results to modify and improve the program as necessary. We assumed that the Ministry of Health

would use any positive results to promote the concept of home detoxification in the province.

Costs

As was stated above, we planned to evaluate the program for one year following the beginning of data collection. Because the detox director was so enthusiastic about the evaluation, she volunteered her service's support staff to enter all of the data collected into the computer as part of her regular job. We also planned to use the service's volunteers to help collect some data. Finally, we tried to use the service's usual data collection forms as much as possible in the evaluation to lessen the burden for the service's workers who must complete them. For these reasons, we hoped to keep the cost of the evaluation low. The services of the evaluator and the Addiction Research Foundation program consultant were provided at no cost to the program. The ARF is funded by the same branch of the government that funds addiction treatment agencies and part of ARF's work is to support those agencies. Of course, these costs, while not involving exchanges of money, could be calculated out, for example, the evaluator had dedicated 20% of her time to the project for the year, the ARF program consultant, 30% of his time, the service support staff, 15% of her time, and so forth. The largest real money outlays were travel expenses for the evaluator to fly 800 kilometres to Timmins to meet with the service's staff and coalition. Some of the travel costs were covered by the fundraising efforts of the service's volunteers and coalition.

As was also stated earlier, it would not be necessary for a service to undertake a process evaluation of all components of the service at once as was being planned here. A service working alone may decide to only evaluate one component at a time and collect data for that component for a shorter period of time than we planned, for example for three months. Over time all the components in which the service was interested could be evaluated and the costs and the resources used would be less intense at any one time.

References for the case example

Addiction Research Foundation *A Guide for Planning Withdrawal Management Services in Rural and Remote Areas and Small Urban Centres of Ontario,* 1994

Cooper, David B. *Alcohol Home Detoxification and Assessment*. Oxford: Radcliffe Medical Press, 1994.

Graham, K., Price, B., Brett, P., Baker, A., Bois, C., Boyle, B., Chapman, L., Eliany, M., Gaskin, J., Martin, G., Sobell, L.C., and Thompson, J. *A directory of client outcome measures for addiction treatment programs.* Toronto: The Addiction Research Foundation, 1993.

Hayashida, M., Alterman, A., McLellan, A., O'Brien, C., Purtill, J., Volpicelli, J., Raphaelson, A., and Hall, C. Comparative Effectiveness and cost of inpatient and outpatient detoxification of patients with mild-to-moderate alcohol withdrawal syndrome. *The New England Journal of Medicine,* 1989, 320: 358-365.

Larsen, D., Attkisson, C., Hargreaves, W., and Nguyen, T. Assessment of client/patient satisfaction: development of a general scale. *Evaluation and Program Planning,* 1979, 2: 197-207.

Stinnett, J.L. Outpatient detoxification of the alcoholic. *The International Journal of the Addictions,* 1982, 17: 1031-1046.

Stockwell, T., Bolt, L., Milner, I., Russell, G., Bolderston, H., and Pugh, P. Home Detoxification from alcohol: its safety and efficacy in comparison with inpatient care. *Alcohol and Alcoholism,* 1991, 26: 645-650.

Stockwell, T., Murphy, D., and Hodgson, R. The Severity of Alcohol Dependence Questionnaire: its use, reliability and validity. *British Journal of Addiction,* 1983, 78(2): 145-155.

It's your turn

What are the strengths and the weaknesses of the presented case example? List three positive aspect and three negative aspects:

Strengths of the case study

1

2

3

Weaknesses of the case study

1

2

3

Case example of implementing evaluations

Development of the Maudsley Addiction Profile (MAP)

Marsden J, Gossop M, Steward D, Best D, Farrell M, Edwards C, Lehmann P & Strang J. (1998) The Maudsley Addiction Profile: A brief instrument for assessing treatment outcome. *Addiction* 93, 1857-1868.
National Addiction Centre
Maudsley Hospital/Institute of Psychiatry
4, Windsor Walk, London, SE5 8AF, United Kingdom

Dr. John Marsden
National Addiction Centre,
Maudsley Hospital/Institute of Psychiatry
4, Windsor Walk, London, SE5 8AF, UK.
Tel: + + 171 919 3830
E-mail: J.Marsden@iop.kcl.ac.uk

Who was asking the question(s) and why did they want the information?

The authors alone are responsible for the views expressed in this case example.

