feature

Michael Donmall, Petra Meier, Tim Millar

Treatment outcome evaluation

Is treatment working?

Evaluation of treatment outcome – finding out what the effects of the treatment process are – has become vital for service providers, commissioners and funders, but what is the best way to go about it?

Michael Donmall, Petra Meier and **Tim Millar** work at the Drug Misuse Research Unit of the University of Manchester. Research supported by Sefton Health Authority on behalf of North West Health Authorities

Many agencies find it difficult to evaluate the quality of treatment outcomes. When evaluation results are not what they expect they are disappointed. So here is a plain guide to some of the issues that people should bear in mind when trying to measure the outcome of treatment.

Increasing demand for health care challenges health service commissioners who have limited resources. They need to evaluate different drug treatment procedures and to assess differences between services, so they can decide which treatments and services to purchase. They also need ways to monitor the outcomes from services that they purchase.

There are no agreed quality criteria to inform such decisions, and as different treatment programmes vary widely in structure and processes of care, they are difficult to compare.[1] Purchasers and agencies themselves are increasingly interested in methods of evaluating treatment outcome, for several reasons:

- to assess the value of the service to the client
- to support funding applications
- to investigate how clients respond to changes in service
- to satisfy purchaser requirements.

Evaluation model

Donabedian[2] developed one of the best models of how to evaluate the quality of medical care. This proposes evaluation of care through analysis of *structure*, *process* and *outcome*.

Structure refers to variables which are independent of individual clients – in drug treatment this could be accessibility of treatment, methods of treatment procedure, type of agency, treatment goals, client and care plan matching, staffing.

Process is about the impact of treatment variables, including duration of programme, frequency and quality of therapeutic intervention, and programme management.

Outcome indicates the condition of the patient during or after treatment, including changes in their behaviour.

A quality programme can be defined as one with a structure and process that leads to desirable outcomes.

Examples of process outcomes:

- take up rates of services
- retention/drop-out and compliance
- taking up of agency's offers

Outcome domains

Domain	Possible outcomes
Drug use	Abstinence, controlled drug use, cessation of illicit use, change of risk behaviour
Physical health	Resolution of health problems, registration with GP and dentist, improvement in general health
Psychological health	Decrease in depression/anxiety, improved self-concept, improved quality of life
Legal situation	Compliance with criminal justice system requirements, resolution of outstanding issues, reduced rate of re-offending
Social situation (housing, employment)	Registration on housing list, planned discharge accommodation, skills training initiated, registration with job centre
Social relationships	Plan concerning child care issues, contact with relevant agencies, improved family relationships

- registration with GP and dentist
- prescription type and duration
- client-staff relationship
- referral and liaison actions
- client satisfaction and staff satisfaction.

Treatment outcomes are highly specific and should be assessed with respect to the specified aims, objectives and philosophy of the agency. For example, a counselling service may focus on different outcomes to a detoxification centre and should be evaluated in a different way. Treatment outcomes may include changes in patterns of drug use, quality of life, criminal activity and health.

Outcome evaluation

Several methodological issues should be addressed before undertaking an outcome study.

The desired outcome should be clearly defined. Many studies (and associated outcome instruments) fall into the trap of attempting to measure too much. As a result, analysis of data collected does not give clear results.

It is better to focus on one or a few key areas and conduct several smaller-scale studies, rather than a single study examining all possible outcomes. Perhaps it is most important to be clear what it is you want to measure before starting your evaluation. It is also important to find out what the commissioner/agency is most interested in.

There are three basic designs which may be combined in large studies:

- longitudinal comparison of the effectiveness of different methods of process for treatment
- pre- and post-treatment comparison, with or without a control group
- interval service evaluation, usually without a control group, sometimes compared to clients on waiting lists.

The choice of design is likely to depend on the resources and time available for the project.

If possible there should be a carefully chosen (untreated) control group, otherwise changes cannot be attributed to treatment. For example, because clients may seek treatment at time of crisis it is plausible that someone's situation may improve over time whether or not they receive treatment.

If a change in the outcome variable is found in the treated group but not in the control group, statistical analysis will tell you whether it is reasonable to conclude that the intervention has caused the change.

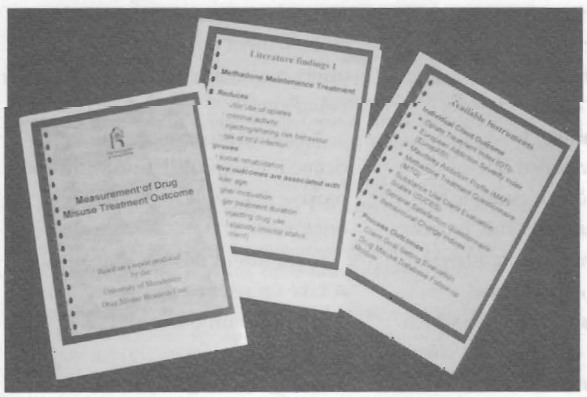
Sample sizes

The smaller the sample, the less likely it is that positive outcomes will be detected. But large sample sizes involve increased costs and time for data collection and entry, so you should attempt to get the right balance.

