

EVALUATION OF TREATMENT PROGRAMMES FOR DRUG-DEPENDENTS IN SPAIN

José Ramón Fernández Hermida and Roberto Secades Villa
Universidad de Oviedo

Studies designed to evaluate programmes for the treatment of drug-dependency are of obvious interest and relevance at the present time. The development of a research protocol to deal with the most common methodological difficulties in this type of study is a priority, not only for researchers and professionals, but also for the administrations, institutions and services involved. It is with this objective that studies carried out in Spain up to the present time are reviewed, analysing the methodological problems most frequently encountered. Finally, some recommendations are made, with the aim of minimising the effects of the problems and increasing the internal and external validity of evaluations.

La realización de estudios de valoración de programas de tratamiento en drogodependencias es una tarea de evidente interés y oportunidad en estos momentos. El desarrollo de un protocolo de investigación que subsane las dificultades metodológicas más comunes en este tipo de trabajos es una prioridad, tanto para los investigadores y profesionales como para las administraciones, instituciones y servicios implicados. Con esta finalidad se pasa revista a las investigaciones españolas que se han realizado hasta la actualidad y se analizan las dificultades metodológicas más comunes. Por último, se dan una serie de recomendaciones que minimicen los efectos de los problemas e incrementen la validez tanto interna como externa de las evaluaciones que se realicen.

The evaluation of treatment programmes for drug-dependents is not a new concept. Since the early 1970s, both in Spain and elsewhere, a great deal of research has been carried out aimed at determining the effectiveness of a variety of systems and procedures for eliminating addictive behaviour and facilitating social and employment reintegration in subjects with heroin addiction or multiple addiction with a strong bias towards heroin. These studies include among their objectives the evaluation of the effectiveness of one or several programmes (or the differential effect of some versus others), and offer results of follow-up studies on a wide variety of cohorts of patients.

The aim of this review is to outline the main features of the studies carried out in Spain and to indicate the essential conditions of this type of study and the strategies necessary for achieving them.

For the purposes of this work, we shall only consider as "treatment evaluation studies" those that identify this as their explicit aim, and that possess certain minimum methodological characteristics, as described below. Consequently, we shall not include follow-up studies

whose sole aim is to monitor the evolution of a group of patients over a given period of time. Also excluded are those studies that evaluate treatment with agonists, since this type of programme pursues different therapeutic objectives, and therefore requires different evaluation methodology and criteria. For example, a maintenance programme with methadone may lack well-defined time limits, being judged fundamentally for its achievement of a high rate of retention and a significant reduction in criminal behaviour (Roca, Gómez and Paños, 1989). In such conditions, it is more appropriate to state that the object of the evaluation is an addictive course with a legal substance (substituting heroin) than a therapeutic procedure with an established temporal framework that seeks, among other things, "drug-free" behaviour.

ANTECEDENTS

The pioneering and reference studies in this field are from the United States. The DARP (Drug Abuse Reporting Program) and TOPS (Treatment Outcome Prospective Study) programmes, costly and wide-ranging, are classic examples of this type of work. Financed initially by the NIMH (National Institute of Mental Health), the DARP programme (Simpson and Sells, 1983) recorded data up to 1973. In the course of its work it examined 44,000 people that had been subjected to one or more of the following treatments: methadone, therapeutic community, outpatient disintoxication program-

The original Spanish version of this paper has been previously published in *Psicothema*, 1999. Vol. 11, nº 2, 279-291

.....
Correspondence concerning this article should be addressed to José Ramón Fernández Hermida. Facultad de Psicología, Universidad de Oviedo, 33003 Oviedo. Spain. E-mail: jrhermid@correo.uniovi.es

mes and day centres with drug-free treatments, in a total of 52 different public centres. The results of patients that completed the treatment were compared with those of patients that had dropped out. Also analysed were the differences between the different treatment modes. The TOPS programme (Hubbard, Marsden, Rachal, Hardwood, Carvanaugh and Ginzburg, 1989) began collecting data in 1979, from a total of 11,000 patients, to whom a follow-up was applied, involving interviews at 3, 6 and 12 months after discharge. Results of the treatments, as in the case of the DARP, were compared with one another and with those of the control group. The data from both TOPS and DARP appear to support the idea that all treatments work well, giving results superior to those of drop-outs; the most effective programmes are those that succeed in maintaining patients within the programme for the longest periods.

