Published in final edited form as:
Alcohol. 2005 June ; 36(2): 19–26.

Treated and Treatment-Naïve Alcoholics Come From Different Populations

George Fein^{*} and Bennett Landman

Neurobehavioral Research, Inc., 201 Tamal Vista Boulevard, Corte Madera, CA 94925, USA

Abstract

Most research on alcoholism uses convenience samples of individuals who have been in some type of treatment; Berkson's fallacy results when the associations found in studies of select samples are incorrectly presumed to apply to *all* alcoholics (i.e. including untreated alcoholics in the general population). This study examines whether treated and untreated alcoholics have similar early alcohol use histories by comparing abstinent alcoholics (treated and sober at least 6 months) to treatment-naïve alcoholics (active drinkers). We studied fourteen pairs of women and twenty-five pairs of men matched on the age at which they first met criteria for heavy alcohol use (80 drinks/month for women and 100 drinks/month for men). The timeline follow-back interview methodology was used to gather retrospective alcohol use information. Alcohol dose and duration was then computed for two intervals: the time between the person's first drink and the date at which they met criteria for heavy drinking, and the period between meeting criteria for heavy drinking and the current age of the treatment-naïve person from each pair. During the period prior to meeting the matching 'heavy drinking' criteria, alcohol dose did not differ between groups. In the period after meeting criteria for heavy alcohol use, the treated alcoholics had higher average and peak alcohol doses than the treatment naïve alcoholics. We rejected the hypothesis that the treatment naïve alcoholics and the treated alcoholics have similar alcohol use trajectories over time, with the treatment naïve sample simply being observed earlier in their alcohol use histories. Instead we concluded that the two groups come from different populations with regard to alcohol use (in fact, the treated alcoholics had alcohol doses over 50% higher than treatment-naïve alcoholics in the years just after they began drinking heavily). This suggests that results from studies of alcoholics in treatment or post-treatment (i.e., most studies of alcoholics) cannot be generalized to untreated individuals (who comprise the majority of alcoholics).

Keywords

Berkson's fallacy; Alcoholism; Treatment; Study design

1. Introduction

1.1. The potential bias in studying clinical samples – Berkson's fallacy

The first clues to the association between diseases and both their antecedents and their consequences often derive from the study of select samples of treated individuals, hospitalized patients, or autopsy cases. This is the case regarding our knowledge of the antecedents of alcoholism and of the effects of alcoholism on brain structure and function. Because not all alcoholics from the general population are equally likely to be in these study samples, bias may result when findings in these samples are presumed to apply to the population at large. This type of bias, known as Berkson's fallacy (after the person who first studied it in detail)

^{*}Corresponding author: Tel: +415-927-7676; fax: +415-924-2903, Email address: george@nbresearch.com (G. Fein).

(Berkson, 1946, 1955), occurs whenever the association between the independent variable (e.g., the diagnosis of alcohol dependence) and the dependent variable (e.g., the antecedents of alcoholism, severity of alcohol use, or the consequences of alcohol dependence) differ between the population from which the sample derives (hospitalized alcoholics or alcoholics in treatment or shortly post-treatment) and (alcoholics in) the general population.

Fleiss (Fleiss, 1973) presents examples of this bias, and the mathematics underlying it. A classic example of Berkson's fallacy occurred when Pearl (Pearl, 1929) found a negative association between the presence of cancer and tuberculosis in autopsy cases. Tuberculosis was less frequent in autopsy cases with cancer than in autopsy cases without cancer. Pearl inferred (erroneously) that the same negative association should apply to live patients, and proposed treating terminal cancer patients with tuberculin to arrest their cancer. He failed to understand that extrapolating an association found in autopsy cases to live patients is a fundamental error, *unless* all deaths are equally likely to be autopsied.

