Evaluating Alcoholics Anonymous's Effect on Drinking in Project MATCH Using Cross-Lagged Regression Panel Analysis

STEPHEN MAGURA, PH.D.,[a],* CHARLES M. CLELAND, PH.D.,[b] AND J. SCOTT TONIGAN, PH.D.[c]

[a]The Evaluation Center, Western Michigan University, Kalamazoo, Michigan
[b]School of Nursing, New York University, New York, New York
[c]Center on Alcoholism, Substance Abuse and Addictions, University of New Mexico, Albuquerque, New Mexico

ABSTRACT. Objective: The objective of the study is to determine whether Alcoholics Anonymous (AA) participation leads to reduced drinking and problems related to drinking within Project MATCH (Matching Alcoholism Treatments to Client Heterogeneity), an existing national alcoholism treatment data set. Method: The method used is structural equation modeling of panel data with cross-lagged partial regression coefficients. The main advantage of this technique for the analysis of AA outcomes is that potential reciprocal causation between AA participation and drinking behavior can be explicitly modeled through the specification of finite causal lags. Results: For the outpatient subsample ($n = 952$), the results strongly support the hypothesis that AA attendance leads to increases in alcohol abstinence and reduces drinking/ problems, whereas a causal effect in the reverse direction is unsupported. For the aftercare subsample ($n = 774$), the results are not as clear but also suggest that AA attendance leads to better outcomes. Conclusions: Although randomized controlled trials are the surest means of establishing causal relations between interventions and outcomes, such trials are rare in AA research for practical reasons. The current study successfully exploited the multiple data waves in Project MATCH to examine evidence of causality between AA participation and drinking outcomes. The study obtained unique statistical results supporting the effectiveness of AA primarily in the context of primary outpatient treatment for alcoholism. (J. Stud. Alcohol Drugs, 74, 378–385, 2013)

ALCOHOLICS ANONYMOUS (AA) REMAINS the most widely used single intervention for alcoholism in the United States and globally. In 2011, there were 57,905 AA groups and 1,279,664 members in the United States and 107,967 groups and 2,057,672 members worldwide (AA, 2011). AA is the only source of help many people with alcohol problems will ever use (Hasin and Grant, 1995).

A substantial research literature indicates a relationship between AA participation and less drinking or abstinence from alcohol. Tonigan et al.'s (1996) review includes studies up to that time; major studies since then have been Kaskutas et al. (2002, 2005), Gossop et al. (2003), Chappel and DuPont (1999), and Schuckit et al. (1997), as well as prior analyses for Project MATCH (Matching Alcoholism Treatments to Client Heterogeneity) (Magura et al., 2012; Kelly et al., 2011; Tonigan et al., 2003). With few exceptions, these studies are essentially correlational in nature, including those with longitudinal data. Even if spuriousness can be ruled out by appropriate statistical controls, this makes it difficult to establish the direction of any effect that is found. That is, although such correlations are invariably interpreted as AA participation causing changes in drinking, the actual direction of effect may be the reverse, or the effects may be in both directions.

Randomized controlled trials of interventions can determine whether interventions lead to desired outcomes, but randomized controlled trials of AA are rare because of practical problems in implementation. The few randomized controlled trials of AA that have been conducted are significantly limited in their methods or interpretability, and none reported a positive AA effect on drinking (Brandsma et al., 1980; Ditman et al., 1967; Walsh et al., 1991).

Instead of attempting to randomly assign alcoholics directly to AA or non-AA, several studies have randomly assigned them to either twelve-step facilitation (TSF) or alternative treatments. The primary purpose of TSF is to educate subjects about 12-step recovery and maximize their participation in community AA groups. TSF aims to prepare subjects for AA participation by facilitating "acceptance" of alcoholism as a disease that has become unmanageable, by facilitating "surrender" through giving oneself over to a higher power and the help of other recovering alcoholics, and by instilling hope for recovery through participation in a 12-step program (Nowinski and Baker, 2003).

Three studies that randomly assigned alcoholics to TSF found that the TSF subjects reported higher rates of alcohol abstinence at follow-up than non-TSF subjects, apparently because of higher AA participation (Litt et al., 2009; Timko and DeBenedetti, 2007; Walitzer et al., 2009). However, in

Project MATCH, drinking outcomes were equivalent among three outpatient treatments, one of which was TSF (Project MATCH Research Group, 1997).

The small number of controlled studies involving AA makes it desirable to learn as much as possible from correlational studies about the direction of effect between AA and drinking. One avenue is to better exploit the statistical potential of longitudinal study designs, in particular through the use of cross-lagged panel regression analysis (Kenny, 1979). The main advantage of this technique for the analysis of AA outcomes is that potential reciprocal causation between AA participation and drinking behavior can be explicitly modeled through the specification of "finite causal lags" (i.e., a cause precedes an effect and the two are separated by some finite amount of time). This seems to be a realistic assumption, because we would not necessarily expect instantaneous effects between AA participation and drinking.

