
Definition and Diagnosis of Relapse to Drinking

Richard K. Fuller

The issues of definition and diagnosis of relapse to drinking are somewhat intertwined because the methods for diagnosing relapse may differ depending on which definition is used. In the initial part of this report, there is a discussion of the various definitions of relapse used by alcoholism treatment researchers. Next is a discussion of the methods used to measure each definition. In the last section is a recommendation as to which definition would be most appropriate for defining relapse to drinking after liver transplantation. *Relapse* is a preferable term to *recidivism*. *Recidivism* is used in the criminal justice system. Because drinking is not a felony, and because alcoholism and liver disease are medical conditions, *relapse* is preferable.

Relapse to drinking is not the same as relapse to alcoholism itself. Therefore, the diagnosis of alcoholism is different than the diagnosis of relapse to drinking, which is the topic of this report. The standard diagnostic criteria for alcoholism are delineated in the latest version of the *Diagnostic and Statistical Manual of Mental Disorders* (DSM-IV).¹ Signs of alcoholism include tolerance; withdrawal; impaired control; neglect of activities; time spent in activities to obtain alcohol, to drink, or to recover from its effects; and continued drinking despite knowledge of physical or psychological problems caused by drinking. Relapse to drinking involves frequency and quantity of alcohol consumption. These parameters of drinking are not specifically part of the diagnostic criteria for alcoholism.

Definitions of Relapse to Drinking

The definitions of relapse vary widely from any deviation from abstinence from alcohol to consequences of drinking such as rehospitalization for alcoholism or physical and social consequences.

From the National Institute on Alcohol Abuse and Alcoholism, Bethesda, Maryland.

Address reprint requests to Richard K. Fuller, MD, National Institute on Alcohol Abuse and Alcoholism, Suite 505 Wilco Bldg, 6000 Executive Blvd, Bethesda, MD 20892-7003.

Copyright © 1997 by the American Association for the Study of Liver Diseases

1074-3022/97/0303-0011\$3.00/0

Intermediate definitions involve frequency of drinking, quantity of drinking, and combinations of frequency and quantity. Alcoholism treatment researchers call these definitions "outcome" variables because they are measures of treatment outcome. In 1991, one clinical trial of alcoholism treatment reported 10 outcome variables!² The authors used 10 variables because they were attempting to capture different aspects of drinking behavior.

Abstinence

Although there is no standard definition of relapse in treatment research, lack of abstinence is the most commonly used definition because of the emphasis placed on abstinence by treatment providers. Treatment providers emphasize abstinence because of the idea that the alcoholic, unlike the nonalcoholic, cannot voluntarily stop drinking after imbibing one or two drinks and that drinking one drink inevitably leads to heavy, destructive drinking. One of the criteria in the DSM-IV diagnostic scheme for alcoholism is "impaired control," i.e., the alcoholic drinks in larger amounts or over longer periods than he or she intends and/or unsuccessfully tries to cut down or control the drinking. Much of the evidence to support this criterion is based on the experience of alcoholics. In support of this concept is the classic longitudinal study of Vaillant,³ which found that "controlled" drinking cannot be sustained for long periods (more than 3 years) without the patient returning to alcohol abuse.

In addition to lack of abstinence, alcoholism treatment researchers use other measures to define a relapse to drinking because the sporadic ingestion of alcohol does not invariably lead to uncontrolled drinking in the short run (less than 3 years)³ and to associated negative physical and social consequences. There may be a threshold of drinking that is more than sustained abstinence but below which negative consequences do not occur. Reporting only the lack of abstinence may disguise the fact that considerable posttreatment improvement has occurred. An analogy may be made to the goal of tight control of blood glucose in the management of diabetes. The optimal goal is eugly-

cemia every day. Nonetheless, euglycemia 80% of the time is better than no euglycemia or euglycemia 20% of the time. In alcoholism, as in diabetes, a substantial partial response is better than no response.

Another consideration for treatment researchers is that abstinence is a categorical (yes or no) variable and not a continuous variable. From a statistical perspective, continuous variables have an advantage over categorical variables because the parametric statistical tests used to test continuous variables for level of significance have more statistical power than the nonparametric tests used for categorical variables.

Definitions of Relapse Other Than Abstinence

Although lack of abstinence as a definition of relapse has its limitations, there is no consensus among alcoholism treatment researchers concerning other measures of relapse to drinking. The most commonly used variable after lack of abstinence is frequency of drinking. This is usually expressed as percentage of days drinking or percentage of days abstinent. The limitation to this variable is that a high percentage is not necessarily worse than a low percentage. For example, a person may drink one drink every day and therefore have 100% drinking days. Another person may drink only on Saturdays but drink 10 drinks on those days. This person has 14% drinking days, but that lower percentage masks the fact that this individual's drinking may be more harmful.