Roberts et al (Roberts et al., 1978) was the first to publish a study empirically demonstrating Berkson's fallacy. Robert's sample was comprised of 257 individuals (of a random sample of 2,784 individuals interviewed in the community) hospitalized during the prior six months. There was a very large positive association between the presence of respiratory disease and the presence of locomotor disease in the hospitalized individuals. However, Roberts correctly ascertained that respiratory and locomotor diseases were essentially independent in the entire random sample. The spurious association between respiratory disease and locomotor disease arose because the hospitalization rate of people with both diseases (29%) was about three times the rate of people with only respiratory or locomotor disease or neither disease (7–10%). As Fleiss (Fleiss, 1973) succinctly stated: "... Unless something is known about differential hospitalization rates..., a good amount of skepticism should be applied to any generalization from associations found for hospitalized patients ... to associations for people at large." Parnas and Teasdale (Parnas and Teasdale, 1987) presented an example of Berkson's fallacy in schizophrenia research with direct applicability to alcoholism research. An American-Danish prospective study (Parnas and Teasdale, 1987) of children of schizophrenic mothers compared psychiatrically hospitalized and untreated cases of schizophrenia spectrum disorders on a number of characteristics. Hospitalized and untreated cases were similar on a number of measures; however, hospitalized individuals exhibited higher levels of substance abuse, affective symptoms, and psychopathic tendencies. The authors suggest that "the clinical population may not be representative of the diagnostic category in question owing to [a greater] co-existence of confounding symptomatology (Berkson's fallacy)".

Drawing convenience samples from treated samples could create another instance of Berkson's fallacy in the study of alcohol dependence. Coexisting pathology (e.g., depression or bipolar affective disorder, antisocial personality disorder, attention deficit hyperactivity disorder, post-traumatic stress disorder, and other substance abuse disorders) may be greater in the treated samples than in alcoholics in the general population. However, this coexisting pathology may not be severe enough to result in clinical diagnoses that would exclude subjects from "alcoholism" research samples. Alternatively, the bias due to Berkson's fallacy may result if the severity of alcoholism is greater in clinical versus general population samples. Once again, (in either case) findings in clinical samples may not generalize to alcoholics in the general population.

1.2. How big is this the potential bias in alcoholism research?

The magnitude of the potential bias consequent to Berkson's fallacy depends on the proportion of alcoholics who are in the treated sub-population. The most current data available indicates that the number of alcoholics in treatment is a small proportion of alcoholics in the general population. The 1992 National Longitudinal Alcohol Epidemiologic Survey (Grant, 1994)

estimates that over 27 million Americans exhibit alcohol abuse or alcohol dependence. At about the same time, Harwood (Harwood et al., 1994) estimated that there were approximately 1.8 million Americans receiving treatment for alcohol problems in non-Federal hospital and community based treatment settings. Grant (Grant, 1994) estimates that only one in 10 individuals who need treatment for alcohol abuse problems have actually sought treatment. These estimates derive from different methodologies and sampling plans; however, even assuming that *three times* the 1.8 million individuals from the Harwood study received some form of treatment for alcoholism, the treatment population is still less than a quarter of the of the number of people who exhibit alcohol problems. Therefore, it makes sense that estimates drawn from clinical samples only represent at most one-quarter of individuals who exhibit alcohol abuse or dependence. We need studies comparing alcoholics in treatment with alcoholics who have not sought treatment to determine whether clinical samples differ from treatment-naïve samples in alcoholism's antecedents, alcoholism severity, and alcoholism's 'consequences' (in the psychological, social, legal and biological arenas).

1.3. The bias resulting from Berkson's fallacy may differ between male and female alcoholics

The bias inherent in studying only treated alcoholics may be different for men and women, and this difference in bias may underlie the different results reported for male versus female clinical samples. The literature showing that women suffer more cerebral consequences from long-term alcohol dependence than men is based on studies of clinical samples (Bergman, 1987; Jacobson, 1986). However, this result may be spurious if clinical samples of treated alcoholic women differ from treatment-naïve alcoholic women more than treated male alcoholics differ from treatment-naïve samples of male alcoholics. This is entirely possible, because it has been claimed that women (for a variety of reasons) are less likely than men to receive treatment for alcohol problems (1996).

