Alcoholics Anonymous Effectiveness: Faith Meets Science

Lee Ann Kaskutas, DrPH

ABSTRACT. Research on the effectiveness of Alcoholics Anonymous (AA) is controversial and subject to widely divergent interpretations. The goal of this article is to provide a focused review of the literature on AA effectiveness that will allow readers to judge the evidence effectiveness of AA for themselves. The review organizes the research on AA effectiveness according to six criterion required for establishing causation: (1) magnitude of effect; (2) dose response effect; (3) consistent effect; (4) temporally accurate effects; (5) specific effects; (6) plausibility. The evidence for criteria 1- 4 and 6 is strong: rates of abstinence are about twice as high among those who attend AA (criteria 1, magnitude); higher levels of attendance are related to higher rates of abstinence (criteria 2, dose-response); these relationships are found for different samples and follow-up periods (criteria 3, consistency); prior AA attendance is predictive of subsequent abstinence (criteria 4, temporal); and mechanisms of action predicted by theories of behavior change are present in AA (criteria 6, plausibility). However, rigorous experimental evidence establishing the specificity of an effect for AA or Twelve Step Facilitation/TSF (criteria 5) is mixed, with 2 trials finding a positive effect for AA, 1 trial finding a negative effect for AA, and 1 trial finding a null effect. Studies addressing specificity using statistical approaches have had two contradictory findings, and two that reported significant effects for AA after adjusting for potential confounders such as motivation to change.

KEYWORDS. Alcoholics Anonymous (AA), 12-step, self-help, mutual aid, outcomes

Lee Ann Kaskutas is affiliated with the School of Public Health, University of California–Berkeley, Berkeley, California.

Address correspondence to: Lee Ann Kaskutas, DrPH, Alcohol Research Group, 6475 Christie Avenue, Suite 400, Emeryville, CA 94608-1010 (E-mail: lkaskutas@arg.org).

INTRODUCTION

Research on the effectiveness of Alcoholics Anonymous (AA) is controversial and subject to widely divergent interpretations. For example, the Cochrane Group published a review of the AA literature that considered outcome studies of AA and of 12-step facilitation (TSF), a form of specialty treatment that introduces clients to the 12-step philosophy and support system. Their review recommended that people considering attending AA or a TSF treatment program should be made aware that there is a lack of experimental evidence about the effectiveness of such programs.[1] This is despite optimal outcomes for TSF at 1 and 3 years for outpatients in the Project MATCH trial.[2,3] At the other end of the spectrum, 12-step scholar Rudy Moos has recommended that referral agencies should consider referring people to AA first rather than to treatment first. This is based on his own observational studies, which have found that longer duration of AA attendance is associated with less drinking at 8 and 16 years,[4] and that those who attend AA

before attending treatment tend to attend AA longer than those who attend treatment first.[5] The goal of this article is to provide a focused review of the literature on AA effectiveness that will allow readers to judge the evidence for AA effectiveness themselves.

Prior efforts to summarize the findings on AA effectiveness have included literature reviews[6,7] and meta analyses.[8–10] The most recent meta-analysis[10] concluded that attending AA led to worse outcomes than no treatment at all. An earlier meta-analysis focusing on moderating effects found that the evidence for AA effectiveness was stronger in outpatient samples, and that poorer quality studies (based on volunteers, self-selection rather than random assignment, and no corroboration of self-report) somewhat inflated the case for AA effectiveness.[9] A review summarizing the state of the literature 7 years later[7] argued that there was a consistent, rigorous body of evidence supporting AA effectiveness. Again, there seems to be something for everybody and the literature seems to be widely subject to interpretation. This may stem from the criterion being used to judge effectiveness.

At the heart of the debate is the quality of the evidence. AA critics have argued that AA is a cult that relies on God as the mechanism of action,[11] and that rigorous experimental studies are necessary to convince them of AA's effectiveness. Their concern is well-founded. As will be evident from this review, experimental studies represent the weakest of the available evidence. However, the review also will highlight other categories of evidence that are overwhelmingly convincing with respect to AA effectiveness, including the consistency with established mechanisms of behavior change. This review will organize the research on AA effectiveness according to 6 formal criterion for establishing causation,[12] which should help readers to integrate the sometimes conflicting conclusions discussed above. These criterion were first introduced to assist policymakers in evaluating the totality of the evidence of a causal effect for smoking on lung cancer in the absence of experimental data (as randomizing individuals to smoker and non-smoker conditions was not an option).[13,14] The criterion offer a framework for judging the "totality" of the evidence,[12] implicitly acknowledging that the evidence may not be strong for all criteria, and leaving the final decision to the individual evaluator. These are the criterion:

1. The relationship between an exposure (here, exposure to AA) and the outcome (here, abstinence because AA does not recommend any drinking for alcoholics) must be strong. According to this criteria, weak relationships between AA and abstinence would not be as convincing of causality as strong ones nor would they be as clinically relevant.
2. There should be a dose–response relationship, such that more involvement in AA relates to higher levels of abstinence. Building on the first criterion, the size of the dose–response effect also is important.
3. The consistency of the association matters. If some studies find a strong relationship between the number of AA meetings attended and the rate of abstinence but many do not, this would call into question whether the dose–response relationship should be trusted, as evidence goes.
4. The timing of the purported influence must be correct. This means that the measurement of AA exposure must be prior to the period of abstinence that is being studied; otherwise, it could mean that abstinent people tend to go to AA rather than AA causing people to be abstinent. Concurrent relationships do not count here; thus, according to this criterion, AA attendance for the past month cannot be considered as causal evidence for being abstinent during the past month.
5. The specificity of the association must be demonstrated. One must be able to rule out other explanations than AA exposure for having led to abstinence. This addresses the concern that those who attend AA are a part of a select sample who would be sober without ever going to AA. For example, if those who attend AA are highly motivated to do something about their drinking, it could be that this motivation is the cause of their abstinence and it would be unfair to credit AA for their successful outcome. Evidence of specificity ideally requires experimental manipulation of exposure to AA. For example, individuals in

a study might be randomized to attend AA or to attend psychotherapy; they do not select their treatment. Because of randomization, motivated people would end up being randomized both to psychotherapy and to AA, so it would not be the case that the "deck was stacked" in favor of AA. If those randomized to attend AA were more likely than those randomized to psychotherapy to be abstinent 2 years later, this would demonstrate an effect specific to AA that could not be due to a selection bias in which only motivated people attend AA. Randomization would also equalize other pre-existing conditions (known and unknown) that might confound AA's effect.