In Spain, few studies have been carried out with the aim of evaluating the effectiveness of treatments, despite the fact that, since the 1970s, the public funds made available for this purpose have increased considerably. In this context, the remarks of Emiliano Martín, Assistant Director of the National Anti-Drugs Plan, are revealing: "We have passed through a phase of initiation and growth of programmes, especially assistance programmes, and we have now reached a stage of consolidation of such programmes and of the improvement of their quality and effectiveness" (Matellanes Matellanes, 1998, p. 40). Quite probably, the need to channel the available resources into satisfying the growing treatment needs of a burgeoning and increasingly demanding population leaves little room for research. In any case, the scarcity of studies carried out and their limitations are notable. Lack of resources and methodological problems with serious design deficiencies are the most common blights.

Below we outline the methodological characteristics of a typical study for the evaluation of drug-dependence treatments and the most common limitations and methodological problems. This is followed by an analysis of the different studies published in Spain.

CHARACTERISTICS OF STUDIES FOR THE EVALUATION OF DRUG-DEPENDENCE TREATMENTS

In general, this type of study is designed according to the following scheme:

1. Follow-up of a cohort of patients that have completed a treatment programme in a given period. Data-collection is carried out before, during (in some cases) and after the treatment. The number of times

data is obtained varies from one study to another, though there should always be pre- and post-treatment measures. Studies may be single-centre or multi-centre, and may be devoted to one or more types of treatment.

2. Measurement of the variations occurring in, at least, the following areas: drug consumption, productive activity (work and/or study) and criminal activity. These would be the main variables, since they would normally be considered priority objectives of the therapy. Also recorded may be sociodemographic data, health, family situation, interpersonal relations, leisure time occupation and, in some cases, psychological (psychopathological) variables.
3. Use of control group. This aspect is especially problematic, given the ethical difficulties involved in assigning patients to different treatments and control or placebo groups at random. Studies usually rely on using as control group those patients that, having begun the programme, drop out before completing it.
4. Comparison of post-treatment results with pre-treatment measures and follow-up data from the control group. It is supposed that the positive effects of the intervention will be more stable the more time that elapses between discharge and the date of the post-treatment measures used for the analysis.

METHODOLOGICAL PROBLEMS AND LIMITATIONS IN THIS TYPE OF STUDY

Despite the apparent simplicity of the design, there are several difficulties involved in carrying out studies for the evaluation of drug-dependence treatments. Melgarejo, Sanahuja, Masferrer, Sala, Pérez and Piña (1988, p. 17) list some of the most relevant: "a) Use of retrospective instead of prospective designs. Assignment of users to different treatment groups; b) Control of the variance of the different treatment groups; c) Use of indicators with poor sensitivity in the evaluation of results; d) Difficulties in locating patients; e) Difficulty in evaluating the reliability and validity of the data".

Sánchez-Carbonell and Camí (1986), in a previous article, pointed to similar design difficulties, among them: a) Lack of a control or comparison group; b) Bias in the sample studied, data being collected only from the cases that present the best evolution; c) Validity and reliability of the information; d) Disparity in the comparison periods from one study to another, with time intervals ranging from one year to ten; e) Disparity between stu-

dies in the parameters used for measuring characteristic variables, such as drug consumption, criminal behaviour, etc.; f) Markedly different criteria of success or cure from one study to another.

These difficulties would hinder comparative interpretations of results and impede, in some cases, the drawing of minimally acceptable conclusions.

There are some important reasons why these methodological problems are common in the field of the evaluation of drug-dependence treatments. Among these, three should be highlighted. First, the need for long evaluation processes with high research costs, which should ideally be the responsibility of institutions independent of those studied, and with no short-term economic or political interests, so common in the field of drug addiction. Second, the peculiar features of the population studied, frequently characterised by social rootlessness, broken homes, chronic illness, etc., which renders extraordinarily costly and difficult any research involving prolonged follow-up. Finally, the characteristics of the treatment programmes themselves, widely disparate in both basic philosophy and method, making problematic any process of standardisation of criteria or assessment instruments.

As pointed out in the previous section, a study for the evaluation of drug-dependence treatments should have, as part of its design, pre- and post-treatment data, measures of both addictive behaviour and work and criminal behaviour, a control or comparison group, and comparative analyses between groups. On the basis of these minimum requirements, we shall discuss briefly the most important methodological difficulties in terms of the effect they have on the validity of the treatment evaluation (Table 1).

1. Reliability and validity of data collection is affected by at least two factors: the nature of the measurement instruments and the relationship between the researchers and the institution in which the treatment is carried out.