Because of this limitation, measures of quantity, e.g., amount consumed during a specified time interval, also have been used. Similar to frequency, quantity does not separate binges from steady drinking. This has led some investigators to regard intensity of drinking as a more important variable. An example of intensity is average number of drinks per drinking day. Measures of combined frequency and intensity such as percentage of "heavy" drinking days have been used; however, no definition of heavy drinking has been used consistently. From a public health perspective, three or more drinks per day may be heavy in the sense of being associated with harmful consequences. Levels higher than three drinks per day often are used as criteria for heavy drinking. The levels appear to differ by country and may reflect cultural norms more than epidemiological data.

Recent clinical trials of a pharmacotherapy have used a definition of relapse that includes a combination of frequency and/or intensity.^{4,5} The investigators defined relapse as follows: (1) drinking on 5 or more days within 1 week, (2) five or more drinks per drinking occasion, or (3) blood alcohol concentration of >100 mg/dL. Clinical trials found that naltrexone did not significantly reduce the number of "slips" but did significantly reduce the incidence of relapse.

Because each variable has advantages and limitations and none captures the entire phenomenon of problem drinking, several variables often are used. However, statistical adjustment must be made for using more than one index of drinking because, if more than outcome variable is used, there is a greater probability of detecting by chance a statistically significant difference.⁶ The reason for this is that the most frequently used statistical tests (*t*-test, chi-squared) were designed to test differences between two groups on one outcome variable. If multiple variables are compared, the *P* value (level of significance) is no longer .05 but larger, and the value balloons as more comparisons are made. A common correction for this problem of multiple comparisons is the Bonferroni correction or inequality,⁶ which "protects" against the type I error. The formula for the Bonferroni inequality is $\alpha^1 = 1 - (1 - \alpha)^{1/k}$, where α is the usual level of significance, .05; α^1 is the "true" level of significance (*P* value); and *k* is the number of comparisons to be made. If 20 comparisons are to be made, the true *P* value is .0026 rather than .05. Therefore, a result is not considered significant unless the *P* value is $\leq .0026$.

To select two or three variables that capture treatment outcome without affecting the level of significance too greatly, alcoholism treatment researchers have tried to prioritize indices of drinking. Investigators designing a multisite patient-treatment interaction study wrestled with this issue of choosing a few important treatment outcome variables. Their review of the literature⁷ found that abstinence; frequency; quantity; intensity; combined frequency and intensity; time to event measures, e.g., time to first drink; negative consequences; relapse; biological indices; cost indicators; and outcome categories created by combining alcohol consumption patterns and negative consequences had been used as measures of treatment outcome.

Because they were testing 10 *a priori* hypoth-

eses, these investigators were constrained for statistical reasons to select two primary outcome variables. They reasoned that if a few variables are to be selected, less correlated ones capture different aspects of drinking behavior and therefore are more informative than highly correlated variables. They compared correlations among continuous measures of frequency, volume, and intensity of drinking using two different data sets that contained these variables. Based on their analysis, they selected two primary outcome variables: proportion of days abstinent (a measure of drinking frequency) and number of drinks per drinking day (a measure of drinking intensity).

Methods of Detecting Abstinence

Treatment researchers use the patient's report (self-report), reports from others (family members, friends, employers), and results of breath alcohol tests to confirm abstinence or detect the lack of abstinence. All of these methods have limitations. The patient may claim abstinence when he or she is not. Family members and others are not with the patient 24 hours a day and may be unaware of the patient's drinking. Friends and associates who are heavy drinkers may not be reliable witnesses, and some alcoholics are estranged from their families. Breath alcohol tests only detect drinking that has occurred recently (less than 12 hours). Because of the limitations of any one method, it has become standard to use all three.

Agreement between patients and relatives about sustained abstinence is often quite good. In a study of more than 1,330 patients and collaterals, there was 84.7% agreement about the presence or absence of abstinence between patients and collaterals at the 1-year posttreatment interview (T. Babor, University of Connecticut, personal communication, September 1966). There was a false-positive rate of 3.7%, i.e., the patient reported abstinence and the collateral did not. Conversely, there was a false-negative rate of 11.6% in which collaterals reported that the patient did not drink when the patient reported drinking. In another study, 6.7% of patients' reports of abstinence were not corroborated by the collaterals.⁸ The authors of this study used Baye's theorem to determine conditional probabilities of abstinence. They found that the probability that a patient is abstinent based on his or her claim of abstinence is 85% assuming that the collateral's report is the true measure of abstinence.

These investigators also found that positive urine alcohol test results repudiated another 23% of patients' reports of abstinence. Combining collateral reports and laboratory test results, these authors found that the probability of a patient's claim of abstinence being true is 65% assuming the combination is the true measure of abstinence.