6. Coherence with existing knowledge is needed to establish causation. In drug trials, this is addressed by considering biological plausibility. For example, the drug neurontin stops seizures because it reduces the electrical activity in the brain. Here, in studying AA effectiveness, biological plausibility is of no help. The notion of theoretical plausibility is suggested as a way of addressing coherence with existing knowledge; that is, are the mechanisms of action that explain behavior change present in AA? Several theories and different aspects of AA exposure will be considered in addressing this final criterion.

METHODS

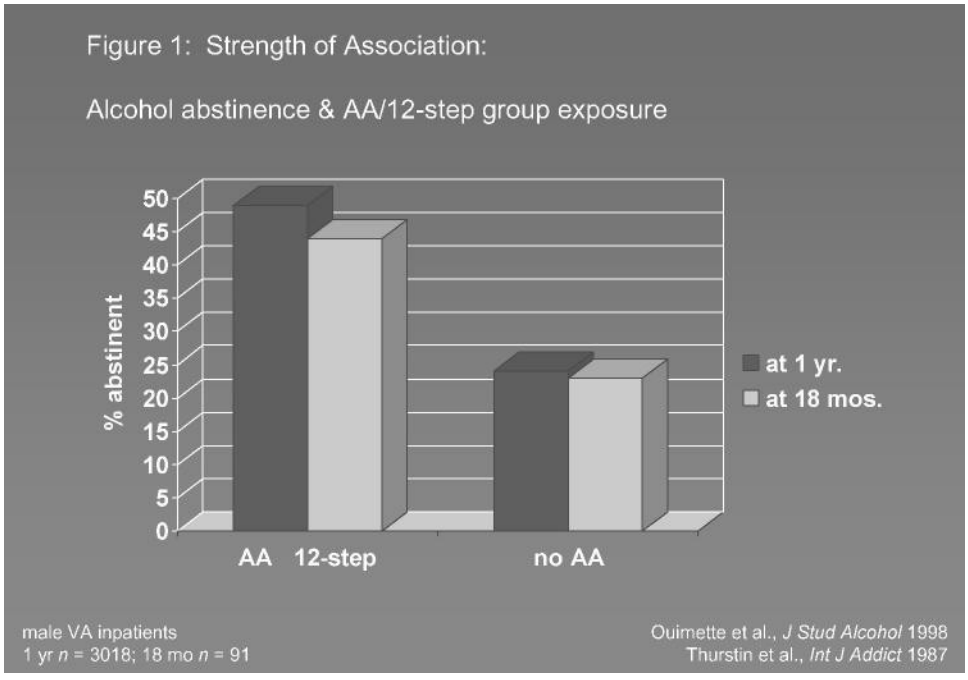
Articles involving Alcoholics Anonymous, Narcotics Anonymous, Cocaine Anonymous, 12-step group, and 12-step facilitation in the title or as a keyword were considered for this review. Electronic searches involved all relevant databases (e.g., Etoh and MedLine) and were augmented by the author's paper files on AA. Based on the title and in some cases the abstract, articles were considered for inclusion and were then read and classified. Representative studies were selected and are presented for each criterion. All located studies reporting a negative role for AA in abstinence are reported, and no studies with negative findings have intentionally been excluded. In the interest of brevity and clarity, many studies with positive findings for AA and several small 12-step facilitation studies with mixed results among subgroups have been excluded. The objective was not to provide another exhaustive literature review on AA effectiveness, but rather to present representative studies of AA effectiveness according to the criterion for establishing causation.

Results are shown using figures, with the percentage abstinent from alcohol along the y-axis and the AA exposure along the x-axis. Some studies combined alcohol and drug abstinence or considered 12-step group attendance, which would have included Narcotics Anonymous and other 12-step groups for drugs (in addition to AA). This is reflected in the figure titles and in the text. Results from studies that did not report rates of abstinence are not shown. The study samples and citations are summarized at the bottom of each figure.

RESULTS

Criterion 1: Strength of Association

How large is the relationship between AA exposure and abstinence? As shown in Figure 1, which draws on a longitudinal study of male inpatients in Veterans Administration programs, rates of abstinence are approximately twice as high for those who attended a 12-step group such as AA following treatment. One-year follow-up results considered 12-step group attendance and abstinence from alcohol and drugs, whereas the 18-month results reported AA attendance and alcohol abstinence. Results are remarkably similar at 1 year and 18 months for these different exposure and abstinence measures. Approximately 20% to 25% of those who did not attend AA or another 12-step group (or receive any other form of aftercare after the inpatient stay) were abstinent from alcohol and drugs at 1 year[15] and from alcohol at 18 months (combined alcohol and drug abstinence were not reported at 18 months).[16] The rates of abstinence were about twice as high among those who had attended AA or another 12-step group (but no other form of aftercare). In terms of effect sizes, this translates to a robust medium-size effect