We could identify only two examples of this approach in the AA research literature. McKellar et al. (2003) studied a sample of 2,319 male alcohol-dependent patients in 15 Veterans Administration inpatient programs. Cross-lagged regression panel analysis using structural equation modeling (SEM) indicated that 1-year posttreatment levels of AA involvement predicted lower alcohol use/problems at 2-year follow-up, whereas level of alcohol use/problems at 1 year did not predict AA involvement at 2-year follow-up. Additional models found that these effects were not attributable to motivation or psychopathology. Although this study was well done technically, the relatively long time lag between follow-up waves—1 year—may correspond poorly to actual finite causal lags between AA involvement and alcohol use/problems. Walitzer et al. (2009) studied 169 alcoholic outpatients randomly assigned to treatment as usual and two types of TSF: a 12-step-based directive approach and motivational enhancement approach. Cross-lagged regression panel analysis using SEM found that AA involvement during treatment significantly predicted alcohol abstinence at 4–6 months after treatment end, and AA involvement at 4–6 months after treatment end predicted alcohol abstinence at 10–12 months after treatment end; the reverse effects (abstinence predicting subsequent AA involvement) were not significant. Again, this was a well-conducted study, but the sample size is relatively small, all patients were recruited from a single alcohol research clinic, and generalizability is limited. It may be that there are few examples of this method in the alcoholism treatment or AA research literature because multiwave follow-up studies with large samples are rare, and, when there are multiwave data, the follow-ups tend to be spaced too far apart (e.g., 1 year or more) to make cross-lagged analysis sufficiently credible.

The goal of the present study was to determine whether AA participation leads to reduced drinking and problems related to drinking, based on cross-lagged panel regression analysis of the existing Project MATCH national alcohol-ism treatment data set. Prior analyses of the AA data in Project MATCH have not used this method (Kelly et al., 2011; Tonigan et al., 2003). The advantages of this data set for the present study are the large sample size, the multiple treatment sites nationally, the comprehensive set of potential covariates, and the availability of five waves of treatment and follow-up data.

Method

Subjects

Project MATCH was a national alcoholism treatment trial conducted between 1989 and 1998. Study subjects were recruited at outpatient sites ($n = 952$ at five sites) and aftercare sites ($n = 774$ at five sites), the latter after discharge from inpatient or day hospital treatment. More than 90% were diagnosed as dependent on alcohol, and none was dependent on illicit drugs with the exception of marijuana. Selected baseline characteristics of the Project MATCH sample are in Babor and Del Boca (2003: Table 5.1). The current secondary analysis of the data set was approved by the institutional review board of Western Michigan University.

Procedures

Volunteers underwent informed consent, completed a baseline assessment battery, and were randomly assigned to one of three treatments: TSF, cognitive–behavioral therapy (CBT), or motivational enhancement therapy (MET). TSF is described above. The primary focus of CBT was to increase ability to deal with high-risk situations that commonly lead to relapse to drinking, and the primary focus of MET was to mobilize the subject's own resources to bring about the changes needed to achieve sobriety. TSF and CBT had 12 scheduled weekly sessions and MET had 4 sessions, all over a 3-month period. Interview follow-ups were conducted at the scheduled end of treatment (after 3 months) and after 6, 9, 12, and 15 months. Follow-up rates were more than 90% in both subsamples at each follow-up. Project MATCH is described in detail in Project MATCH Research Group (1993) and Babor and Del Boca (2003).

Measures

The primary outcome measure for the present analysis is alcohol abstinence, as measured by "percentage days abstinent" (PDA) for each 3-month period after study intake (PDA3, PDA6, PDA9, PDA12, PDA15), which was arcsin transformed because of skewness in accord with all analyses previously reported for Project MATCH. The secondary drinking outcome is a previously validated outcome composite (OTCM) coded as (1) *abstinent from alcohol,* (2) *moderate drinking without problems related to drinking,*

(3) *heavy drinking or problems related to drinking,* and (4) *heavy drinking and problems related to drinking* (Cisler and Zweben, 1999; Zweben and Cisler, 2003). This outcome composite was also measured for each 3-month period after study intake (OTCM3, OTCM6, OTCM 9, OTCM12, OTCM15). This outcome composite has been shown to be a reasonable proxy measure for role functioning and quality of life (Zweben and Cisler, 2003); it was included in the analysis to examine whether the potential effects of AA appear to extend beyond effect on drinking, which certainly is the immediate objective of AA.