1.4. Examination of Berkson's fallacy in alcoholism research

In our laboratory, we have two ongoing studies comparing alcohol dependent samples to age comparable light/non-drinking samples. In one study we are examining alcoholics with six or more months of abstinence. All of those individuals (35 to 55 years old) went through treatment (we include Alcoholics Anonymous as one form of treatment). The other study examined alcohol dependent individuals between 20 and 50 years of age who had never been in treatment (in fact, none of those subjects identified themselves as alcoholic, although all met DSM-IV-R criteria for alcohol dependence). In this paper, we compare alcohol dependent subjects from the two studies as to their alcohol use history, examining both quantity and use trajectory. We are testing the null hypothesis that the treatment naïve alcoholics and the long-term abstinent alcoholics have similar alcohol use trajectories over time (the treatment naïve sample being observed earlier in their alcohol use histories). In addition, if their trajectories differ, we will test the hypothesis that that difference is larger in females than it is in males.

2. Materials and methods

2.1. Subjects

As noted above, subjects for this manuscript come from two different studies, one of abstinent alcoholics age 35–55, and the other of treatment naïve alcoholics age 20–50. In both studies, alcohol dependent and control samples were recruited; for the analyses reported here only the alcohol dependent samples were used. Abstinent alcoholic's needed to meet the lifetime criteria for alcohol dependence, and be abstinent for at least six months. Treatment naïve alcoholics needed to meet lifetime criteria for alcohol dependence and to have never been in treatment. All participants for either study were informed of the study's procedures and signed a written consent form prior to their participation. There were a total of four sessions for each study, with each lasting between an hour to two-and-a-half hours, involving clinical,

neuropsychological, electrophysiological and neuroimaging assessments. All subjects who participated in any session were paid for both their time and travel expenses. Subjects who completed all four aspects of either study received a completion bonus. An independent review committee (the Institutional Review Board, Independent Review Consulting, Corte Madera, CA) approved all procedures prior to study, and all procedures were carried out in compliance with the Helsinki declaration of 1975, as revised in 1983.

Exclusion criteria for both studies were: i) history or presence of an Axis I diagnosis on the DIS; ii) history of drug dependence other than nicotine; iii) significant history of head trauma or cranial surgery; i.v.) history of diabetes, stroke, or hypertension which required medical intervention, or of other significant neurological disease; v) clinical or laboratory evidence of active hepatic disease; vi) clinical evidence of Wernicke- Korsakoff syndrome; or vii) current substance abuse other than alcohol (aside from caffeine and nicotine). Table 1 shows the recruitment and inclusion/exclusion data for both studies. Table 2 presents subject demographics.

2.2. Assessment

All subjects were assessed using a computerized psychiatric Diagnostic Interview Schedule (DIS). Subjects were also interviewed on their drug and alcohol use using the lifetime drinking history methodology (Skinner and Sheu, 1982; Sobell and Sobell, 1990; Sobell et al., 1988), medical histories were reviewed, liver functions tested, and Family Drinking Questionnaires were administered based on the Family Tree Questionnaire by Mann et al. (Mann et al., 1985; Stoltenberg et al., 1998).