This case study summarises the development of the Maudsley Addiction Profile (MAP). The MAP is a brief, interviewer-administered questionnaire for use with problem drug users (DUs) and alcohol users (AUs). It has been developed by a team of researchers in association with clinical staff at the Maudsley Hospital/Institute of Psychiatry, London. The MAP is intended for use in treatment outcome research and for routine use by treatment programmes for evaluation. The items have been selected to be sensitive to change and the instrument can be administered at one or more points during and after treatment.

In our development studies we wanted the MAP to have: (i) good content and face validity; (ii) brief administration; (iii) simple scoring with clear and unambiguous interpretation; and (iv) acceptable standards of test-retest reliability. The first stage involved a review of the relevant literature and discussions with key informants in the research and clinical fields in the UK and internationally about the relevant assessment domains for this client population

(Marsden, 1994). Four domains emerged as essential components of an outcome questionnaire: (a) drug and alcohol consumption; (b) health risk behaviour; (c) health problems; and (c) personal/social functioning (the latter usually spanning relationship problems, employment and crime involvement) (McLellan *et al.* 1980; Babor *et al.* 1994; Simpson & Chatham 1995). For the first domain, it is important to note that we elected to assess the typical quantity of substance use. Whilst we acknowledge the accuracy of self-reports of drug doses is problematic, we considered that this is a desirable clinical and research measure.

The development of measures in each of these four domains was guided by existing concise omnibus instruments for treatment outcome evaluation — the Addiction Severity Index (ASI, McLellan *et al.* 1980; McLellan *et al.* 1992a) and the Opiate Treatment Index (OTI, Darke *et al*. 1992). We were also informed by our work on developing measures and working with treatment providers during the planning phases for the UK's National Treatment Outcome Research Study (see Gossop, Marsden & Stewart, 1998, for description). On the basis of this review phase, we established for following structure for the MAP.

Structure of the Maudsley Addiction Profile (MAP)

Domain	Measure	Variable (behaviour in past 30 days)
Substance use	Consumption	• Days used • Usual amount taken on using day • Usual route (oral, intranasal, inhalation, injection)
Risk behaviour	Injecting	• Days injected and times injected per day • Times shared needle/syringes
	Sexual	• Number of sexual partners (non-condom) • Times had sex when not using a condom
Health	Physical	• Frequency of major symptoms (general, gastrointestinal, neurological, musculo-skeletal, cardio-respiratory)
	Psychological	• Frequency of major symptoms (anxiety and depression)
Social functioning	Employment	• Days in paid work/unemployment • Work days lost due to absence
	Relationships	• Days contact with partner/relatives/friends • Days conflict with partner/relatives/friends
	Crime	• Days committed drug selling • Days committed shoplifting and other offences

The MAP comprises an introductory section and, in the version used for field testing, contains a total of 60 items. A recall period of the past 30 days before intake to treatment is used. Given the purpose of the instrument, no lifetime or history measures are included. The purpose and structure of the MAP and confidentiality issues are explained to the client at the start of the interview and their age, gender, and ethnic group recorded.

Substance use

The first section has 22 items. In the field test version, the following substances were recorded: illicit heroin, prescribed and non-prescribed methadone, prescribed and non-prescribed benzodiazepines (commonly diazepam and temazepam), cocaine base (crack), cocaine hydrochloride and alcohol. To assist recall, the client is shown a response card which lists seven common frequency patterns (ranging from one day per week to every day). In this way, the total number of days in the month before intake when use of each substance took place is recorded first. The intensity (or quantity) of use is then recorded from the client's verbatim report of the amount consumed on a typical using day in the past month. In instances where the quantity has varied, the client is asked to recall the amount used on the two-three most recent days in the recall period when use took place. The amount used is recorded for each of these days and later averaged. With the exception of alcohol, the usual route(s) of substance administration during the recall period is also recorded, using the following categories: oral, intra-nasal, inhalation and injection.

Injecting and sexual risk behaviour

The second domain has five measures. The number of days on which the client injected is recorded together with the number of injections on a typical day. If injecting has been episodic, it is recorded using the same procedure for the frequency of episodic substance use. The client is also asked about the number of times they have injected using a needle/syringe which they believe has already been used by someone else (our definition of needle/syringe sharing). As a proxy of recent sexual risk behaviour, the client is asked to estimate the number of people and the total number of times that they have had penetrative sexual intercourse without using a condom.

