

---

# **4. Methodological Issues in Evaluating the Effectiveness of Alcoholism Treatment**

# Methodological Issues in Evaluating the Effectiveness of Alcoholism Treatment

---

The development of a body of research on alcoholism treatment is fairly recent, having occurred mainly during the past 20 years. Until the 1950's, treatment for alcoholism was more likely to have been incarceration or custodial care in State mental hospitals than to have been medical or psychological therapy (cf. 325). Thus, the lag in the development of a scientific research base is not surprising.

Even today, despite the rapid growth of a formal treatment system, evidence of treatment effectiveness<sup>●</sup> is often based on unsystematically collected data (46,76,135,297). Because of the limitations of available research, conclusions about the effectiveness of treatment for alcoholism are necessarily limited, although some tentative conclusions can be drawn. Such conclusions are presented in chapter 5. The present chapter analyzes the methodological problems in conducting and

---

● Effectiveness is the benefit as measured under average conditions of use (229). Efficacy is the health benefit as measured under controlled conditions such as those in a randomized clinical trial.

analyzing alcoholism treatment research. The goal in this chapter is to place the current state of scientific knowledge about treatment alternatives into a research perspective.

Because patient characteristics, treatment settings, services offered, and practitioner characteristics interact to affect treatment outcome, research on alcoholism treatment is complicated. The evaluation issues, however, are parallel to those involved in assessing other health care interventions (see, 226,227,229).

In assessing the quality of research conducted on alcoholism treatment effectiveness, the validity of the research evidence and the ability to generalize from it will be emphasized (229,350). For research to be valid and permit generalizations to be drawn, there must be clarity about what is being tested, what is being compared, which subject populations are involved in the research, and what is being measured. Operationally, these four factors refer to: 1) treatment design, 2) research design, 3) sampling, and 4) outcome measures.

## TREATMENT DESIGN

Treatment design issues involve the extent to which clarity about the "active ingredients" of the program being tested can be achieved. Questions such as whether the program involves a single treatment, a combination of treatments, or a combination of treatment and nontreatment factors must be answered. Often, because alcoholism treatment programs are multivariant, researchers *are* unable to identify separate effects (18,317). The extent to which researchers can measure the impact of any one component of the treatment is limited, of course, when all patients receive or have access to multiple components concurrently. Clarity about what the program includes is essential for attributing outcomes to particular treatments or treatment packages.

One solution to these problems is to group treatments in such a way as to be able to form treatment packages. Few would argue that a single treatment (e.g., psychotherapy) could alone reduce alcoholism problems (cf. 241). A single patient, especially when hospitalized or in a residential setting, may receive group therapy, antidepressants, and attend sessions of Alcoholics Anonymous (AA). There are various problems in analyzing treatment groupings, however. Lumping treatment programs under umbrella terms such as "inpatient" or "outpatient," without clarifying which specific services are offered or utilized, may obscure differences between treatment components. By lumping multiple treatment programs together, one is unable to decipher

which treatment is effective, for whom, and under what conditions (cf. 306).

Furthermore, even if coherent treatment packages can be developed, they may be difficult or undesirable to administer. This is particularly problematic if assignment of patients to treatment packages or components is required for research purposes. A research design that requires systematic assignment of treatment or segregation of services might undermine a basic treatment principle—that of involving patients in decision-making (306). Such research-based assignment criteria may also be troublesome for practitioners who, on the basis of clinical criteria, may want to control treatment regimens.

## RESEARCH DESIGN

A valid research design requires systematic comparisons. At minimum, the comparisons must involve a single group of patients measured before and after treatment. Optimally, they will involve two or more randomly assigned groups (an experimental group and a comparison or control group) of patients who are tested before and after treatment (see 350). The latter design is usually called a “true” experiment (67), or in health care research, a randomized clinical trial (RCT). RCTS are considered the most definitive experimental method for evaluating the efficacy or health benefits of a technology (229). The advantage of this design, in comparison to a nonrandom design, is that it allows differences in outcomes to be attributed more confidently to the treatment, and not to preexisting differences in the samples tested.