Drinking was captured on Form 90 (Miller and Del Boca, 1994), and problems were identified on the Drinker Inventory of Consequences (Miller et al., 1995). Percentage days AA attendance was measured on Form 90 for each 3-month period after study intake (AA3, AA6, AA9, AA12, AA15) and was also arcsin transformed.

Selection of covariates

Other Project MATCH variables can be roughly grouped into two categories, baseline characteristics of the subjects and postbaseline, time-varying potential mediators of the relation between AA attendance and outcomes. Baseline characteristics were screened for inclusion in the analyses, because these may be a source of confounding in assessing the relation between AA attendance and drinking. Potential mediators should not be included as covariates, because these could "wash out" any relation between AA attendance and outcomes, and modeling indirect effects between AA attendance and outcomes over time would greatly complicate the analyses. It is also important to note that previous studies using Project MATCH data already have identified putative mediators of AA effects on outcomes (e.g., Kelly et al., 2012).

The potential baseline confounding variables were identified as follows: Two new variables were computed that were the means of PDA and OTCM across the five follow-up waves. The pool of baseline variables included extensive sociodemographics, employment status, alcohol dependence score, psychiatric severity, cognitive impairment, and treatment assignment (TSF, MET, or CBT). These variables were screened for significant correlations with AA attendance and the two outcome variables ($p < .05$). Any baseline variable that had a significant correlation with either AA attendance or outcome was selected for inclusion as a covariate. Six covariates were identified: subject's age, White ethnicity, alcohol dependence score, living alone, living with spouse, and CBT/MET assignment. In addition, the baseline (3 months pre-treatment) measures of alcohol abstinence, AA attendance, and the respective outcome (PDA or OTCM) were included as covariates in all analyses. The same set of covariates was included in all analyses so that results would not be influenced by differences in covariate inclusion.

Data analysis plan

The method used is SEM of panel data with cross-lagged partial regression paths, as shown in Figure 1. The analysis can distinguish between effects of AA attendance on subsequent level of drinking and the effects of level of drinking on subsequent AA attendance (Kenny, 1979). This technique is well documented in the applied statistical literature (Cook and Campbell, 1979; Pugesek et al., 2003; van Montfort et al., 2004). Two assumptions required for this logic to hold are synchronicity (both variables at issue must be measured at the same point or period in time) and correct time lags (the measured time lags should be consistent with the actual causal lags). The data conform to the first assumption, but it is speculative whether a 3-month measurement interval captures the actual causal lag. (If it does not and there are null results, that may be the explanation.) The SEM analysis as conducted freely estimates regression coefficients without imposing equality constraints; thus, potential changes in the strengths of causal relations over time can be detected. The analysis is conducted for five successive 3-month data waves after study intake; baseline measures of AA participation and drinking are included as covariates. The analyses control for baseline variables (covariates), which are possible confounders. As explained above, this strengthens the conclusions that may be drawn about causality or causal direction. To summarize, the SEM models to be tested include the first-order cross-lagged regression paths between AA and the two outcome measures, the synchronous (same time) correlations between AA and the outcomes, the first-order regression paths between adjacent AA and adjacent outcome variables over time, and a consistent set of covariates selected as described above.

For PDA and AA attendance, the standardized partial regression coefficients are reported. For OTCM, the partial probit coefficients are reported.

Several "goodness-of-fit" indices are reported for the SEM models. These include one absolute fit index for all analyses (root mean square error of approximation, ≤.07 indicates good fit; Steiger, 2007) and a second absolute fit index suitable for either PDA or OTCM (standardized root mean square residual for PDA only, ≤.08 indicates good fit; Hu and Bentler, 1999) or, alternatively, weighted root mean square residual for OTCM only (<.90 indicates good fit; Muthén and Muthén, 1998–2010). Two incremental fit indices are reported: comparative fit index (≥.95 indicates good fit) and the Tucker–Lewis index (≥.95 or greater; Hu and Bentler, 1999).

Most previous analyses of Project MATCH have analyzed the outpatient and aftercare subsamples separately, and we continue that practice here. Missing data are handled by the full information maximum likelihood method. The SEM analysis was conducted using the Mplus software (Muthén and Muthén, 1998–2010). For OTCM as the outcome, the

Mplus option of an ordinal measurement scale was selected, which yielded estimates of ordinal probit coefficients via weighted least squares.

Power analysis

The analysis was planned to estimate SEM models for the outpatient and aftercare subsamples as shown in Figures 1 and 3. The most pertinent question is, "What is the power of the analysis for estimating the magnitudes of the cross-lagged regression paths?" In the smaller aftercare sample (n = 697 for those with complete data), power is .80 to detect a cross-lagged effect as small as R^2 = .017 (semipartial r = .13) with up to 50 covariates. A second question is, "What is the power to detect overall lack of fit (i.e., discrepancies between the model-implied and observed covariance matrices)?" Degrees of freedom were 32 for the models with PDA and 36 for the models with OTCM. Power was greater than .99 for overall tests of exact fit, close fit, and not close fit (MacCallum et al., 1996). Therefore, power is sufficient to detect small cross-lagged effects and overall lack of fit of the data to the model.