As far as measurement instruments are concerned, studies use more or less structured interviews designed *ad hoc*. This is due to the fact that the programmes vary in terms of content and context of the treatment (outpatient treatment with or without drugs, therapeutic communities, treatment with drug substitutes and disintoxication programmes), objectives (e.g., abstinence or damage limitation) and patient characteristics. The particular and unique nature of some treatment programmes (and therefore their evaluation) precludes the use of assessment instru-

ments with guarantees, already proven in other studies. Data that was reliable in the sense of being extracted from an instrument with guarantees, but without a clear reference to the reality of the programme or the research objectives, would be of little use. Consequently, it is difficult to design a measurement instrument valid for all existing programmes. The principal virtue of these *ad hoc* instruments is that they are specifically adapted to the objectives of the research, but their main drawback is that their psychometric quality is usually unknown, as regards both validity and reliability. A method frequently used for determining the validity of an instrument is to compare the responses given in the interview phase, relating to consumption, work activity or family situation, with other measures, such as urine analyses, information from the family or data from the official records of institutions

Despite the fact that research attempts to adapt the instrument to the specific needs of the study, on some occasions the problem is situated on the other side. An example of this is provided by cases in which the measurement instrument used in the pre-treatment phase is the same as that used in the follow-up. Pre-treatment and follow-up are two different phases with different objectives, and involve different information. In the initial evaluation it is fundamental to collect data necessary for getting to know the patient and planning the treatment (socio-demographic data, consumption history, previous treatment, family consumers, etc.), information which is of no use in the follow-up, where the relevant data refer to variation, from the moment of discharge, in other types of variable considered sensitive to the treatment programme, such as drug consumption, productive activity (work or study) and criminal activity. Furthermore, the majority of studies concern themselves with discovering the current or recent behaviour of the subject in relation to any type of drug consumption, even though the

Table 1
Most frequent methodological limitations in studies for the evaluation of drug-dependence treatments

Threats to internal validity	1. Reliability and validity of data collection
	2. High statistical mortality
Threats to external validity	3. Effects of interaction between selection bias and dependent variable.
	4. Reactive effects of the programme.
	Effect of shaping the response or of rejection due to some of the questions asked.

consumption revealed may have been an isolated event, with little significance from the clinical point of view. Nevertheless, it is necessary to assess the patient's previous behaviour to determine whether these changes are the result of the treatment received or are due to other factors, such as other treatments subsequent to the programme under assessment. The fact that a subject has spent, say, six years without consuming drugs since discharge from the assessed programme does not allow us to infer that his or her current state is the result of that treatment, rather than being a consequence of other changes (social, personal, further treatment, etc.), unrelated to the programme, that have taken place in the intervening six years.

However, in many cases, the main cause of bias in data-collection derives from the fact that the data is recorded by the staff of the institution that applies the treatment programme, with no type of control over the distortions that may affect the obtaining and interpreting of the information received. When the same people are responsible for both the application of the treatment and the collection of data, the advantage lies in the possibility of improving the precision of that which is measured (given the knowledge available on the evolution and state of the patient); however, considerable bias may also be involved, since the person obtaining the information belongs to the institution and the programme under assessment. Whilst the level of knowledge of the patient's state for an external observer can be improved through various data-collection strategies, the bias involved in internal assessment is extremely difficult to control. It is therefore advisable to use external evaluation, carried out by researchers independent of the programme's design and application, so that the objectivity and credibility of results and conclusions is maximised.

2. Loss of subjects, in the samples or populations studied, is one of the main problems facing the assessment of drug-dependence treatment programmes. This statistical mortality leads to considerable bias that is itself difficult to evaluate. For some authors (Gould and Lukoff, 1977; Apsler and Harding, 1991), patient loss over the course of follow-up means that only those that present adapted behaviour are assessed, since they are easier to locate and more willing to collaborate. On the other hand, Desmond, Maddux, Johnson and Confer (1995) maintain that rehabilitated subjects are more diffi-

cult to locate, as they no longer keep in contact with the institutions that serve as points of capture for assessment studies. In such conditions, it is a highly complex task to interpret a statistical mortality that nevertheless may be as high as 40 or 50 percent. Polich, Armor and Braiker (1980) suggest a loss rate of no more than 20% as the objective of any evaluation study, whilst Gould and Lakoff (1977) clearly state that once this rate is exceeded efforts must be made to reduce it, since it would be inadvisable to analyse results in a state of ignorance of the direction of their bias. There is no reliable criterion for determining a figure, but it would seem reasonable to suggest, as these authors do, that at least 80% of subjects initially included in a study complete it. Nevertheless, it is quite clear that discussions on maximum permissible levels of losses could be endless, since there is no reliable criterion for stating whether or not a maximum loss of 20, 30 or 40 percent is acceptable and permits valid interpretation of the data.