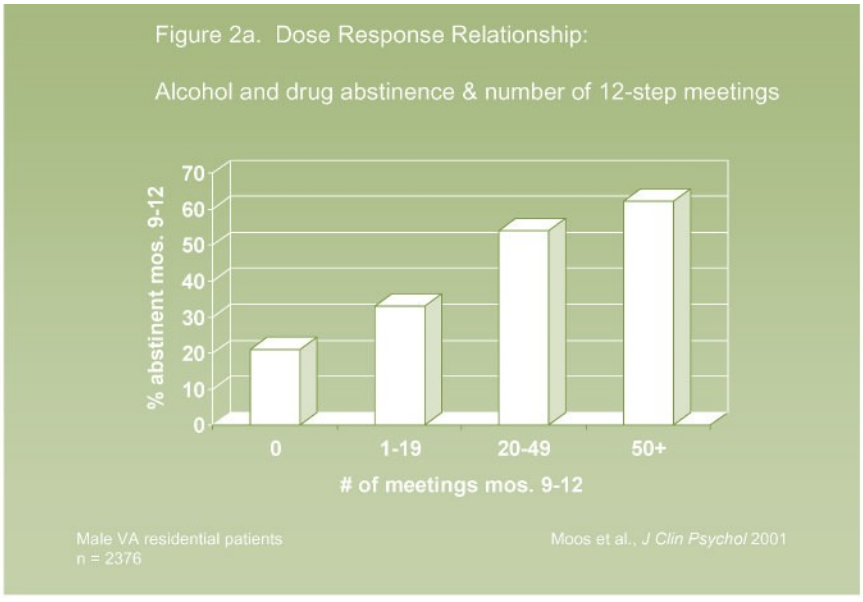
(h = .5).[17] Other studies are available that report on other substance use measures (such as percent days abstinent [PDA]) and samples. This study is selected to demonstrate the strength of the association because it comes from a large sample (n = 3,018 at 1 year); it reported simple dichotomous measures of AA or 12-step group exposure and abstinence; and it reported separately for those who attended AA or 12-step groups during follow-up but were not exposed to subsequent formal treatment.

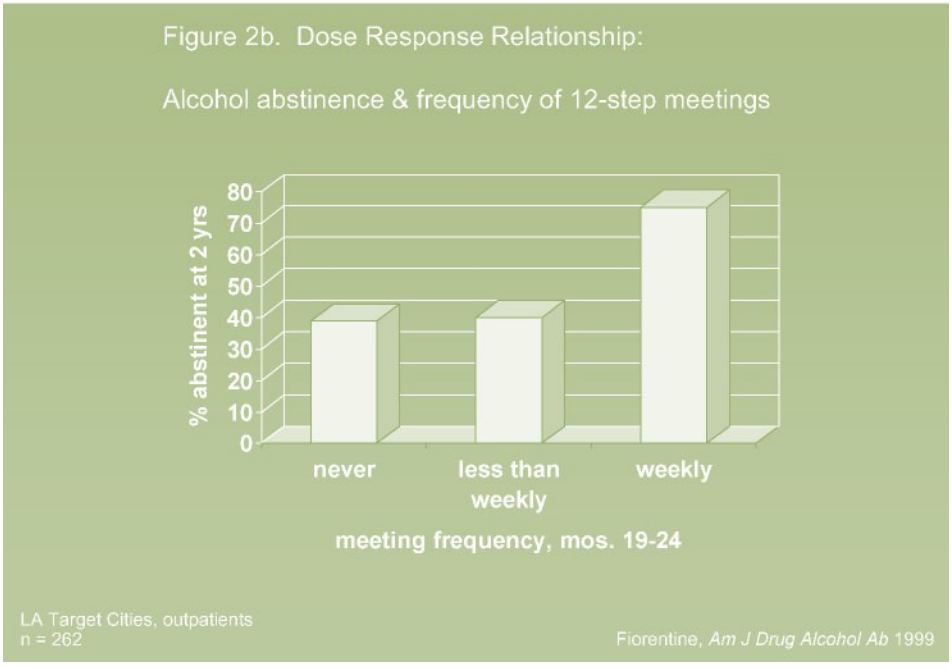
Criterion 2: Dose Response Relationship

Do higher levels of AA attendance or involvement relate to higher levels of abstinence? There is evidence of a dose response relationship for number of 12-step meetings (Figure 2a), frequency of 12-step meetings (Figure 2b), and duration of AA meeting attendance (Figure 2c). Again, studying male residential patients in the Veterans Association system and considering AA meeting attendance for the 90 days prior to the 1-year follow-up, the dose response curve looks almost linear (Figure 2a), with more 12-step meetings associated with higher rates of alcohol and drug abstinence.[4] In a smaller outpatient sample, more than 70% of those attending 12-step groups weekly for the 6 months prior to the 2-year follow-up were alcohol abstainers, whereas alcohol abstinence rates among those attending less than weekly were the same as those who never attended during that period[18]; this suggests a threshold dose-response effect for weekly attendance at 12-step groups (Figure 2b). In a longitudinal study of previously untreated problem drinkers, 70% of those with 27 weeks or more of sustained AA meeting attendance any given year (whether at year 1, years 2 to 3, or years 4 to 8) were abstinent from alcohol at the 16-year follow-up;[4] those with shorter duration of attendance had lower rates of abstinence, with the dose response most evident for AA attendance years 1 and years 4-8 (Figure 2c). This study is the reason for Moos' recommendation (see Introduction) to send people to AA first because those who went to AA first were more likely to be involved in AA for longer duration.[5]

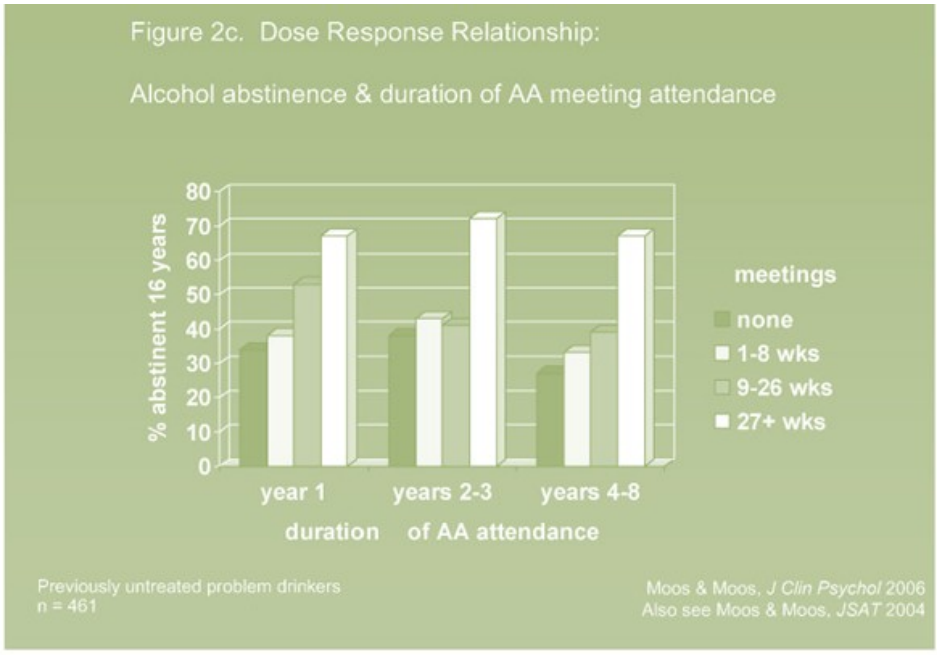
Criterion 3: Consistency of Association

The similarities in abstinence rates between the weekly or near-weekly AA attendees (70%) in these two latter studies with different