Results

Structural equation modeling results: Outpatient subsample

Figure 1 presents the SEM results for PDA in the outpatient subsample. As expected, the amount of AA attendance

during any previous 3 months strongly predicts the amount of AA attendance in the subsequent 3 months; the same holds true for PDA. All four lagged regression paths from AA to PDA are statistically significant, indicating that higher AA attendance in the previous period predicts higher abstinence in the subsequent period. None of the lagged paths from PDA to AA are statistically significant. Three of the fit indices conform to the standards for a good fit, whereas the Tucker–Lewis index does not.

Figure 2 presents the results for OTCM in the outpatient subsample. Two of the four lagged regression paths from AA to OTCM are statistically significant, and one of the four lagged paths from OTCM to AA is statistically significant. The significant negative coefficients indicate that higher AA attendance in the prior period predicts lower drinking/problems in the subsequent period, whereas the significant positive coefficient indicates that higher drinking/problems predicts higher AA attendance in the subsequent period. All four of the fit indices conform to the standards for a good fit.

Considering both PDA and OTCM (the outcomes), six of eight path coefficients from AA to these outcomes were statistically significant. The probability of this occurring, given a .05 type I error of falsely rejecting a null hypothesis of nonsignificance for any single coefficient, is $p < .001$ (by exact binomial distribution). One of the eight path coefficients from the outcomes to AA was statistically significant; the probability of this occurring is consistent with chance occurrence ($p > .05$). Last, the probability of this observed difference in the significance of the cross-lagged paths (6 / 8
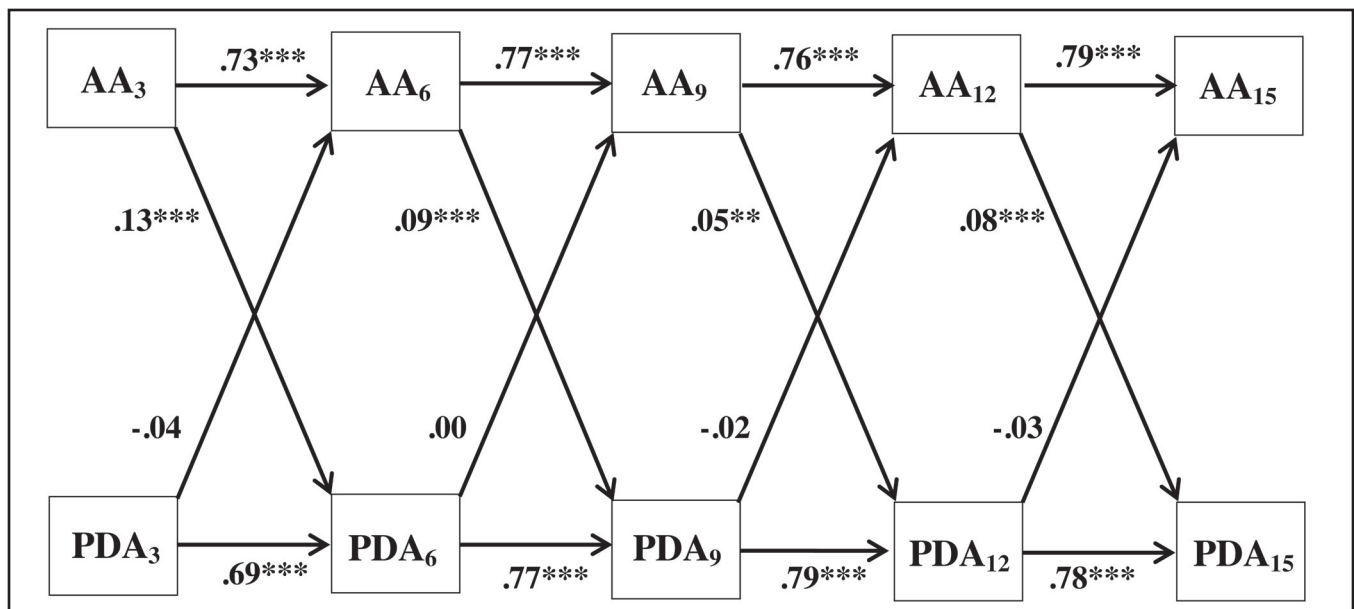