Liver function tests have also been used to corroborate patients' reports of abstinence. Irwin et al⁹ studied alcoholics 3 months after discharge from an inpatient program. They compared the results of posttreatment liver tests with baseline values and concluded that parallel increases in the values of γ -glutamyl transferase (GGT) of $\geq 20\%$, alanine aminotransferase (ALT) of $\geq 40\%$, and aspartate aminotransferase (AST) of $\geq 20\%$ have a sensitivity of 100% and a specificity of 82% in differentiating alcoholics who are abstinent from those who are not.

In a clinical trial of disulfiram, liver function tests were performed every 2 months for 1 year to monitor for disulfiram-induced hepatotoxicity. Increases in liver enzyme levels almost always indicated resumption of drinking rather than disulfiram hepatotoxicity.¹⁰

Tests for detecting drinking for periods of time are under development or being evaluated. One promising marker is carbohydrate-deficient (desialylated) transferrin (CDT). A recent study¹¹ measured CDT and GGT levels in 35 male alcoholics before they entered treatment and every 4 weeks during 12 weeks of outpatient treatment. Approximately half of the men had elevated CDT ($n = 17$) and/or GGT ($n = 18$) levels. The percent of change from baseline was calculated. In the aggregate, CDT levels decreased approximately 30% in abstinent patients ($n = 14$) but increased approximately 10% in relapsed patients ($n = 10$) by week 12. Eleven of the 35 men were defined as neither abstinent nor relapsed but as having had a "slip," defined as five or more drinks on 2 consecutive days at some time during the 12 weeks. The change from baseline levels in this group was intermediate between those of the abstinent group and of the relapsed group. On the average, the GGT level was approximately 45% below baseline by week 12 in the abstinent group, whereas in the other two groups the GGT level was approximately 25% below baseline levels. It must be noted that the only significant difference was the difference in CDT levels between the abstinent men and the relapsed men. Although not explicitly stated, the

other differences probably were not significant because of the variation in CDT and GGT values and the small sample sizes. This study suggests that measuring changes in CDT and GGT levels from baseline is useful for providing additional verification for reports of abstinence in groups of subjects enrolled in treatment trials. However, this study did not report false-positive or false-negative rates. Therefore, the applicability to individual patients is uncertain. It is not clear at this time what the utility of CDT or GGT level would be as a marker for verifying abstinence in patients with alcoholic liver disease who have undergone liver transplantation.

Methods of Detecting Frequency, Quantity, and Intensity of Drinking

Self-reports and collaterals' reports are used to measure frequency, quantity, and intensity of drinking. Although the correlation between self-reports and collaterals' reports of drinking frequency is good, collaterals' reports do not account for all the variance in self-reports. One study found that collaterals reported twice as many drinking days as patients reported during a 1-year period.⁸

A problem with collaterals' estimates of drinking quantity is that data often are missing primarily because collaterals are unable to estimate volume of alcohol consumed because they are not with the patient during the entire day. To improve the validity of self-reports, researchers are evaluating a computerized touch-tone telephone system. The individual calls a toll-free number daily and reports if he or she has drunk that day and how much.

Liver function test results have been used to corroborate self-reports of drinking frequency and quantity. Keso and Salaspuro¹² studied 60 male and 13 female employed alcoholics at 2-month intervals for 8 months after discharge from an inpatient program. They reported that substantial reductions in number of drinking days and average daily alcohol intake after treatment were accompanied by substantial decreases from baseline levels of GGT, AST, and ALT. Of the three tests, they concluded that the GGT level was "clearly superior" to the other two. Although changes in the liver enzyme results paralleled changes in alcohol consumption, those results were for the patients as a group. The data do not indicate whether these results can be applied equally well to individual patients.

Breath, blood, urine, and saliva alcohol tests are

good for determining drinking or abstinence at the time the test is obtained. However, because of the relatively rapid metabolism and excretion of alcohol, they are not useful for verifying outcome variables other than abstinence or intensity at a specific point.

Most alcoholism treatment studies have measured drinking behavior and not the consequences of drinking, although measuring the negative consequences of drinking is becoming more common. The assumption is that cessation of drinking or a substantial reduction in alcohol consumption eventually results in fewer adverse consequences, and consequences may not occur frequently in the short term to detect significant differences between treatment groups on these variables. This is analogous to smoking cessation programs that measure reduction in smoking and not lung cancer. When consequences are measured, archival data, e.g., medical records, employment records, motor vehicle records, are often used.