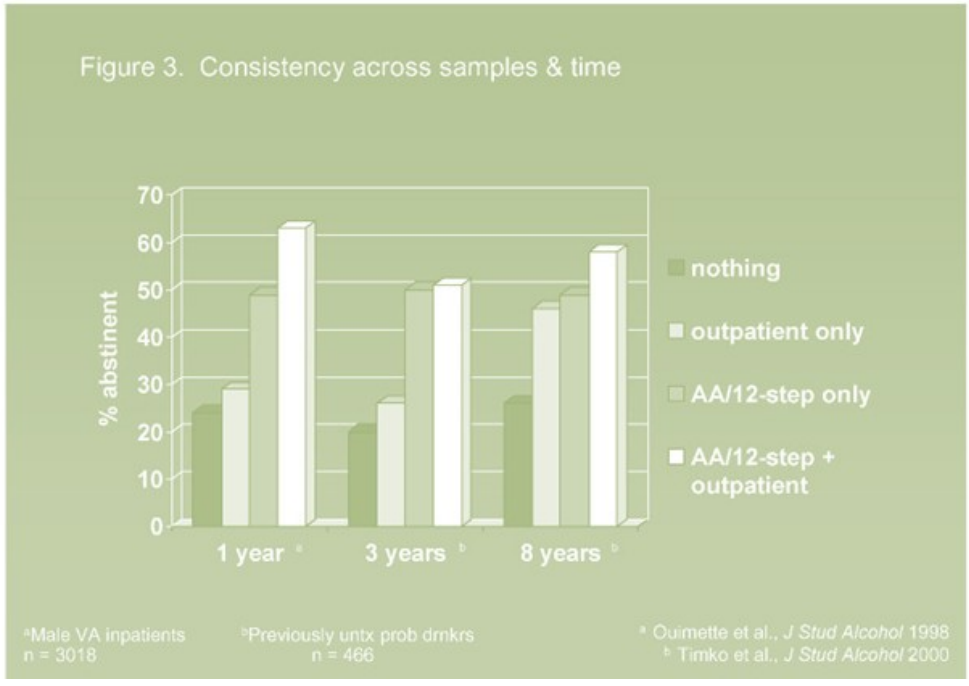
populations and follow-up periods is relevant to criterion 3. Another example is shown in Figure 3, which presents the rates of abstinence for those who attended AA but no other treatment (third bar, labeled "AA only") in two different samples (Veterans Association inpatients and previously untreated problem drinkers in the general population) with different follow-up periods (1, 3, and 8 years). The 1-year study considered alcohol and drug abstinence as a function of 12-step group attendance, whereas the 3-and 8-year data focused specifically on AA attendance and alcohol abstinence. Approximately 50% of those who had attended AA or 12-step meetings only were abstinent at 1 year[15] and at 3 and 8 years;[19] approximately one-fifth of those who did not attend AA or 12-step meetings or treatment were abstinent at the parallel



follow-up interviews. Another study of the general population[20] found that individuals with lifetime alcohol dependence who went to 12-step meetings but did not have formal treatment were more likely to be abstinent than those who did nothing (not shown).

Criterion 4: Temporally Correct Association

Most of the above studies considered concurrent AA attendance, and thus do not meet the 4th criterion for evidence of causality. An exception



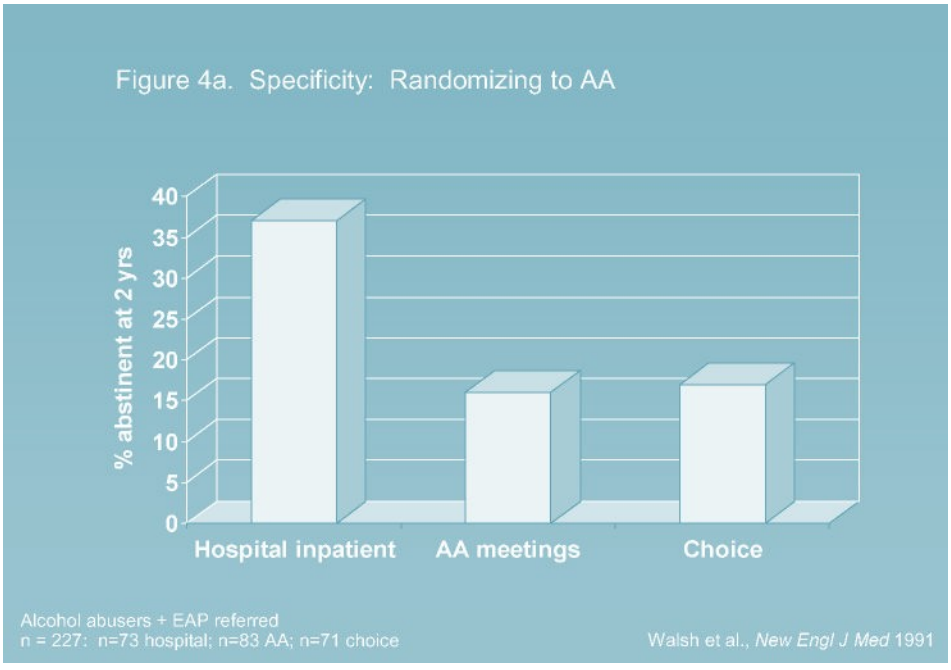
is Moos' work, which studied 16-year alcohol abstinence in a previously untreated problem drinking sample as a function of AA during years 2 to 3 and years 4 to 8 (Figure 2c).[4] Project MATCH also has evidence of a temporally correct association, reporting that frequency of AA meeting attendance as well as overall AA involvement in months 1 to 6 significantly predicted the percentage of days of alcohol abstinence during months 7 to 12. This was the case for Project MATCH subjects who attended inpatient treatment prior to entering the study ("aftercare" arm) as well as those who attended only the Project MATCH treatment ("outpatient" arm); the beta coefficients for AA involvement predicting abstinence were 0.34 in the aftercare arm and 0.29 in the outpatient arm (results not shown).[21,22]