FIGURE 1. Cross-lagged panel analysis: Percentage days abstinent (PDA) and Alcoholics Anonymous (AA) attendance at 3, 6, 9, 12 and 15 months after intake–outpatient sample (N = 952). Comparative fit index = .98; Tucker–Lewis index = .92; root mean square error of approximation = .07; standardized root mean square residual = .02. Covariates measured at baseline: age, White ethnicity, living alone, living with spouse, cognitive–behavioral therapy/motivational enhancement therapy assignment, AA attendance, alcohol dependence score, and alcohol abstinence (PDA). Standardized regression coefficients are reported. **$p < .01$; ***$p < .001$.
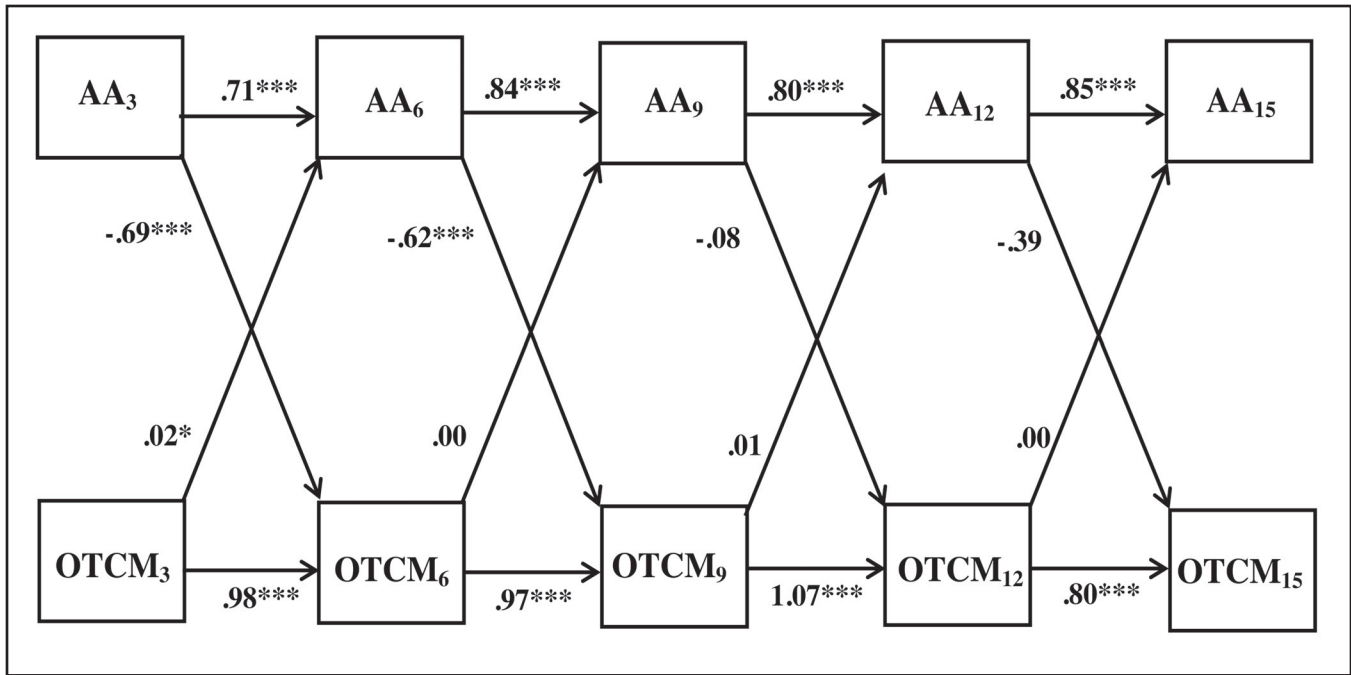
FIGURE 2.   Cross-lagged panel analysis: Composite outcomes (OTCM) and Alcoholics Anonymous (AA) attendance at 3, 6, 9, 12 and 15 months after intake—outpatient sample ($N = 952$). Comparative fit index = .99; Tucker–Lewis index = .96; root mean square error of approximation = .06; weighted root mean square residual = .45. Covariates measured at baseline: age, White ethnicity, living alone, living with spouse, cognitive–behavioral therapy/motivational enhancement therapy assignment, AA attendance, alcohol dependence score and outcome composite (OTCM). Partial probit coefficients are reported. *$p < .05$; ***$p < .001$.
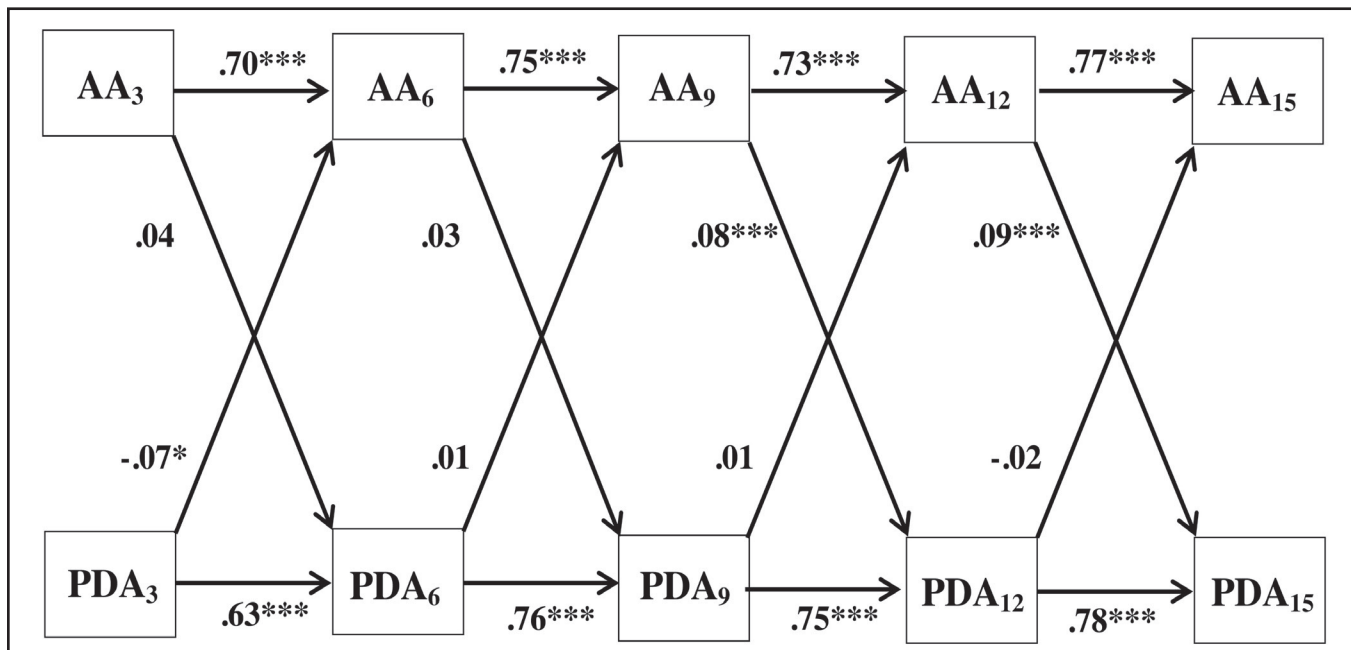