Physical and psychological health

The third domain has 20 items. For physical health, a 10-item symptom scale was adapted from the 51-item checklist developed by Darke and colleagues for the Opiate Treatment Index (Darke *et al*. 1991). A five-point Likert-type scale was included to assess the frequency of experiencing each symptom (never, rarely, sometimes, often, always (scored 0-4) in order to maximise change sensitivity. The scale is simply scored by summing the values across each item (range 0-40). A 10-item scale to assess general emotional problems was derived and adapted from the six-item anxiety and depression sub-scales of the Brief Symptom Inventory (BSI, Derogatis 1975). The scale is simply scored by summing the values across each item (range 0-40). Separate scores for anxiety and depressive ideation may also be computed.

Social functioning

The fourth domain has nine items in three sub-sections:

(i) Relationship conflict. A measure of serious conflict experienced by the client is taken by recording the number of days on which the client had contact with his/her usual sexual partner, relative(s) and friends, and the number of days on which there was serious conflict between the client and each one or group. These measures

were based on the measurement of family conflict developed for the Addiction Severity Index (McLellan *et al.* 1992a). If the client was not in a personal relationship, or did not have any relatives or friends, a score of zero is recorded. These measures are subsequently expressed as a percentage of conflict to contact days.

(ii) Employment. Three measures of employment are taken: the number of days of formal unemployment; the number of days on which the client undertook paid work; and the number of working days on which the client did not attend work due to sickness or unauthorised absence.

(iii) Criminal behaviour. In this final section, the three categories of criminal activity are recorded: sales of illegal drugs, shoplifting, and other crime. The latter category groups the following offences together: theft from a property, theft from a person, theft from or of a vehicle, and fraud/forgery. The client is asked to recall the frequency of days on which the crime type was committed during the past month. Finally, as a measure of intensity of crime involvement, the respondent is asked to estimate the number of times each crime type has been committed on a typical day.

What resources were needed to collect and interpret the information?

Internal research and clinical resources were used for the development of the MAP questionnaire and its initial psychometric evaluation. The research team, led by the first author, comprised two male and two female researchers. All were experienced members of the Addiction Research Unit and trained to first, masters or doctoral degree level in psychology or sociology. Two male and two female clinical nurses participated as interviewers under the supervision of the fifth author (a consultant psychiatrist). All were experienced members of our clinical services and were trained in general or psychiatric nursing with addictions specialties. Our internal research and clinical resources were supplemented with a research grant from the English Department of Health in support of the project. In support of the project, we prepared a detailed protocol and instruction set and used a single two hour training session to induct the interviewers into the conduct of the interviews. Periodic meetings were held during the course of field testing to discuss and resolve any difficulties encountered in interviewing clients.

How were the data collected?

We conducted field testing studies with samples of DUs and AUs at our community and in-patient services at the Maudsley Hospital, London. For a detailed description of the initial psychometric evaluation of the instrument see Marsden et al (submitted for publication). Two multi-disciplinary community programmes participated: an opioid substitution and detoxification service for drug users and an assessment, brief intervention, and detoxification service for AUs. Two in-patient programmes also participated: a 30-day detoxification and relapse prevention service for DUs and a 10-day assessment and detoxification service. The purpose of these studies was to evaluate several aspects of the psychometric performance of the MAP. A quota-recruitment procedure was used to obtain a total of 240 clients (160 men and 80 women) from the four treatment programmes for the test-retest and concurrent validation assessments. The eight inter-

viewers each interviewed 30 clients (20 DUs and 10 AUs). Participation in field testing was voluntary and no client was paid for their time.

How were the data analysed?

Psychometric evaluations of test-retest were assessed using the intra-class correlation coefficient (ICC) for interval measures and Cohen's Kappa for categorical measures.

What did they find out?

The items in the MAP were found to be highly acceptable to a majority of the clients. The average completion time for the test interview was 11.7 minutes (range = 6-27 minutes). Satisfactory results were obtained from the assessment of self-report with urine testing, concurrent validity of scales with comparable measures and 3-day intra-interviewer test-retest reliability.

How were the results used?