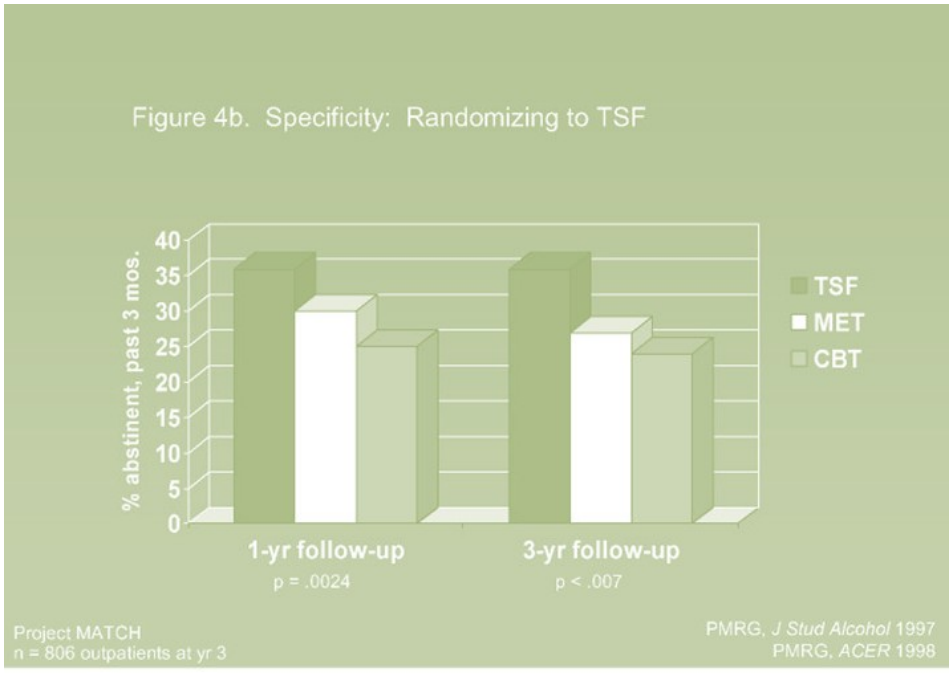
Criterion 5: Specificity

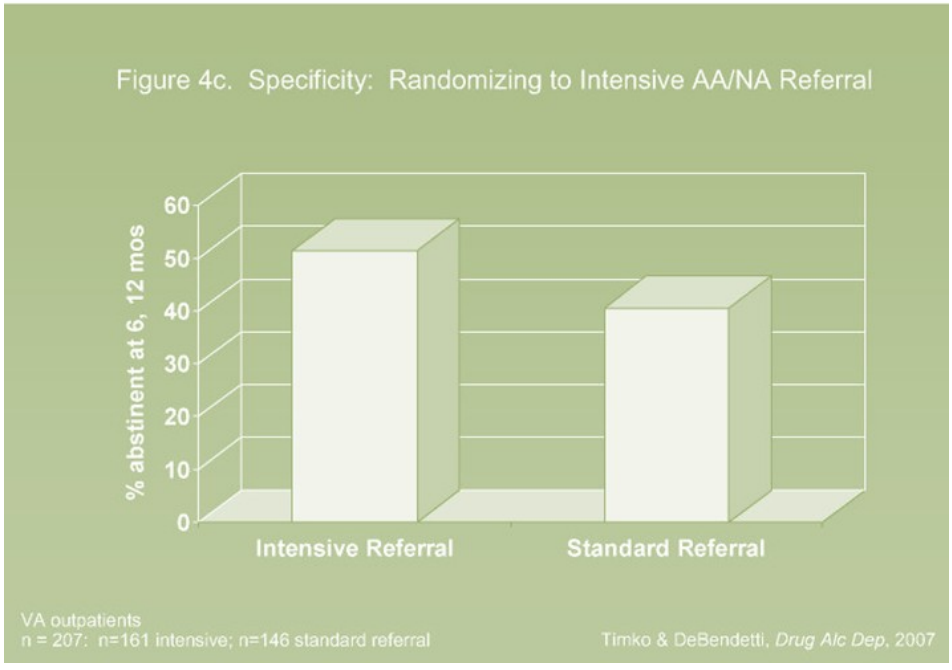
Experimental evidence is generally considered evidence of specificity. Three rigorous studies are particularly relevant here. The first, a clinical trial of compulsory treatment that randomized individuals to attend AA, attend hospital inpatient treatment, or choose their own treatment or service provider[23] found significantly lower rates of alcohol abstinence for the AA and the choice conditions, with over twice as many individuals abstinent at 2 years in the hospital inpatient condition (Figure 4a).

The second study, Project MATCH (discussed in criterion 4), randomized subjects to TSF, cognitive behavioral therapy, or motivational enhancement. In the aftercare arm, there were no significant differences between the three treatments, with more than two-fifths abstinent at the 1-year follow-up (results not shown). In the Project MATCH outpatient arm, rates of alcohol abstinence were significantly higher for those treated in TSF at 1 year[2] [Table 4] and 3 years[3] (Figure 4b). As noted above in Criterion 4, AA participation among Project MATCH clients predicted subsequent abstinence, regardless of study arm or condition.

The third trial randomized Veterans Association outpatients to an intensive 12-step referral condition or to standard AA referral,[24] finding significantly higher rates of total abstinence (from alcohol and drugs) at both the 6- and 12-month follow-ups for the intensive referral condition (Figure 4c). Higher AA or Narcotics Anonymous involvement in the intensive







referral condition fully mediated the condition effect on abstinence, but AA participation predicted abstinence regardless of condition.