2.3. Subject matching and computation of dependent variables

The Lifetime Drinking History (LDH) assessment which uses the timeline follow-back interview methodology (where subjects break their drinking history into periods with consistent alcohol use) was used to gather retrospective alcohol use information (Sobell and Sobell, 1990, 1992; Sobell et al., 1988). We used data from the LDH to match abstinent alcoholic subjects and treatment naïve alcoholic subjects on a one-to-one basis. Subjects were matched on gender and age at the onset of heavy drinking. Heavy drinking was defined as the age when a subject first reached a monthly dose of 80 drinks per month for females and 100 drinks per month for males. There were 47 abstinent alcoholic subjects, but only 39 of these were able to be matched to a treatment naïve subject. The remaining subjects first met the criteria for heavy drinking relatively late in life (in their mid thirties), and we did not have any same gender treatment naïve subjects who matched them on that variable. The average difference in the age at which subjects met the criteria for heavy drinking was 1.85 months (s.d. = 15.4 months) across the 14 female and 25 male matched subject pairs. Once the matches were completed, the alcohol use variables from the LDH of the abstinent alcoholic subject from each pair were computed as if that subject were the age of his or her matched treatment naïve subject. For example, if an abstinent alcoholic subject was 55 years old, the matched treatment naïve subject was 30 years old, and both subjects met the heavy drinking criteria at age 23, then the alcohol use variables for the abstinent alcoholic subject were recomputed using the drinking history up to the point when that subject reached the age of 30 years. Alcohol dose, duration of use, and duration of abstinence variables were computed from the LDH for two intervals: the time between the person's first drink and the date at which they met criteria for heavy drinking, and the period between that date and the current age of the treatment-naïve person from each pair. This procedure is illustrated in Figure 1.

2.4. Analysis

The groups were compared on alcohol use variables using a repeated measures Analysis of Variance within the Statistical Analysis System (SAS, 2001). The trajectories of alcohol use

were also examined visually. For that purpose, because subject pairs first met the criteria for heavy use at very different ages (range 13 to 40 years), lifetime drinking histories were normalized with the age of meeting the criteria for heavy drinking set to zero. Each subject's use history was then plotted as a function of time. Subjects from the two groups were plotted using different colors to help visualize differences in drinking trajectories between the two groups.

3. Results

Table 3 presents the drinking history data divided into two intervals. The first interval is the period from an individual's first drink until the beginning of heavy drinking (as defined above). The second interval is the period from the beginning of heavy drinking to the age of the treatment naïve subject in each subject pair; thus the duration of this interval is the same for the two subjects in each abstinent alcoholic – treatment naïve alcoholic pair.

3.1. Interval from first drink to the beginning heavy drinking

In this interval, there were strong gender effects for all alcohol dose variables (average dose, peak dose and last six months dose), with males consistently having higher doses than females. All effects were large, with gender accounting for 14.9% of the variance of average dose, 22.1% of the variance for peak dose, and 16.9% of the variance for last six months dose. The only group difference was on abstinence duration during this period, with alcoholics who eventually were treated having some periods of abstinence in this interval before they had even begun to drink heavily. Treatment naïve alcoholics had no abstinence periods during this interval. Five of 39 alcoholics (4 males and 1 female) who eventually were treated had periods of abstinence in this interval.

3.2. Interval from beginning of heavy drinking to end of matched period

In this period, there were strong gender effects for average dose and peak dose, with gender accounting for 12.6% and 15.2% of the variance. There were also group effects that were larger than the gender effects, accounting for 27.8% of the variance of average dose, 20.8% of the variance of peak dose, and 21.8% of the variance of final six month dose. For all three variables, alcoholics who eventually were treated had much higher alcohol doses. There were no significant group by gender interaction effects. The effects for average alcohol dose for this interval are illustrated in Figure 2. There were also significant group effects for duration of alcohol use and duration of abstinence for this interval. Alcoholics who eventually were treated had more abstinence during this interval; because the entire interval was the same across groups, they had a smaller period of alcohol use during this interval as a consequence of their greater abstinence (Fig. 3).

Figure 4 presents the alcohol use trajectories of all research participants relative to the age at which the individual met the criteria for heavy drinking. The figure shows that the treated and treatment naïve trajectories almost entirely overlap in the period prior to meeting criteria for heavy drinking, but that the treated alcoholics have higher average doses in the period after meeting criteria for heavy drinking.