3. Studies on the effectiveness of treatment programmes should take into account the bias produced if the analyses include only those subjects that have obtained the most favourable results and benefited most from the treatment. This is especially true in the field of drug-dependency, since for ethical reasons it is not possible to assign subjects at random to different treatment groups or to a control group. In the admission to treatment programmes there are strict (if not necessarily explicit) protocols whereby subjects are selected according to the level of success they are expected to achieve, with rejection of those for whom the programme "will not work". Moreover, once the treatment has begun, its requirements in terms of devotion of time and resources also imply a selection mechanism, resulting in the expulsion of those that fail to fulfil the conditions set, or whose evolution does not fit the model. This selection bias and its interaction with the dependent variable makes it difficult to generalise results, since what is demonstrated is that the programmes work for those for whom they work, a useless tautologism as far as research is concerned. In this context, the problem is that apart from those programmes with some kind of regional or administrative admission criteria, and in which inclusion is not linked to high expectations of success (which is neither logical nor ethical), there is no deontological possibility of defence against this attack on the

external validity of assessment programmes. In any case, it would be advisable for the programmes assessed to state clearly and explicitly the requirements for admission and continuance, so that the effects on the external validity of the results can be evaluated.

4. It is not frequent in research on the methodology followed in studies for the assessment of drug-dependence programmes to control the bias produced by subjects' reaction to the assessment study itself. This is a not inconsiderable effect that we have identified in our own research on the results of the *Proyecto Hombre* ("Project Mankind") treatment program in the region of Asturias (Spain), which is still being carried out (Fernández Hermida, Secades, Magdalena and Riestra, 1998). One of the instruments we intended to use for validating responses to the questionnaire was urine analysis. When subjects were told that this analysis was going to be carried out, their reluctance was so strong that it was considered inadvisable not to do so, as use of this instrument could affect the already complicated process of obtaining the sample. News of other aspects of the research, such as the collection of data on HIV, quickly spread among the subjects who were to be interviewed, which may have had effects on subject loss that are difficult to interpret. This is a little-studied aspect that requires more attention, since it can greatly affect the possibility of generalisation of the results obtained.

STUDIES IN SPAIN

In this section we shall review the studies that have been published in Spain, identifying their principal characteristics and limitations, in the light of that which has been discussed above.

The first observation researchers will make is the extreme scarcity of studies in this field in Spain. There are a considerable number of follow-up studies (Guardia Serecigni and González, 1983; Sánchez-Carbonell, Brigos and Camí, 1989; GID, 1992; García Martín, Roldán and Comas, 1995; Melús Moreno and Gutiérrez, 1996; Marina, Vázquez, Jiménez and Erkoreka, 1996; Comás, García and Roldán, 1996), which are usually referred to in the sections on evaluation of the programmes, but which will not be included in this review for two reasons. The first reason is that follow-up studies are considered to have different objectives from those of treatment evaluations. In the former, the priority is to obtain information on patients' evolution and its rela-

tionship to different pre-, intra- and post-treatment variables; in the latter the aim is to determine whether the treatment under study is preferable to non-treatment (although comparisons may also be made with other treatments, including, in theory, placebo), in terms of their effects on the modification of certain variables already mentioned above. The second reason for excluding follow-up studies is that their different objectives require different methodologies, such as the use of, at least, a control or comparison group. These differences mean that any joint evaluation of the two types of study would be confused and of little informative value.

Thus, if we eliminate follow-up studies, to our knowledge there remain only two that should concern us here. The first is the study for the evaluation of the Dross programme in Cataluña (Melgarejo et al., 1988) and the second is that carried out on the treatment at the Provincial Drug-dependency Centre (*Centro Provincial de Drogodependencias*, CPD) in Cádiz, southern Spain (Girón García, 1997; Girón García and Martínez, 1998). Some of the characteristics of these two studies are shown in Table 2.