Another relevant trial randomized individuals (mainly court-referred) to attend a weekly AA meeting run by the investigative team but not part of mainstream AA in the community, to attend weekly one-on-one therapy sessions led by lay individuals, or to a control condition in which subjects may have attended AA in the community, other available treatment, or no treatment.[25] Significantly more binge drinking at the 3-month follow-up was found for individuals randomized to the special AA meeting (2.37 binges in the

past 3 months) than to the other conditions (0.26 in lay therapy and 0.56 for the controls), but there was no reported difference in abstinence. However, at the 1-year follow-up, all drinking measures including rates of abstinence were similar across the conditions (result not shown). A 5th experiment randomized convicted drunk drivers to AA, outpatient treatment, or a no treatment condition; the study did not report drinking outcomes but found no differences in recidivism for drunk driving[26] (result not shown).

Criterion 6: Coherence with Existing Knowledge

To evaluate the literature on AA effectiveness according to this criterion (which usually is studied by considering biological plausibility), theoretical plausibility will be discussed; that is, does AA work in a way that is consistent with major theoretical perspectives on health behavior and behavior change? For example, a recent interpretation of contemporary psychodynamic theory has characterized alcoholism as an interaction between one's abilities to express feelings and self-regulate one's behavior.[27] The theory argues that despite low self-esteem, many alcoholics have a narcissistic personality[28] and a sense of omnipotence. They drink to self-medicate as a way of addressing unmet needs and uncomfortable psychological states. AA solutions consistent with this characterization of the problem are evident at meetings, in the AA steps, and through people in the AA fellowship. Meetings provide an opportunity to share one's own struggles, to learn how to talk about one's feelings, to increase one's motivation to abstain, and to get outside of one's self and change one's mood by hearing others talk about their problems and how AA helped them. The steps help with self-governance, narcissism, and omnipotence: accepting powerlessness over alcohol (step 1); recognizing that one cannot do it alone but that a higher power, which can be operationalized as the AA group, is there to help (steps 2-3); realizing how one's behavior affected and affects others (steps 4–9); treating other people better (step 10); finding meaning in life (step 11); and relinquishing one's negative self-focus by helping others (step 12). Through the people in AA, one learns how to live a sober life and how to regulate one's behavior one day at a time.

Bandura's social learning theory[29] adds to the psychodynamic perspective, saying that a large part of the problem arises from social influences and from self-efficacy: if everyone around you drinks and if you don't think it is within your ability to not drink, you will be unable to abstain. The antidote includes changing environmental cues (such as staying away from bars), role modeling (seeing others succeed at not drinking), and self-efficacy (believing you can abstain). AA meetings and spending time with people in AA represent changes in environmental cues (i.e., you're not at a bar seeing alcohol and watching people drink alcohol when you're at a meeting or out with AA friends). At an AA meeting, you are exposed to successful role models, instead of current drinkers, who suggest a new approach to abstinence: not drinking 1 day at a time (instead of saying you are "quitting forever"). Seeing yourself able to abstain for one day begins to build self-efficacy, which accumulates with the passage of every sober day. Spending time at AA meetings and with people in AA also leads to relapse prevention mechanisms put forward by standard behavioral modification techniques. These include learning how to say no to a drink when offered, having a plan of action when confronted with likely drinking conditions, and choosing alternative behaviors to take the place of drinking.

Several studies offer empirical support for these mechanisms. The positive relationship between AA involvement and abstinence has been shown to be partially mediated (explained) by (1) psychological and spiritual mechanisms including finding meaning in life,[30] greater motivation for abstinence,[31] and changes in religious beliefs and spiritual experiences;[32] (2) social influences such as fewer pro-drinking influences,[33] more friends in general,[34] having AA friends supportive of abstinence,[35] and enhanced friendship networks;[36] and (3) social learning and behavioral mechanisms including improved self-efficacy,[31,37] and effective coping and relapse prevention skills[34,36] to abstain. These mechanisms (and theories) are inter-related. For example, AA friends represent a particularly effective

source of social support because they provide expertise in preventing relapse.

DISCUSSION

Limitations

This is not a thorough review of the literature on AA effectiveness. For example, we did not keep track of the number of relevant studies located or the relative numbers of studies with positive versus negative findings for AA or TSF effectiveness. However, we did take care to present any study where the effect of AA was negative. The goal was not to provide an exhaustive review of the evidence, but rather to present representative studies that address AA effectiveness according to six accepted criterion for establishing scientific causation. This framework may be especially appropriate for considering AA effectiveness because it acknowledges the value and limitations of experimental evidence in the context of other criterion for determining treatment effectiveness.

Another limitation is the choice of theoretical frameworks for consideration. Biological theories were not considered here because their solutions are not behavioral but rather pharmacological: genetic theory (one is predisposed to develop alcoholism) and neurobiological theories (the brain becomes addicted to alcohol). For ideas about other behavioral theories that might be at work in AA, readers are referred to Moos' recent article on the active ingredients of substance use-focused self-help groups, which considers social control theory, behavioral economics, and stress and coping theory in addition to social learning theory.[38] The breadth of theoretical frameworks through which AA mechanisms can be understood is encouraging.

CONCLUSIONS

As stated at the outset, the experimental evidence for AA effectiveness (addressing specificity) is the weakest among the six criteria considered crucial for establishing causation. Only two studies provided strong proof of a specific AA or TSF effect: the outpatient arm of Project MATCH (with effects at 1 and 3 years)[2,3] and the intensive referral condition in Timko's trial (with effects for abstinence at 6 months and 1 year).[24] The effect sizes were similar, with the TSF/intensive referral conditions having a 5% to 10% advantage in abstinence rates. It is noteworthy that neither of these studies attempted to randomize patients to AA per se; instead, they focused on interventions intended to facilitate AA involvement.