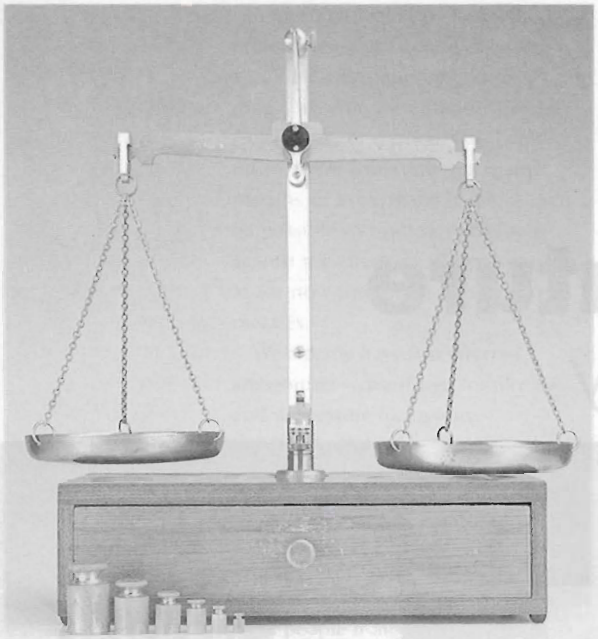
There are statistical procedures, called a-priori power analyses, that should be used to determine sample size.

An important issue is the selection of appropriate target groups. When comparing different treatment services, or even process methods, samples should be sufficiently similar, for example in age and gender distribution, severity of use, mental health and priority status.

The study should have a carefully planned time schedule. A very short follow-up period may not provide opportunities for change, too long a period may conceal differences. There is little literature to guide this decision and it will depend on planned treatment duration. Three-monthly reviews may be appropriate to assess impact and long-term consequences. Whenever possible a pilot study should precede the actual study.



Purchasers and agencies themselves are increasingly interested in methods of evaluating treatment outcome



If an instrument asks for the frequency of drug use and gives the scale 'seldom, often, or always', one person may think three times a week is relatively seldom whereas another might consider this as quite often

Frequency and duration

If the study is to make comparisons between methods of process it should take into account variations in duration of treatment and frequency of sessions. This can be done either by comparison of services with similar treatment duration or statistically controlling for the effect.

Data attrition

Researchers need to deal with missing data at different points:

- contacted, not assessed
- assessed, not treated
- drop-outs
- no follow-up.

The last group is especially important for the correct interpretation of the results.

With a longitudinal design the research usually starts with an initial sample of people entering treatment. But, if only 70 per cent of this initial sample is available for follow-up then there is no information about the outcome for the other 30 per cent.

It is likely that the outcome for those that received treatment but cannot be followed up differs systematically from the outcome for those who are still in contact with the service. Thus, large drop-out rates threaten the validity of a study.

Validity, reliability, objectivity

Validity, reliability and objectivity are quality indicators for an evaluation instrument. Validity is about whether the instrument is able to measure what it aims to measure.

Reliability tells whether different assessors would achieve the same results.

Objectivity implies that the instrument does not leave room for subjective interpretation. For example, if an instrument asks for the frequency of drug use and gives the scale 'seldom, often, or always', one person may think three times a week is seldom whereas another might consider this as quite often.

Scaling

Another important issue is the type of answer the instrument asks for. There are several possibilities:

- simple yes/no questions
- continuous scales (how many days in the past four weeks have you used this drug, answer 0-28)
- categories (never, sometimes, often, very often, every day).

For efficiency, simplicity of statistical analysis, and data quality it is best to use as many continuous and multiple-category questions as possible and avoid overusing yes/no questions.

Analytical expertise

Often, service based research runs into problems because decisions about how the data is analysed have been left too late. Non-research organisations need to make sure, prior to setting up data collection, that they have access to suitable expertise to deal with statistical analysis. Results are only reliable if the appropriate inferential analyses are chosen ■

Evaluating treatment outcome – recommendations

Be clear about what you want to measure

Define your target outcomes

If possible use a control group

Make sure your sample size is large enough

Think carefully about the time of reassessment

Chose suitable answer formats for the kind of analyses you want to do

Chose carefully which instrument you want to use

Consider what you want to do about drop-outs

Make sure you know how to analyse your data before you start or that you can get suitable expertise

1. Ghodse H. 'When too much caution can be harmful'. *Addiction*: 1996, 91(6), p.764-766.
2. Donabedian A. *The definition of quality and approaches to its assessment*. Ann Arbor: Health Administration Press. 1980.
3. Meier P. Donmall M. & Millar T. *Measurement of drug misuse treatment outcome: a training guide*. Manchester: University of Manchester 2000.

The Drug Misuse Research Unit of the University of Manchester has published a training pack[3] on treatment outcome, which summarises available literature, and explains evaluation instruments used in the UK. It can be downloaded from the Drug Misuse Research Unit website: http://www.medicine.man.ac.uk/epidem/dmru/ or sent on request against a small charge for printing costs phone 0161 772 3790.

The Drug Misuse Research Unit is conducting an ongoing study, the Treatment Outcome Project (TOP).

DrugScope has published *QuADS*, which is a set of measurable standards agencies can use to evaluate their service provision phone DrugScope publications 020 7928 1211 fax 020 7928 1771 email: services@drugscope.org.uk website: www.drugscope.org.uk