In both cases it is clearly stated that the study aims to assess a treatment programme, and control groups are used that have received no treatment or that have been treated for a period insufficient to be considered effective. The follow-up period in these two studies differs greatly: while in the case of Melgarejo et al. (1988) it is one year, in that of Girón García (1997) it is six years. This difference would appear to be reflected in the percentage of subjects with favourable evolution, which is far higher one year after treatment. In both cases there is a notable difference in the number of subjects included in the control groups, which are very small, compared to the numbers in the groups of discharged patients (see Table 2), and this can lead to certain difficulties in interpreting the results. In either study there appears to be a differential effect of treatment on the diverse variables analysed, as compared to non-treatment, which would indicate the effectiveness of the therapeutic services studied.

However, the two studies have some important shortcomings. If we analyse the results in terms of the biases described in the previous section, we find that both that of Melgarejo et al. (1988) and that of Girón García (1997) are subject to some of the problems affecting internal and external validity.

In the study by Melgarejo et al. (1988), data-collection was entrusted to a group of researchers that had to administer a questionnaire to patients, either in person or by

telephone, without any type of cross-validation measure (urine analysis, use of external records, family information, etc.). In the study by Girón García (1997), the instrument for collecting data in the interviews was constructed on the basis of some test items, such as the ASI (Addiction Severity Index), the results referring to recent consumption being compared with those from a urine analysis carried out after the interview. None of the information relating to other variables analysed (family, employment, sociolegal situation, etc.) was checked against external records.

Statistical mortality was in neither case very high, if we take into account the results of other research in the field, such as the DARP (Simpson, 1993), with 39% of patients lost, and the TOPS (Hubbard et al., 1989), with losses of between 42 and 29 percent. Even so, losses were above the level recommended by authors such as Gould and Lukoff (1977), who set a maximum acceptable figure of 20%.

In neither of the two studies is any mention made of the possible effect on results of the mechanisms of selection for and admission to therapy. Nor is there any reference to the potential effect of the evaluation procedure on the sample studied, and how this may influence the results of the evaluation.

Another shortcoming of these studies is the lack of analysis of the active principles of the treatment that explain its differential effect compared to non-treatment. The mere confirmation that “the longer the time the subject has remained on the programme the lower the probability of relapses” explains very little if there is no control of the mechanisms of admission and exit from the programme. The explanation becomes completely circular: they are abstinent and they fit the programme’s criteria so they remain in it for longer, and because they remain in it for longer they stay abstinent and fit more closely the improvement criteria.

There is a need in the reports of the two studies for a more detailed description of the therapeutic procedures assessed. In particular, it would be appropriate to inclu-

de a more detailed account of the mechanisms for selecting patients and of the therapy or therapies used in each case, with a view to assessing improvement effects within this framework.

RECOMMENDATIONS FOR CARRYING OUT STUDIES FOR THE EVALUATION OF DRUG-DEPENDENCY TREATMENT PROGRAMMES

As can be deduced from that which has been discussed up to now, very few studies for the evaluation of drug-dependency treatment programmes have been carried out in Spain; moreover, those that have been implemented have been hindered by certain problems that have made their interpretation difficult and reduced their general utility. Any study carried out in the field of treatment assessment must comply with the methodological criteria of this area of research (Fernández Ballesteros, 1995). Some of these have already been mentioned, such as the need for external evaluation, for the assessed programmes to be described in detail, and for the clear specification of the objectives of the assessment carried out.

Furthermore, there are some specific recommendations that should be taken into account by researchers in the field of drug-dependency, related to general planning of the research and precise methodological issues. In the following section we shall review some of these.

Prospective or retrospective research

The majority of research carried out both for the study of the natural evolution of addiction and for the assessment of treatment has been of a retrospective nature. In no case has provision been made for an assessment process within the treatment programme itself. This is an important drawback, as it constitutes an obstacle to the connection between therapeutic operations and results, and to the corresponding adjustment of the assessment process to the different stages of treatment. In the conditions in which this type of assessment is usually made, only a prospective study could provide information on the therapeutic operations that maximise effectiveness of the intervention and its appropriateness for the profile of each patient. Retrospective studies (given that they evaluate the results of a complex problem) allow, at the most, to report on the effectiveness of a service and of a treatment programme as a whole, excluding from the analysis the variations of patients’ individual responses to each therapeutic stage of the programme. Another beneficial effect of prospective research is that it permits the optimisation of strategies for reducing statistical

Table 2
Characteristics of studies for the evaluation of drug-dependency treatment programmes carried out in Spain

Study	DROSS	CPD-Cádiz
% of interviews out of total initial sample	74.7	75.3
N of discharge sample	136	65
N of control group	23	7
Follow-up period	1 year	6 years
% of favourable evolution in discharge sample	77.8	55.1
% of favourable evolution in control group	47.8	29.6

mortality (informing patients that they will be assessed after discharge, obtaining precise information for their subsequent location, etc.).