4. Discussion

The central finding of the current study is that treatment naïve alcohol dependent individuals in the community come from a population with much lower alcohol use than do treated alcoholics who have been successful in maintaining abstinence. In other words, we rejected the null hypothesis (that treatment naïve alcoholics have similar alcohol use trajectories to treated alcoholics, but are just identified earlier in their drinking histories). This hypothesis

was tested with matched pairs of subjects consisting of a treatment naïve and a treated alcoholic of the same gender, who both met criteria for heavy drinking at the same age. The drinking pattern for both subjects in the pair was then examined for an identical period of time (after meeting criteria for heavy drinking). During this period (on average about eight to nine years in duration), the average alcohol dose for the treated alcoholics was much higher than that for the treatment naïve alcoholics (56% higher for males and 68% higher for females).

The current study demonstrates Berkson's fallacy with regard to the association of a diagnosis of alcohol dependence with the magnitude of alcohol use. This association is markedly different in treated vs. treatment naïve samples (in the years immediately after meeting criteria for heavy drinking). We cannot generalize results from clinical samples of alcoholics (those in treatment or post-treatment - most studies of alcoholics) to untreated individuals (who comprise the majority of alcoholics) with regard to measures of the severity of alcohol use. This means that findings on any measures of the antecedents of alcohol dependence that may be predictive of differences in alcohol use (e.g., pre-existing comorbid psychopathology) or findings on measures of the consequences of alcohol dependence that may be affected by differences in alcohol use (alcohol use associated morbid changes in brain structure and function and exacerbation of comorbid psychopathology) also are not likely to extend from treated samples of alcoholics to alcoholics in the general population. The difference in alcohol dose (both average and peak dose) between treated and treatment naïve alcoholics was similar for men and women. The comparability of the differences between groups across genders implies that treated samples of alcoholic women differ from treatment-naïve alcoholic women comparably to treated versus treatment-naïve samples of alcoholic men. This argues against the contention that women are less likely than men to receive treatment for alcohol problems (1996), and suggests that the literature showing that women suffer more cerebral consequences from long-term alcohol dependence than men (Bergman, 1987; Jacobson, 1986) is true, and not a spurious finding due to sampling bias.

Another interesting result is that the treated alcoholics had more episodes of (brief) abstinence after starting to drink heavily than did treatment naïve alcoholics. In fact, they even had more episodes of abstinence before they met the criteria for heavy drinking. E. (It is important to note here that all persons in both groups met criteria for alcohol dependence.) Even though all of the individuals in both groups reported on interview that alcohol use had interfered with their lives, only people who eventually sought treatment and achieved abstinence identified this early in their drinking history that their drinking was problematic. . Attempts at abstinence this early in their drinking histories may not be characteristic of all alcoholics who go on to treatment. We must keep in mind that our treated sample consisted entirely of alcoholics who were eventually successful in achieving long-term abstinence. It is possible that early attempts at abstinence are more characteristic of the subset of alcoholics who are eventually successful at achieving long-term abstinence. However, it is surely a commentary about the difficulty of achieving long-term abstinence that these subjects continued drinking on average another 5.50 years for men and 10.24 years for women after the observation period of the current analysis.

There are limitations to the current study. First, the primary data for analysis is each subject's recall of the duration and dose of periods of prior alcohol use, and the results may reflect differences in recall rather than differences in actual prior use. The treated sample was older than the treatment naïve sample, and had long periods of abstinence before their data was collected. It is possible that they exaggerated their use in comparison to the recollections of the treatment naïve sample who were recalling relatively recent experience. We do not believe this is likely because the treated sample, in addition to recalling higher alcohol doses, also recalled more episodes of abstinence; therefore, if the data reflected primarily a recall error, that error would manifest both as estimates of less use (more abstinence) and of more use (higher doses).

It is hard to imagine a recall error that would result in both of these findings. The simplest hypothesis is that the recall data accurately reflect prior use.