One reason that several of the other trials may not have found positive effects for AA/TSF is because many individuals randomized to the non-AA/non-TSF conditions also attended AA; thus, the AA or TSF condition ended up being compared to a condition consisting of an alternative treatment plus AA. This was the case in Walsh's hospital inpatient treatment versus AA study[23] and in the aftercare arm of Project MATCH,[22] and arose because the patients in the non-AA/non-TSF conditions also had attended 12-step-based inpatient treatment, which in turn engendered strong participation in AA. Thus, AA attendance levels were high in the inpatient hospital condition in the former study and in the cognitive behavioral therapy and MET conditions among the Project MATCH aftercare subjects. In fact, cognitive behavioral therapy and MET aftercare patients attended more meetings than the TSF outpatients, and the aftercare patients overall attended twice the number of meetings at every follow-up compared to the outpatients.[22]

There are other concerns with the Brandsma trial,[25] which call its experimental results into question. The control condition allowed for participation in actual AA meetings, whereas those in the AA condition attended a weekly AA-like meeting administered by the study that was not an actual AA meeting. The description of the AA condition states that the steps were used for discussion content, the group focused on newcomers, and they told patients about sponsors,[25] but it is not clear whether the meetings were led by AA members, whether crosstalk was allowed, whether the meeting leader shared their story as part of the meeting, or whether the meeting format was what one would encounter at an actual AA meeting. The meetings may not have been

open to other AA members in the community and may not been listed in the AA meeting directory, which would mean that a potentially important therapeutic ingredient of AA—the experience of longer-term members—would not have been present in the AA condition. This is of special concern because the control condition did allow for attendance at such meetings.

Given these challenges in conducting rigorous randomized trials of AA effectiveness, researchers have turned to statistical methods to address the selection bias associated with AA attendance in observational studies. These efforts are intended to address criteria 5, specificity of the AA effect. The goal with these methods is to statistically adjust for study participants' likelihood or propensity to attend AA prior to evaluating AA's impact on subsequent drinking. One approach, used in two studies of AA effectiveness, is an econometric method using so-called "instrumental variables" to parse-out AA attendance. The instrumental variables in one study were the availability of AA meetings in one's community and being able to drive to meetings;[39] after adjusting for these potential confounders, AA's effect on abstinence was reduced from OR = 3.70 ($P < .05$) to OR = 1.69 (not significant). Using different instrumental variables (perceived seriousness of drinking, and having a coping style tending towards information-seeking solutions), another study[40] found that AA's impact on heavy drinking was significant and doubled in magnitude after correcting for the instrumental variables. A third study[30] adjusted for baseline motivation and psychopathology as potential confounders and found that those with more AA involvement at 1 year had fewer alcohol problems at the 2-year follow-up interview. Another statistical study of selection bias used Propensity Scores to adjust for study participants' propensity to attend AA[42] and found that the odds of abstinence associated with AA attendance were reduced but remained significant after adjusting for individuals' propensity to attend AA. The method allowed investigators to study whether the selection bias operationalized by the Propensity Scores varied based on whether an individual had a low versus a high propensity to attend AA. AA's effect was minimal (e.g., OR = 1.3) among those with a high propensity to attend AA; however, the odds of abstinence associated with AA attendance were significant and of considerable magnitude, ranging from 2.7 to 6.9, among those with a lower propensity to attend AA.

What, then, is the scorecard for AA's effectiveness in terms of specificity? Among the rigorous experimental studies, there were two positive findings for AA effectiveness, one null finding and one negative finding. Among those that statistically addressed selection bias, there were two contradictory findings and two studies that reported significant effects for AA after adjusting for potential confounders such as motivation to change. Readers must judge for themselves whether their interpretation of these results, on balance, supports a recommendation that there is no experimental evidence of AA effectiveness (as put forward by the Cochrane review). As for the scorecard for the other criteria, the evidence for AA effectiveness is strong: rates of abstinence are approximately twice as high among those who attend AA (criteria 1, magnitude); higher levels of attendance are related to higher rates of abstinence (criteria 2, dose-response); these relationships are found for different samples and follow-up periods (criteria 3, consistency); prior AA attendance is predictive of subsequent abstinence (criteria 4, temporal); and mechanisms of action predicted by theories of behavior change are evident at AA meetings and through the AA steps and fellowship (criteria 6, plausibility).

REFERENCES

1. Ferri M, Amato L, Davoli M. Alcoholics Anonymous and other 12-step programmes for alcohol dependence. Cochrane Database of Systematic Reviews. 2006; 3: Art. No. CD005032.

2. Project MATCH Research Group. Matching alcoholism treatment to client heterogeneity: Project MATCH posttreatment drinking outcomes. J Stud Alcohol. 1997;58:7–29.

3. Project MATCH Research Group. Matching alcoholism treatments to client heterogeneity: Project MATCH three-year drinking outcomes. Alcohol Clin Exp Res. 1998;22:1300–11.

4. Moos RH, Moos BS. Participation in treatment and Alcoholics Anonymous: a 16-year follow-up of initially untreated individuals. J Clin Psychol. 2006;62:735– 50.

5. Moos RH, Moos BS. Help-seeking careers: connections between participation in professional treatment and Alcoholics Anonymous. J Subst Abuse Treat. 2004;26:167–73.

6. Emrick CD. Alcoholics Anonymous: affiliation processes and effectiveness as treatment. Alcohol Clin Exp Res 1987;11:416–23.

7. Kelly JF. Self-help for substance-use disorders: history, effectiveness, knowledge gaps, and research opportunities. Clin Psychol Rev. 2003;23:639–63.

8. Emrick CD, Tonigan JS, Montgomery HA, Little L. Alcoholics Anonymous: what is currently known? In: McCrady BS, Miller WR, eds. Research on Alcoholics Anonymous: opportunities and alternatives. New Brunswick, NJ: Rutgers Center of Alcohol Studies, 1993:41–78.

9. Tonigan JS, Toscova R, Miller WR. Meta-analysis of the literature on Alcoholics Anonymous: sample and study characteristics moderate findings. J Stud Alcohol. 1996;57:65–72.

10. Kownacki RJ, Shadish WR. Does Alcoholics Anonymous work? The results from a meta-analysis of controlled experiments. Subst Use Misuse. 1999;34:1897–916.