Assignment of subjects to groups. Control or comparison group

There are unavoidable ethical difficulties involved in the random assignment of addict subjects to the different treatment groups and the control or non-treatment group. The inclusion of subjects in the different treatment groups is made according to subjects' characteristics and their addiction, always with the intention of obtaining maximum benefit and reducing the risk of abandonment and relapse. Thus, it is quite obvious that a patient cannot be assigned to a control group that will receive no effective treatment. Given this situation, it is difficult to establish the basic conditions that make possible a comparison between groups of patients, in which the differences between them would be attributable to the therapeutic operations to which they have been subjected, and not to the previous conditions required for forming part of them. This is especially the case given the impossibility of forming a pure experimental control group.

The solution offered for these problems is not a definitive one, but it may mitigate some of the negative effects on the validity of results. This solution has two conditions. On the one hand, the treatment groups should have explicit conditions (made explicit in an operative way) of admission and exit of the subjects of each intervention. In this way it would be possible to estimate the influence of subjects' particular conditions on the therapeutic results. On the other hand, it is necessary to form a comparison group made up of subjects that, although identical to those of the treatment groups in terms of indicators related to the evolution of the addiction, have not received significant therapy, and whose situation, therefore, can be considered as following a "natural course" of addiction. These types of subject tend to be found among those that, having fulfilled the conditions for receiving treatment and admitted to a programme, abandon it after a short time.

Minimising statistical mortality

The presence of a large number of subjects that are impossible to locate, or that refuse to participate in the treatment assessment, constitutes a risk of bias in the results that it is difficult to evaluate. In a work devoted specifically to this problem, Desmond et al. (1995) recommended the following measures: a) To collect as much information as possible on the subject during the

treatment; b) To inform the subjects that they will be interviewed after the treatment; c) To give material incentives to subjects for participating in the research; d) To select suitable staff for carrying out the follow-up; e) To keep a detailed record of the activities carried out during the research work; f) To use information from other institutions; g) To carry out a brief interview; h) To carry out the interview as soon as possible after location of the subject; i) To make it easy for the subject to attend the interview; j) To allow plenty of time for carrying out the field work.

Some of these recommendations are intended to make subjects more motivated to participate, whilst in others the aim is to compensate for the lack of information resulting from the inevitable loss of part of the sample. The authors of this work (Desmond et al., 1995) maintain that the implementation of these measures has produced significant reductions in statistical mortality, keeping it within the recommended limits, according to the results of a field study.

In any case, if considerable losses of subjects occur, a useful procedure for palliating this problem is the comparison of subjects lost to the study with those that participate in it, in terms of a series of basic variables (pre-treatment consumption history, sociodemographic data, criminal record, etc.). The differences found in these types of variable may provide clues as to the direction of bias in the sample (Caspar, 1992).

Validity and reliability of the data. Assessment instruments

There would appear to be a general consensus on the main variables to be studied and certain disagreement on methods and measurement instruments. Consumption behaviours, together with the data referring to work or criminal activity, are included in practically all follow-up and treatment evaluation studies, since they relate to a variety of objectives pursued by the majority of therapeutic programmes in this field. Nevertheless, it is highly recommendable to collect data on the patient's evolution corresponding to the follow-up period, and not only on the point at which the assessment is made. In this way, more reliable conclusions can be drawn with regard to the effects of the programme assessed, and firmer relationships established between the variables studied (consumption, work activity, criminal behaviour, etc.) and other types of event subsequent to completion of the programme.

As far as instruments are concerned, disagreements are more evident. The use of tests providing results that can

be compared from one assessment to another is a highly desirable goal within this field of research. It would not appear to be an easily achievable one, however, given that each study seems to require the adaptation of existing instruments to the peculiarities of the treatment under analysis. Graham (1994), taking this into account, developed a series of general recommendations for the selection of measures and instruments that should be adhered to by all follow-up studies of addiction treatments. Nevertheless, there are certain areas, frequently included in research, in which, wherever possible, the use of standardised instruments should be considered that take account of the validity and reliability of results, two essential measurement properties in the social sciences (Muñiz, 1998). In this context we might mention, for example (and among many others), the BDI (Beck Depression Inventory) for the assessment of depression, the GHQ (Goldberg Health Questionnaire) for the evaluation of state of health, or the AUDIT for obtaining information on alcohol-related problems.