A second limitation is that the study focuses on the subset of the treated population that eventually attains long-term abstinence. Observed differences in consumption patterns may be associated with the treated sample's ability to eventually achieve long-term abstinence. Different drinking histories may be present for treated samples that do not achieve long-term abstinence; however, it is also possible that drinking histories are comparable across treated samples.

A third limitation is that the current study does not actually measure antecedent factors (e.g., psychiatric and other comorbidities) and consequences of alcohol abuse (e.g., effects of chronic alcohol abuse on brain structure and function). The main finding of this research is that treated and treatment naïve alcoholics come from different populations with regard to alcohol use histories. This is a sobering finding for the field. It suggests that it is improper to generalize results and conclusions from convenience samples to alcoholics in the general population. This is demonstrated in the current study for alcoholism severity, but is also likely to be the case for antecedents and consequences of alcohol dependence (because such phenomena are associated with differences in alcohol dose). Our study underlines the importance of direct comparison between treated and treatment naïve alcohol dependent samples on measures of the antecedents and consequences of alcohol dependence. In order to elucidate the public health implications of research findings, we must understand how to extrapolate such findings to alcohol dependent individuals in the general population.

Acknowledgements

This work was supported by Grants AA11311 (GF) and AA13659 (GF), both from the National Institute of Alcoholism and Alcohol Abuse. This study would not have been possible without the dedicated recruitment team at NRI and our volunteer participants.

References

- Substance Abuse and the American Woman. Columbia University, Center on Addiction and Substance Abuse; New York: 1996.
- Bergman, H. Brain dysfunction related to alcoholism: Some results from the KARTAD project. In: Parsons, OA.; Butters, N.; Nathan, P., editors. *Neuropsychology of Alcoholism: Implications for Diagnosis and Treatment*. Guilford Press; New York: 1987. p. 21-45.
- Berkson J. Limitations of the application of fourfold table analysis to hospital data. *Biom Bull (now Biometrics)* 1946;2:47-53.
- Berkson J. The statistical study of association between smoking and lung cancer. *Proc Staff Meet Mayo Clin* 1955;30:319-348. [PubMed: 13245839]
- Fleiss, J. *Statistical Methods for Rates and Proportions*. Wiley; New York: 1973.
- Grant, B. National Institute on Drug Abuse Technical Review Meeting: Comorbid Mental and Addictive Disorders: Treatment and HIV-Related Issues. Rockville, MD: 1994. The influence of comorbid major depression and substance use disorders on alcohol and drug treatment: results of a national survey.
- Harwood, H.; Thomson, M.; Nesmith, T. *Healthcare Reform and Substance Abuse Treatment: The Cost of Financing Under Alternative Approaches*. Lewin-VHI; Fairfax, VA: 1994.
- Jacobson R. The contributions of sex and drinking history to CT scan changes in alcoholics. *Psychological Medicine* 1986;16:547-559. [PubMed: 3763773]
- Mann RE, Sobell LC, Sobell MB, Pavan D. Reliability of a family tree questionnaire for assessing family history of alcohol problems. *Drug Alcohol Depend* 1985;15:61-7. [PubMed: 4017879]
- Parnas J, Teasdale T. A matched-paired comparison of treated versus untreated schizophrenia spectrum cases. A high-risk population study. *Acta Psychiatrica Scandinavica* 1987;75:44-50. [PubMed: 3577840]
- Pearl R. Cancer and tuberculosis. *Am J Hyg (now Am J Epidemiol)* 1929;9:97-159.