FIGURE 3.   Cross-lagged panel analysis: Percentage days abstinent (PDA) and Alcoholics Anonymous (AA) attendance at 3, 6, 9, 12 and 15 months after intake—aftercare sample ($n = 774$). Comparative fit index = .98; Tucker–Lewis index = .90; root mean square error of approximation = .07; standardized root mean square residual = .02. Covariates measured at baseline: age, White ethnicity, living alone, living with spouse, cognitive–behavioral therapy/motivational enhancement therapy assignment, AA attendance, alcohol dependence score, and alcohol abstinence (PDA). Standardized regression coefficients are reported. *$p < .05$; ***$p < .001$.
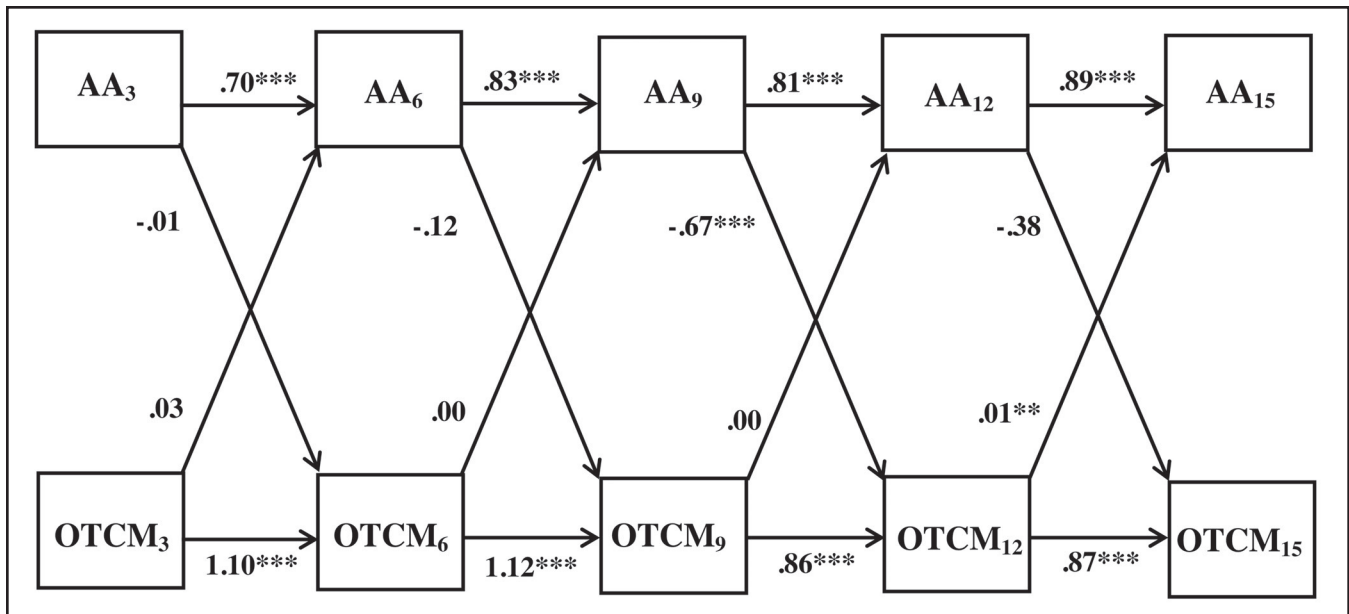
FIGURE 4.   Cross-lagged panel analysis: Outcome composite (OTCM) and Alcoholics Anonymous (AA) attendance at 3, 6, 9, 12 and 15 months after intake—aftercare sample (*n* = 774). Comparative fit index = .99; Tucker-Lewis index = .97; root mean square error of approximation = .05; weighted root mean square residual = .39. Covariates measured at baseline: age, White ethnicity, living alone, living with spouse, cognitive–behavioral therapy/motivational enhancement therapy assignment, AA attendance, alcohol dependence score and outcome composite (OTCM). Partial probit coefficients are reported. **p < .01; ***p < .001.