Recommendation

What should be considered a "relapse" after transplantation for alcoholic liver disease? If we knew the level of alcohol consumption below which hepatotoxicity did not occur, we could define relapse as any drinking above that level. An American Cancer Society prospective study¹³ of 276,802 US men aged 40 to 59 years who were followed up for 12 years found that compared with nondrinkers, the relative risk for liver cirrhosis was 3.15 (95% confidence interval, 2.39 to 4.16) for men who had two drinks per day and progressively increased to 18.1 for those who drank 6 or more drinks per day. This suggests that the threshold may be as little as two drinks per day. Although two drinks per day may be the threshold, these data do not indicate how long or short the duration of this level of consumption must be before the risk of cirrhosis is increased. Two drinks per day may increase the risk of dying of cirrhosis, but the situation for mortality from coronary artery disease is not the same. For coronary artery disease, abstinence increases the risk of dying compared with one to two drinks per day.

Although abstinence is the goal, treatment for alcoholism can result in considerable improvement even when complete abstinence is not achieved. This is illustrated by the results of a recent study of more than 700 patients treated with a minimum of

10 days of inpatient or day hospital treatment followed by three months of "aftercare" treatment.¹⁴ The sustained abstinence rate by the end of 12 months after discharge was 35%. The sustained abstinence rate represents the ideal outcome but underestimates the rate of favorable outcomes. These patients, in terms of drinking frequency, were abstinent more than 90% of the days during the 12 months of follow-up, and 60% never had three or more consecutive days of heavy drinking (defined as six or more drinks per day for men and four or more for women). Abstinence must be the goal for an individual with alcoholic liver disease after transplantation because it is impossible to predict *a priori* who will slip and who will relapse when an alcoholic begins drinking. Although a slip is an unwelcome event, if it occurs and does not progress to a relapse it is unlikely to cause harm and should not be treated punitively.

To determine the presence or absence of abstinence, I recommend interviewing the patient and a relative and performing a breath alcohol test. Administration of a breath alcohol test may imply a lack of trust on the physician's part. However, it is important to know whether the patient is abstinent because if he or she is not, this fact must be part of the therapeutic dialogue. A slip often provides an opportunity for the patient to reexamine his or her life and renew the commitment to abstinence.

References

1. American Psychiatric Association. Diagnostic and statistical manual for mental disorders (ed 4). Washington, DC, American Psychiatric Association, 1994.
2. Walsh DC, Hingson RW, Merrigan DM, Levenson SM, Cupples LA, Heeren T. A randomized trial of treatment options for alcohol-abusing workers. *N Engl J Med* 1991;325:775-782.
3. Vaillant GE. A long-term follow-up of male alcohol abuse. *Arch Gen Psychiatry* 1996;53:243-249.
4. Volpicelli JR, Alterman AI, Hayashida M, O'Brien CP. Naltrexone in the treatment of alcohol dependence. *Arch Gen Psychiatry* 1992;49:876-880.
5. O'Malley SS, Jaffe AJ, Chang G, Shottenfeld RS, Meyer RE, Rounsaville B. Naltrexone and coping skills therapy for alcohol dependence: A controlled study. *Arch Gen Psychiatry* 1992;49:881-887.
6. Lachin JM. Statistical analysis of the randomized clinical trial. In: Tygstrup N, Lachin JM, Juhl E (eds). *The randomized clinical trial and therapeutic decisions*. New York: Marcel Dekker, 1982:183-184.
7. Babor TF, Longabaugh R, Zweben A, Fuller RK, Stout RL, Anton RF, et al. Issues in the definition and measurement of drinking outcomes in alcoholism treatment research. *J Stud Alcohol* 1994 (Suppl. 12):101-111.
8. Fuller RK, Lee KK, Gordis E. Validity of self-report in alcoholism research: Results of a Veterans Administration cooperative study. *Alcohol Clin Exp Res* 1988;12:201-205.
9. Irwin M, Baird S, Smith TL, Schuckit M. Use of laboratory tests to monitor heavy drinking by alcoholic men discharged from a treatment program. *Am J Psychiatry* 1988;145:595-599.
10. Iber FL, Lee KK, Lacoursiere RB, Fuller R. Liver toxicity encountered in the Veterans Administration trial of disulfiram in alcoholics. *Alcohol Clin Exp Res* 1987;11:301-304.
11. Anton RF, Moak DH, Latham P. Carbohydrate deficient transferrin as an indicator of drinking status during a treatment outcome study. *Alcohol Clin Exp Res* 1996;20:841-846.
12. Keso L, Salaspuro M. Laboratory markers as compared to drinking measures before and after inpatient treatment for alcoholism. *Alcohol Clin Exp Res* 1989;13:449-552.
13. Boffetta P, Garfinkel L. Alcohol drinking and mortality among men enrolled in an American Cancer Society prospective study. *Epidemiology* 1990;1:342-348.
14. Project MATCH Research Group. Matching alcoholism treatments to client heterogeneity: Project Match post-treatment drinking outcomes. *J Stud Alcohol* 1997;58:7-29.