- Roberts, RS.; Spitzer, WO.; Delmore, T.; Sackett, DL. J Chronic Dis. 31. 1978. An empirical demonstration of Berkson's bias; p. 119-128.
- SAS, I. I. The SAS/STAT System for Windows, Release 8.02. SAS Institute; Cary, NC, USA: 2001.
- Skinner HA, Sheu WJ. Reliability of alcohol use indices: The lifetime drinking history and the MAST. Journal of Studies on Alcohol 1982;43:1157-1170. [PubMed: 7182675]
- Sobell LC, Sobell MB. Self-reports issues in alcohol abuse: State of the art and future directions. Behavioral Assessment 1990;12:77-90.
- Sobell, LC.; Sobell, MB. Timeline Follow-back: A technique for assessing self-reported ethanol consumption. In: Allan, J.; Litten, RZ., editors. Measuring Alcohol Consumption: Psychosocial and Biological Methods. Mumana Press; New Jersey: 1992. p. 41-72.
- Sobell LC, Sobell MB, Riley DM, Schuller R, Pavan DS, Cancilla A, Klajner F, Leo GI. The reliability of alcohol abusers' self-reports of drinking and life events that occurred in the distant past. Journal of Studies on Alcohol 1988;49:225-232. [PubMed: 3374136]
- Stoltenberg SF, Mudd SA, Blow FC, Hill EM. Evaluating measures of family history of alcoholism: density versus dichotomy. Addiction 1998;93:1511-20. [PubMed: 9926555]

Methodology for Computing Alcohol Use Trajectories

Fig. 1.

This figure illustrates the matching procedure. The top section of the figure presents the raw data for a matched pair of subjects (a 46 year old male abstinent alcoholic and a 34 year old male treatment naïve alcoholic) who both met criteria for heavy drinking at about 18 years of age. The middle section shows the two steps of the matching and alignment procedure. In the first step, the two subjects are aligned at their age of meeting the criteria for heavy drinking (this amounts to a shift of 4 months for the abstinent alcoholic subject), and that age is set to zero. In the second step, the data for the abstinent alcoholic subject is truncated so that it is followed for the same duration after meeting criteria for heavy drinking as was the treatment naïve subject. The bottom section shows the data after alignment and truncation.

Fig. 2.

Average Dose For the Period From the First Heavy Use to End of Match Period. In the matched period, the average alcohol dose was significantly lower in the treatment naïve subjects than in the treated subjects who eventually attained long-term abstinence ($p < 0.01$). Females had lower doses than males ($p < 0.05$), but no group by gender interactions were observed.

Fig. 3. Alcohol Use Duration For the Period From the First Heavy Use to End of Match Period. In the matched period, the duration of alcohol use was lower for the abstinent alcoholics than for the treatment naïve alcoholics ($p < 0.01$) due to increased periods of abstinence in the treated subjects who eventually attained long-term abstinence. No gender or group by gender interaction were observed.

Fig. 4.

This figure displays each subject's alcohol use trajectory. When the trajectories of an abstinent alcoholic and a treatment naïve alcoholic overlap, it is indicated in red. The female trajectories are lower than male trajectories (note the difference in scale between the female and male sections of the figure). After meeting criteria for heavy drinking, the abstinent alcoholic trajectories (gray lines) are higher than the treatment naïve (orange lines) trajectories.

Table 1
Exclusions for the Abstinent Alcoholic and the Treatment Naïve Alcoholic Studies

Reason For Exclusion	Number of Abstinent Alcoholics	Number of Treatment Naïve Alcoholics
Drug use	26	123
Did not meet alcohol use inclusion criteria ^a	48	94
Met medical exclusion criteria	33	9
Met psychiatric exclusion criteria	32	7
Other ^b	21	30
Total Excluded	160	263
Total that met all inclusion criteria and completed study	47	70

^a i.e. for abstinent alcoholics, either did not meet requirements for length of sobriety or did not meet requirements for alcohol dependence or alcohol use; for treatment naïve alcoholics, were either formerly treated for alcohol abuse or did not meet requirements for alcohol dependence or alcohol use

^b e.g. did not want to commute, lost to follow-up, or refused continuation or participation