vs. 1 / 8) occurring by chance is p <.05 (Fisher's exact test, two tailed).

SEM results: Aftercare subsample

Figure 3 presents the SEM results for PDA in the aftercare subsample. Two of the four lagged regression paths from AA to PDA are statistically significant, indicating that higher AA attendance in the previous period predicts higher abstinence in the subsequent period. One path from PDA to AA is statistically significant; the negative coefficient indicates that less abstinence in the prior period predicts higher AA attendance in the subsequent period. Three of the fit indices conform to the standards for a good fit, whereas the Tucker–Lewis index does not.

Figure 4 presents the results for OTCM in the aftercare subsample. One of the four lagged paths from AA to OTCM is statistically significant, and one of the four lagged paths from PDA to AA is statistically significant. The negative path from AA to OTCM signifies that higher AA attendance in the prior period predicts lower drinking/problems in the subsequent period. The positive path from OTCM to AA indicates that higher drinking/problems predict higher AA attendance. All four of the fit indices conform to the standards for a good fit.

Considering both PDA and OTCM, three of eight path coefficients from AA to these outcomes were statistically significant. The probability of this occurring, given a .05 type I

error of falsely rejecting a null hypothesis of nonsignificance for any single coefficient, is p <.01. Similarly, the probability of two out of eight coefficients from the outcomes to AA appearing as significant was not quite significant (p = .06). There was no significant difference in the proportions of significant effects between the cross-lagged paths (3 / 8 vs. 2 / 8).

Discussion

Randomized controlled trials of mutual aid, including AA, are very difficult to conduct. Correlational studies of mutual aid, even if they have a longitudinal design, are difficult to interpret due to potential confounds in the analysis that are not amenable to control. The current study successfully exploited the multiple data waves in Project MATCH to examine evidence of causality between AA participation and drinking outcomes. The study obtained unique statistical results supporting the effectiveness of AA primarily in the context of primary outpatient treatment for alcoholism. In particular, for the outpatient subsample, the results strongly support the hypothesis that AA participation leads to increases in alcohol abstinence and reduces drinking problems, whereas a causal effect in the reverse direction is unsupported. For the aftercare subsample, the results are less clear because fewer significant effects were found. However, even in the aftercare sample, AA predicted outcomes more often than attributable to chance, whereas the number of

significant coefficients indicating the reverse effect could have occurred by chance. Thus, overall, based on statistical considerations, the findings support an effect of AA participation on outcomes in both subsamples, whereas the reverse effect is not supported for either subsample.

This analysis is quite different from previous studies examining relationships involving AA in Project MATCH. Tonigan et al. (2003) reported significant correlations between AA attendance and positive drinking outcomes during the same periods. Kelly et al. (2012) found that effect of AA attendance (as occurring during the 3-month treatment phase) on positive drinking outcomes was mediated by several theoretically driven factors such as adaptive social network changes and increases in social abstinence self-efficacy. Neither of these studies examined a potential reciprocal causal relationship between AA and drinking.