Reliability studies that compare self-reports with the results of urine analyses show a high level of agreement—over 90% (Maisto, McKay and Connors, 1990). Nevertheless, the validity and reliability of the data remain open to question. Studies that have related subjects' responses to questionnaire items with results obtained from other indicators (mainly consumption information validated through urine analysis) have been limited to short periods, and it is not clear that the validity and reliability of responses will be similar in studies in which subjects must report their behaviour (addictive, criminal, work, etc.) over a longer period of time (one year or more). In such cases, for example, subjects' reports of their behaviour cannot be validated through the use of a urine analysis at the time of the interview. In view of this, it is necessary for subjects' reports to be backed up by as much external validation as possible, using criteria that take account of the period to be studied. Two important sources of external information are institutional data (health service, judicial, employment, etc.) and data from the family context.

CONCLUSIONS

In recent years there has been a rapid proliferation of assistance programmes for drug-dependents in Spain. There is currently a consensus among the majority of professionals on the urgent need to develop studies for assessing the effectiveness of these programmes, with a view to making possible improved planning and decision-making. Nevertheless, studies for the assessment of

the results of treatment carried out in Spain have been scarce. In the majority of cases they are follow-up studies, which, given their methodological characteristics, can only inform of the state of subjects at a particular moment, and from which we can make no reasonable inferences about the effects of the intervention.

Moreover, most of these studies are subject to methodological limitations. Some of these limitations affect studies for the assessment of programmes in any area of research, while others are peculiar to the field of drug-addiction, such as those related to the population characteristics and to the wide conceptual variety of treatment programmes.

In this context, there is a need for the planning and design of studies for programme assessment that are well controlled and that adhere to the methodological conditions that have demonstrated their usefulness for dealing with the different problems that may present themselves. Some of the recommendations for such conditions have been referred to in this work, though it does not constitute an exhaustive compendium of solutions to all the problems to which this type of study is subject. In any case, it can be considered that observation of these recommendations will lead to more valid and informative research on the effectiveness of treatment in the field of drug-dependency.

Eventually, this line of research should lead us to the design and preparation of assessment protocols for each one of the treatment modes currently in use (programmes with agonists, therapeutic communities, outpatient programmes with antagonists, etc.). This is a suggestion that coincides, moreover, with that which is proposed for other areas of social intervention (Fernández del Valle, 1992). Even so, a first essential step would be the consolidation and homogenisation of the content and objectives of each type of programme.

REFERENCES

- Apsler, R. and Harding, W.M. (1991). Cost-effectiveness analysis of drug abuse treatment: Current status and recommendations for future research. In N. I. D. Abuse (Ed.), Background papers on drug abuse financing and services research (Vol. NIDA Drug Abuse Service Research Series No. 1, DHHS Publication No. ADM 91-1777, pp. 58-81). Rockville, MD: U.S. Department of Health and Human Services.
- Caspar, R. A. (1992). Follow-up of Nonrespondents in 1990. In C. F. Turner and J. T. Lessler (Eds.), Survey Measurement of Drug Abuse. Methodological Studies. Rockville, Maryland: National Institute on