The results of the initial field testing of the MAP achieved satisfactory performance in terms of the development objectives we set for the instrument and the design and field testing phases of the study provided evidence to satisfy each of these. These results have several implications for outcome research in practice. In practice, the eight substances recorded in the field-tested version of the MAP will need to be expanded to include other substances as prevalent patterns of use change over time and location. Alternative recall periods could also be used: either the past 60 days (two months); the past 90 days (three months) or even the past 180 days (six months).

If the MAP is used widely as part of routine clinician practice, there are advantages in incorporating it within existing clinical assessment records and procedures. This is the approach we have taken in our own services. We recommend that a modular approach to outcome research is adopted in which a primary set of outcome measures are recorded, with other outcome measures included as required. Measures of treatment itself are also clearly needed and there are several instruments available which record different aspects of treatment processes and programme environment (Simpson & Chatham 1995; Moos 1988b; McLellan *et al*. 1992b). As a minimum, the setting, modality and duration of treatment should be recorded. Additional measures of the client's choice of treatment goals, the amount of programme services received, their treatment discharge status and additional non-index treatments received, could also be used in a more comprehensive outcome evaluation. We have also recently developed the Treatment Perceptions Questionnaire (TPQ), a brief 10-item measure of addiction treatment satisfaction, which can be used as an adjunct to quantitative outcome assessments (Marsden et al., in press).

We conclude that the MAP development study has demonstrated that collection of a set of quantitative measures of problems experienced by clients should not place a significant administrative burden on treatment or research personnel. In consequence, we hope that it may stimulate the development of research activity by treatment providers.

References for the case example

Babor, T.F., Longabaugh, R., Zweben, A., Fuller, R., Stout, R., Anton, R.F. & Randall, C.L. Issues in the definition and measurement of drinking outcomes in alcoholism treatment research. *Journal of Studies on Alcohol*, 1994, 12 (Supplement):101-111.

Darke, S., Hall, W., Wodak, A., Heather, N. & Ward, J. Development and validation of a multi-dimensional instrument for assessing outcome of treatment among opiate users: the Opiate Treatment Index. *British Journal of Addiction*, 1992, 87: 733-742.

Derogatis, L.R. *Brief Symptom Inventory* (Baltimore, Clinical Psychometric Research), 1975.

Gossop, M., Marsden, J. & Stewart, D. *NTORS at one-year: Changes in substance use, health and criminal behaviour one year after intake*. London: Department of Health, 1998.

Marsden, J. *Treating drug and alcohol problems: A review of assessment domains, outcome measures and instruments*. London: Department of Health, 1994.

Marsden, J., Gossop, M., Stewart, D., Best, D. Farrell, F. Lehmann, P., Edwards, C. & Strang, J. (1998) The Maudsley Addiction Profile (MAP): A brief instrument for assessing treatment outcome. *Addiction*, 93, 1857-1868.

Marsden, J., Stewart, D., Gossop, G., Rolfe, A., Bacchus, L., Griffiths, P., Clarke, K. & Strang, J. (in press) Assessing client satisfaction with treatment for substance use problems and development of the Treatment Perceptions Questionnaire (TPQ) *Addiction Research*

McLellan, A.T., Luborsky, L., O'Brien, C.P. & Woody, G.E. An improved evaluation instrument for substance abuse patients. *Journal of Nervous and Mental Diseases*, 1980, 68: 26-33.

McLellan, A.T., Alterman, A.I., Cacciola, J., Metzger, D., & O'Brien, C.P. A new measure of substance abuse treatment: initial studies of the Treatment Services Review. *Journal of Nervous and Mental Disease*, 1992b, 180:101-110.

McLellan, A.T., Kushner, H., Metzger, D., Peters, R., Smith, I., Grisson, G., Pettinati, H. & Argeriou, M. The Fifth Edition of the Addiction Severity Index. *Journal of Substance Abuse Treatment*, 1992a, 9:199-213.

Moos, R.H. *Community-Oriented Programs Environment Scale manual: Second edition* (Palo Alto, CA, Consulting Psychologists Press), 1988b.

Simpson, D.D. & Chatham, L.R. *TCU/DATAR Forms Manual*. Improving Drug Abuse Treatment, Assessment, and Research (DATAR) Project. Fort Worth, Texas: Institute of Behavioural Research, Texas Christian University, 1995.

It's your turn

What are the strengths and the weaknesses of the presented case example? List three positive aspect and three negative aspects:

Strengths of the case study

1

2

3

Weaknesses of the case study

1

2

3