11. Bufe C. Alcoholics Anonymous: cult or cure? San Francisco, CA: Sharp Press, 1991.

12. Mausner JS, Kramer S. Epidemiology: an introductory text. 2nd ed. Philadelphia: W. B. Saunders Company, 1985.

13. Kleinbaum DG, Sullivan KM, Barker ND. ActivEpi companion textbook: a supplement for use with the ActivEpi CD-ROM. New York: Springer Science+Business Media, Inc., 2005.

14. Hill AB. Principles of medical statistics. 9th ed. New York: Oxford University Press, 1971.

15. Ouimette PC, Moos RH, Finney JW. Influence of outpatient treatment and 12-step group involvement on one-year substance abuse treatment outcomes. J Stud Alcohol. 1998;59:513–22.

16. Thurstin AH, Alfano AM, Nerviano VJ. The efficacy of AA attendance for aftercare of inpatient alcoholics: some follow-up data. Int J Addict. 1987;22:1083–90.

17. Cohen J. Statistical power analysis for the behavioral sciences. 2nd ed. Hillsdale, NJ: Lawrence Erlbaum Associates. 1988.

18. Fiorentine R. After drug treatment: are 12-step programs effective in maintaining abstinence? Am J Drug Alcohol Abuse. 1999; 25:93–116.

19. Timko C, Moos RH, Finney JW, Lesar MD. Long-term outcomes of alcohol use disorders: comparing untreated individuals with those in Alcoholics Anonymous and formal treatment. J Stud Alcohol. 2000;61:529–40.

20. Dawson DA, Grant BF, Stinson FS, Chou PS. Estimating the effect of help-seeking on achieving recovery from alcohol dependence. Addiction. 2006; 101:824–34.

21. Connors GJ, Tonigan JS, Miller WR. A longitudinal model of intake symptomatology, AA participation, and outcome: retrospective study of the Project MATCH outpatient and aftercare samples. J Stud Alcohol. 2001;62:817–25.

22. Tonigan JS, Connors GJ, Miller WR. Participation and involvement in Alcoholics Anonymous. In: Babor TF, Del Boca FK, eds. Treatment matching in alcoholism. Cambridge UK: Cambridge University Press. 2003:184–204.

23. Walsh DC, Hingson RW, Merrigan DM, Levenson SM, Cupples LA, Heeren T, Coffman GA, Becker CA, Barker TA, Hamilton SK, McGuire TG, Kelly CA. A randomized trial of treatment options for alcohol-abusing workers. N Engl J Med. 1991;325:775–82.

24. Timko C, Debenedetti A. A randomized controlled trial of intensive referral to 12-step self-help groups: one-year outcomes. Drug Alcohol Depend. 2007;90:270–9.

25. Brandsma JM, Maultsby MC Jr, Welsh RJ. Outpatient treatment of alcoholism: a review and comparative study. Baltimore, MD: University Park Press. 1980.

26. Ditman KS, Crawford GG, Forgy EW, Moskowitz H, MacAndrew C. A controlled experiment on the use of court probation for drunk arrests. Am J Psychiatry. 1967;124:160–3.

27. Khantzian EJ, Mack JE. Alcoholics Anonymous and contemporary psychodynamic theory. In: Galanter M, ed. Recent developments in alcoholism. New York: Plenum Press, 1989:67–89.

28. Tiebout HM. Therapeutic mechanisms of Alcoholics Anonymous. Am J Psychiatry. 1944;100:468–73.

29. Bandura A. Social learning theory. Morristown, NJ: General Learning Press, 1971.

30. White W, Laudet A. Life meaning as potential mediator of 12-step participation benefits on stable recovery from polysubstance use. In: The college on problems of drug dependence. Conference 2006, June 17–22: Scottsdale, AZ.

31. Kelly JF, Myers MG, Brown SA. Do adolescents affiliate with 12-step groups? A mulitvariate process model of effects. J Stud Alcohol. 2002;63:293–304.

32. Zemore SE. A role for spiritual change in the benefits of 12-step involvement. Alcohol Clin Exp Res. 2007;31:76S–9S.

33. Kaskutas LA, Bond J, Humphreys K. Social networks as mediators of the effect of Alcoholics Anonymous. Addiction. 2002;97:891–900.

34. Timko C, Finney JW, Moos RH. The 8-year course of alcohol abuse: gender differences in social context and coping. Alcohol Clin Exp Res. 2005;29:612–21.

35. Bond J, Kaskutas LA, Weisner C. The persistent influence of social networks and Alcoholics Anonymous on abstinence. J Stud Alcohol. 2003;64:579–88.

36. Humphreys K, Mankowski ES, Moos RH, Finney JW. Do enhanced friendship networks and active coping mediate the effect of self-help groups on substance abuse? Ann Behav Med. 1999;21:54–60.

37. Morgenstern J, Labouvie E, McCrady BS, Kahler CW, Frey RM. Affiliation with Alcoholics Anonymous following treatment: a study of its therapeutic effects and mechanisms of action. J Consult Clin Psychol. 1997;65:768–77.

38. Moos RH. Active ingredients of substance use-focused self-help groups. Addiction. 2008;103:387–96.

39. Fortney J, Booth B, Zhang M, Humphrey J, Wiseman E. Controlling for selection bias in the evaluation of Alcoholics Anonymous as aftercare treatment. J Stud Alcohol. 1998;59:690–7.

40. Humphreys K, Phibbs CS, Moos RH. Addressing self-selection effects in evaluations of mutual help groups and professional mental health services: an introduction to two-stage sample selection models. Eval Program Plann. 1996;19:301–8.

41. McKellar J, Stewart E, Humphreys K. Alcoholics Anonymous involvement and positive alcohol-related outcomes: cause, consequence, or just a correlate? A prospective 2-year study of 2,319 alcohol-dependent men. J Consult Clin Psychol. 2003;71:302–8.

42. Ye Y, Kaskutas LA, Bond J. Using propensity scores to adjust for selection bias when assessing the effectiveness of Alcoholics Anonymous in observational studies. 2008 (under review). Alcohol Research Group, Emeryville, CA.