Table 2**Subject Demographics**

Variables	Abstinent Alcoholics		Treatment Naïve	
	Males (n=25)	Females (n=14)	Males (n=25)	Females (n=14)
Age mean ($\pm$ S.D)	45.6 ($\pm$ 7.0)	47.0 ($\pm$ 6.3)	32.2 ($\pm$ 6.9)	30.3 ($\pm$ 6.7)
YrsEduc mean ($\pm$ S.D)	15.5 ($\pm$ 2.0)	15.0 ($\pm$ 2.7)	16.3 ($\pm$ 1.7)	15.9 ($\pm$ 1.8)
Ethnicity ^a	1 AA, 22 C, 1 H, 1M	14 C	2 A, 18 C, 4 H, 1M	11 C, 3 M

^a A – Asian, AA – African American, C – Caucasian, H – Hispanic, M - Multiracial

Table 3**Alcohol Use Measures by Group and Gender**

Alcohol Use Measures	Females (n=14 pairs)		Males(n=25 pairs)		Gender	Effect Size (%)	
	AbstAlc mean (± S.D)	Tx-Naïve mean (± S.D)	AbstAlc mean (± S.D)	Tx-Naïve mean (± S.D)		Group	Group by Gender
<i>For the Period From the First Drink to First Heavy Use</i>							
Duration of Use (months)	55.3 (±38.0)	59.1 (±41.4)	75.1 (±57.1)	71.5 (±66.1)	2.41	0.00	0.69
Average Dose (drinks/month)	24.7 (±16.6)	25.4 (±17.5)	33.9 (±18.0)	41.2 (±22.0)	14.9 [*]	2.45	1.68
Duration of Peak Use (months)	23.6 (±11.6)	34.3 (±27.0)	29.3 (±19.6)	37.4 (±40.6)	0.97	6.27	0.11
Peak Dose (drinks/month)	34.8 (±25.5)	31.2 (±21.4)	53.8 (±27.6)	56.2 (±28.1)	22.1 ^{**}	0.03	0.78
Average Dose in the Most recent 6 months	33.6 (±25.7)	27.3 (±20.4)	50.8 (±28.9)	48.2 (±30.7)	16.9 ^{**}	1.45	0.25
Duration of Abstinence In Period (months)	13.7 (±24.8)	0.0 (±0.0)	3.6 (±18.0)	0.0 (±0.0)	5.49	13.9 [*]	4.73
<i>Level at First Heavy Use</i>							
Dose (drinks/month)	153 (±67.6)	112 (±37.8)	215 (±138)	149 (±42.2)	14.2 [*]	14.4 [*]	0.74
<i>For the Period From the First Heavy Use to End of Matched Period</i>							
Duration of Use (months)	97.2 (±58.8)	116 (±64.5)	104 (±63.4)	121 (±59.2)	0.27	23.6 ^{**}	0.04
Average Dose (drinks/month)	165 (±70.6)	98.1 (±30.1)	210 (±110)	134 (±56.8)	12.6 [*]	27.8 ^{**}	0.14
Duration of Peak Use (months)	65.0 (±50.5)	45.4 (±33.4)	39.0 (±29.4)	45.9 (±31.7)	5.05	1.75	7.54
Peak Dose (drinks/month)	190 (±72.0)	128 (±46.5)	297 (±199)	175 (±67.3)	15.2 [*]	20.8 ^{**}	2.25
Average Dose in the Most recent 6 months	75 (±87.4)	108 (±47.5)	225 (±170)	120 (±66.4)	3.62	21.8 ^{**}	1.08
Duration of Abstinence In Period (months)	17.0 (±45.5)	6.00 (±22.4)	21.6 (±32.2)	0.48 (±2.40)	0.01	21.2 ^{**}	2.11

Gender is a between subject effect. Group and group by gender are within subject effects. Measures are reported mean $\pm$ standard deviation. Effect is significance:

*
p $\leq$ 0.05

**
p $\leq$ 0.01