Evaluating research on the effectiveness of alcoholism treatment poses several difficulties for researchers interested in valid conclusions. Comparative data on treatment groups are typically not available, and most research merely tracks patients during and after treatment (135,189,297). For example, one study documented a 38-percent abstinence record at a 1-year followup for problem drinkers who received treatment at a 4-week residential treatment center (see 16). In the absence of comparative data—e.g., on individuals who did not receive treatment or who received alter-

The alcoholism field is rife with intense feelings about treatment effectiveness and safety. Individual clinicians can cite examples of patients who have faced death, high economic costs, or health problems because of irresponsible treatments. Practitioners often have intense opinions, convictions, and reservations about the use of particular treatments. Putting treatments to the test, limiting the treatments available, withholding services, or providing treatments presumed ineffective will be resisted by many practitioners (46,47,189). Thus, practitioners have had, and will continue to have, a strong influence on the type of effectiveness research conducted.

native treatments—it is not possible to determine if the observed 38-percent abstinence rate represents natural improvement or if a higher or lower abstinence rate would have resulted from another type of treatment.

Although the absence of comparative data is the most fundamental deficit of the literature on the effectiveness of alcoholism treatment (16,317), other methodological problems also limit the implications that can be drawn. First, data are often presented in aggregate form—i.e., data on patient outcomes are often not differentiated by severity of initial symptoms or other patient characteristics. Social class information may be lacking, and important subpopulation differences may be obscured. In a study that does provide such socio-demographic breakdowns, patients treated at the Raleigh Hills Fair Oaks Hospital in California were reported (221) to have had 1-year abstinence rates that ranged from 36 percent (for Medicare-eligible disabled patients) to 73 percent (for married, employed patients). Age, gender, and social situation seem to affect significantly treatment effectiveness.

A related issue is that multivariate analyses that are useful for examining differences by factors such as age, race/ethnicity, social class, employment status, sex, and disability are typically un-

available (cf. 77). Although there are several important exceptions, including a study by Armor and colleagues (13), studies that statistically control outcome data by demographic or other factors have not been conducted with many treatments. Such analyses present difficulties both in data collection and analysis and require large patient populations. The lack of controlled research hinders informed development of treatment strategies tailored to the needs of subpopulations.

Despite the methodological problems just discussed, alcoholism treatment researchers seek to

generate systematic, experimental designs with comparison group information and multiple, longitudinal outcome measures. Practical dilemmas, however, may undermine this aim. For example, random assignment of patients to conditions does not ensure that patients will accept their assigned treatment, nor that they will remain in treatment (18), although, in some cases, acceptance of or dropping out of treatment is a useful outcome measure.

## SAMPLING

Sampling refers to decisions concerning the subjects selected for research. Issues of sampling concern: 1) eligibility for treatment, 2) selection for participation in research, and 3) availability for followup research. If the general population of alcoholics is not represented in the research samples, or if certain groups (e.g., working class adults or women; cf. 18) are underrepresented, the ability to generalize from research findings is limited.

Perhaps the most important sampling problem is that individuals who receive treatment services cannot be assumed to form a representative group of problem drinkers (18). Many programs explicitly exclude those patients who have poor prognoses for recovery—particularly those from lower income groups and/or the unemployed. Even without exclusion criteria, individuals who elect treatment undoubtedly differ from those who do not (46,135). Those who receive treatment may be more visible (hence, their referral to treatment), more socially connected to others (who encourage treatment seeking), more motivated (and so seek treatment), and more confident of success (willing to undergo treatment). It is also possible that those who seek treatment see themselves as more helpless (and thus reliant on others for assistance), more intrusive (and so referred more readily into treatment), more troublesome (and, perhaps, pushed into treatment), or more abusive (and so more likely to be mandated into treatment). The absence of data on alcohol abusers who do not

seek treatment limits the ability to generalize and the establishment of realistic spontaneous remission rates (see 290,315).

Sampling biases involve not only who constitutes the client population, but who is available for and willing to be involved in research, especially in the case of research that involves follow-up and long-term commitment to a research project (cf. 18,189,272). The probability of obtaining a representative population of alcoholics in treatment and not in treatment is remote.

Even if one were to obtain a representative population of alcoholics or problem drinkers, differences remain in terms of which patients receive different treatments, which patients are available for research, and which patients can be followed up on. Mandell (189), among others, has demonstrated that middle- and upper-middle-class patients are more likely to receive treatments covered by private insurance policies and to be referred by employers. Lower-class alcoholics, in contrast, are more likely to receive services paid for by State and local governments. Although the evidence is not clear cut, there appear to be differences in what types of treatments are received by each of these groups. If, indeed, different kinds of patient groups receive different kinds of services, merely comparing the outcomes will convey little about the treatments' effects across patient populations.