- Drug Abuse. Division of Epidemiology and Prevention Research.
- Comas Arnau, D., García Martín, E. and Roldán Inchusta, G. (1996). Políticas sobre drogas a partir de los resultados de un estudio longitudinal retrospectivo de adictos a la heroína (1982-1992). *Adicciones*, 8(1), 75-89.
- Desmond, D. P., Maddux, J. F., Johnson, T. H. and Confer, B. A. (1995). Obtaining follow-up interviews for treatment evaluation. *Journal of Substance Abuse Treatment*, 12(2), 95-102.
- Fernández Ballesteros, R. (1995). Cuestiones conceptuales básicas en evaluación de programas. In Fernández Ballesteros, R. (Ed.) *Evaluación de programas. Una guía práctica en ámbitos sociales, educativos y de salud*. Madrid: Síntesis.
- Fernández del Valle, J. (1992). Evaluación de programas residenciales de servicios sociales para la infancia. Situación actual y aportaciones de los enfoques psicología. *Psicothema*, 4(2), 531-542.
- Fernández Hermida, J. R., Secades Villa, R., Magdalena Benavente, Y. and Riestra del Rosal, C. (1998). La evaluación de programas de tratamiento para drogodependientes. Address presented at the V Congreso de Evaluación Psicológica, Benalmádena - Málaga.
- García Martín, E., Roldán Intxusta, G. and Comas Arnau, D. (1995). Un estudio de seguimiento de la Comunidad Terapéutica Haize-Gain. San Sebastián: Comunidad Terapéutica Haize-Gain.
- GID. (1992). Los estudios de seguimiento, una experiencia metodológica. Madrid: Fundamentos.
- Girón García, S. (1997). Estudio de Seguimiento de una Muestra de Heroínómanos a los seis años de iniciado el tratamiento. *Folia Neuropsiquiátrica*, 33(2), 17-46.
- Girón García, S. and Martínez Delgado, J. M. (1998). Seguimiento de una muestra de heroínómanos a los seis años de tratamiento: Relación del tiempo de permanencia con la evolución. *Revista Española de Drogodependencias*, 23(1), 25-34.
- Gould, L. C. and Lukoff, I. (1977). Selecting a study design. In L. D. Johnston, D. N. Nurco, and L. N. Robins, (Ed.), *Conducting follow-up research on drug treatment programs* (Vol. DHEW Publications No. ADM 77-487, pp. 29-46). Rockville, MD: U.S. Department of Health and Human Services.
- Graham, K. (1994). Guidelines for using standardized outcome measures following addictions treatment. *Evaluation and the Health Professions*, 17(1), 43-59.
- Guardia Serecigni, J. and González Navarro, S. (1983). Estudio de seguimiento (un año) de adictos a la heroína. *Revista del Departamento de Psiquiatría de la Facultad de Medicina de la Universidad de Barcelona*, 10(6), 357-381.
- Hubbard, R. L., Marsden, M. E., Rachal, J. V., Harwood, H. J., Carvanagh, E. R. and Ginzburg, H. M. (1989). *Drug abuse treatment: A national study of effectiveness*. Chapel Hill: University of North Carolina Press.
- Maisto, S. A., McKay, J. R. and Connors, G. J. (1990). Self-reports issues in substance abuse: state of the art and future directions. *Behavioral Assessment*, 12(2), 117-134.
- Marina, P. A., Vázquez Barquero, J. L., Jiménez, J. M. and Erkoreka, I. (1996). Los beneficios de la abstinencia: Un estudio de seguimiento de adictos a la heroína. *Adicciones*, 8(3), 295-309.
- Matellanes Matellanes, M. (1998). Interview with D. Emiliano Martín González. Subdirector General del Plan Nacional sobre drogas. *Infocop*, (2), 37-41.
- Melgarejo, M., Sanahuja, P., Masferrer, J., Sala, Ll., Pérez, F. and Piña, L. (1988). Tratamiento de 182 heroínómanos y resultados: Seguimiento a un año (1985-86). *Fons Informatiu*, 13(1), 1-140.
- Melús Moreno, R. and Gutiérrez Rosado, T. (1996). Programa de mantenimiento con Naltrexona: Seguimiento a los 12 meses. *Adicciones*, 8(1), 5-18.
- Muñiz, J. (1998). La medición de lo psicológico. *Psicothema*, 10(1), 1-21.
- Polich, J. M., Armor, D. J. and Braiker, H. B. (1980). *The course of alcoholism: Four years after treatment*. Santa Monica, CA: The Rand Corporation.
- Roca, F., Gómez Beneyto, M. and Paños, M. (1989). Evaluación de los programas de mantenimiento con metadona desarrollados en Valencia desde 1983 hasta 1987. *Comunidad y Drogas*, 13(14), 45-52.
- Sánchez-Carbonell, J., Brigos, B. and Camí, J. (1989). Evolución de una muestra de heroínómanos dos años después del inicio del tratamiento (proyecto EMETYST). *Medicina Clínica*, 92(4), 135-139.
- Sánchez Carbonell, J. and Camí, J. (1986). Recuperación de heroínómanos: definición, criterios y problemas de los estudios de evaluación y seguimiento. *Medicina Clínica*, (87), 377-382.
- Simpson, D. D. (1993). Drug treatment evaluation research in the United States. Special Series: Psychosocial treatment of the addictions. *Psychology of Addictive Behaviors*, 7(2), 120-128.
- Simpson, D. D. and Sells, S. B. (1983). Effectiveness of treatment of drug abuse: An overview of the DARP research program. *Advances in Alcohol and Substance Abuse*, (2), 7-29.