Who is available for followup, and how that affects the results of outcome studies, is a problem that has plagued much outcome research. As Baekeland, Lundwall, and Kissin (18) point out, mortality rates for alcoholics are high. Alcoholics drop out, disappear, and reject treatment at numerous points throughout the process. If there are systematic differences between dropouts and those

alcoholics who remain in treatment (and it is reasonable to assume that there are), followup data are limited to an understanding of those who remain in treatment. There is some evidence that those who are difficult to follow up have the poorest treatment outcomes (204), although contrary evidence is also available (cf. 250).

## OUTCOME MEASURES

Finally, the way in which treatment outcomes are measured significantly affects the interpretation of alcoholism treatment research. Studies of alcoholism treatment often use indirect outcomes or a combination of outcome measures. Others may measure the same effect differently. Self-reports on drinking behavior, interviews with spouses, supervisor-based job productivity reports, blood-alcohol levels, psychological improvement, or even recorded attendance at a treatment center may be used as outcome measures. The degree to which different studies use different conceptual outcomes, operational outcomes, and measurement techniques limits the comparisons that can be made about treatments.

Much of the discussion about outcome measures has focused on self-reports. Self-reports are often believed to be low in accuracy (96). Alcoholics may deny that they have a problem (351), the alcohol may have affected their memory (25), and, in general, it is socially undesirable to report alcohol intake. Reporting high use may affect the patient's job and self-esteem, or maybe perceived as unhelpful to the researcher (cf. 18). However, despite sound reasons why self-reports should yield unreliable results, a number of studies report high concordance between self-reports of use and physiological measures (112,250), although physiological measures may be less-than-valid indicators (160).

Controversy over what should be the measure of successful outcomes in treating alcohol problems has not really been resolved. The criteria of abstinence from alcohol use has, traditionally, been used as the single measure of treatment effectiveness. Some studies have also measured various

behaviors related to drinking—e.g., frequency of drinking, number of ounces of alcohol ingested, number of binges, days of abstinence, number of relapses, and percentages of days without alcohol-related problems. More recently, other outcome measures, such as work adjustment, family adjustment, problems with the law, psychological well-being, and continuation of treatment, have been utilized. Physical health has been another important criterion and relates importantly to cost-benefit assessments of treatment (see ch. 6).

A major debate in recent years has focused on whether "controlled drinking" or "nonproblem drinking" can be considered a successful outcome of alcohol treatment (cf. 132). Sobell and Sobell (302), in particular, challenge the unitary view that alcoholism is a single syndrome whose treatment goal is abstinence (cf. 245). They argue that the exclusive use of abstinence as the outcome obscures partial improvement, neglects improvements in other areas of life functioning, is difficult to validate, and represents a narrow understanding of the multifaceted alcoholism syndrome. As described in chapter 5, however, the view that controlled drinking is the desirable outcome of treatment has been challenged by data indicating that the ability to learn "controlled drinking" is unrelated to long-term remission (246).

Pattison (237) and Gerard, Saenger, and Wile (113) present data indicating that abstinence does not necessarily result in improvement in an alcoholic's problems. In some cases, once abstinence is achieved, problems in other areas increase. The meaning of such outcome assessments is not clear. It may reflect either longstanding health problems or relapse. Emrick (94) reviewed 265 alcohol treat-

ment studies to test the relationship of drinking to other outcome measures. He found that in more than two-thirds of the cases, drinking outcome related positively to outcomes in other dimensions.

The solution to these methodological problems would seem to be multidimensional measurements of outcome. Indeed, such a recommendation is strongly encouraged by recent methodological reviews (47,102,297).

## CONCLUSIONS

Conducting outcome evaluation research on alcoholism treatments is difficult. Since these difficulties are reflected in current evaluations of the alcoholism problem, many presently urgent policy questions can probably not be answered by available research. Much of this research is flawed by problems in design, sampling, or outcome meas-

urement. Nevertheless, the “whole” of available research on alcoholism is probably greater than the sum of its parts. By carefully considering the results of individual studies, each of which handles somewhat different methodological problems, conclusions can be drawn from the substantial body of literature.