The stronger results for the outpatient as compared with the aftercare sample cannot be definitively explained, but one possible explanation relates to the observed higher AA attendance of the aftercare subjects. As indicated by Tonigan et al. (2003), this higher AA attendance may be due to their prior inpatient treatment experience, which generally encouraged clients to attend AA. As a result, some of these aftercare clients may have been less self-motivated to attend AA and consequently perhaps less susceptible to influence by AA. The possibly lower self-motivation to attend AA by the aftercare subjects may be indicated by their steeper decline in AA attendance over the 15 month study period as compared with the outpatient subjects (Tonigan et al., 2003).

The study bolsters the existing evidence base for AA by demonstrating a probable causal link between AA attendance and drinking outcomes. In that regard, clinicians are justified in recommending AA to their patients as an effective tool of recovery. The only caveat would be that, at least according to Project MATCH, AA as aftercare may be less effective than AA during primary outpatient treatment. This is somewhat contrary to the prevailing view that AA is especially useful as an aftercare modality.

Limitations

The effects of AA were found in the context of an outpatient or aftercare treatment platform, albeit the treatment only occurred within the first 3 months of the 15-month study period. Nevertheless, we cannot conclude that exactly the same AA effects would be found without that initial treatment experience.

The causal lag of 3 months in the analysis may or may not be the ideal period of lag for identifying AA effects. A shorter lag time (e.g., 1 month) may have shown larger effects, but this is speculative. Certainly long lags (e.g., 1 year) are usually presumed to be unsuitable for cross-lagged panel analysis of individual behavior.

Directions for future research

The findings suggest that future research must distinguish clearly among the phases of treatment or posttreatment in which AA participation occurs in order to better understand why AA during aftercare might be less effective than typically assumed. Although the results did not justify concluding that level of drinking or drinking problems affect level of AA attendance, several significant coefficients were found, indicating that this possibility should be examined in greater detail by future researchers. The study also has implications for any research involving AA or other substance use interventions or treatments in which randomization is not feasible. Although frequent follow-ups add expense to a study, this must be weighed against the advantage of ascertaining causal effects of interventions on outcomes in observational or naturalistic studies.

References

Alcoholics Anonymous. (2011). Estimates of A. A. groups and members as of January 1, 2011. Retrieved from http://www.aa.org/en_pdfs/smf-53_en.pdf

Babor, T. F., & Del Boca, F. K. (Eds.). (2003). *Treatment matching in alcoholism.* Cambridge, England: Cambridge University Press.

Brandsma, J. M., Maultby, M. C., & Welsh, R. J. (1980). *Outpatient treatment of alcoholism: A review and comparative study.* Baltimore, MD: University Park Press.

Chappel, J. N., & DuPont, R. L. (1999). Twelve-step and mutual-help programs for addictive disorders. *Psychiatric Clinics of North America, 22,* 425–446.

Cisler, R. A., & Zweben, A. (1999). Development of a composite measure for assessing alcohol treatment outcome: Operationalization and validation. *Alcoholism: Clinical and Experimental Research, 23,* 263–271.

Cook, T. D., & Campbell, D. T. (1979). *Quasi-experimental design and analysis: Issues for field settings.* Skokie, IL: Rand-McNally.

Ditman, K. S., Crawford, G. G., Forgy, E. W., Moskowitz, H., & Macandrew, C. (1967). A controlled experiment on the use of court probation for drunk arrests. *American Journal of Psychiatry, 124,* 160–163.

Gossop, M., Harris, J., Best, D., Man, L. H., Manning, V., Marshall, J., & Strang, J. (2003). Is attendance at Alcoholics Anonymous meetings after inpatient treatment related to improved outcomes? A 6-month follow-up study. *Alcohol and Alcoholism, 38,* 421–426.

Hasin, D. S., & Grant, B. F. (1995). AA and other helpseeking for alcohol problems: Former drinkers in the U.S. general population. *Journal of Substance Abuse, 7,* 281–292.

Hu, L.-T., & Bentler, P. M. (1999). Cutoff criteria for fit indices in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling: A Multidisciplinary Journal, 6,* 1–55.

Kaskutas, L. A., Ammon, L., Delucchi, K., Room, R., Bond, J., & Weisner, C. (2005). Alcoholics Anonymous careers: Patterns of AA involvement five years after treatment entry. *Alcoholism: Clinical and Experimental Research, 29,* 1983–1990.

Kaskutas, L. A., Bond, J., & Humphreys, K. (2002). Social networks as mediators of the effect of Alcoholics Anonymous. *Addiction, 97,* 891–900.

Kelly, J. F., Hoeppner, B., Stout, R. L., & Pagano, M. (2012). Determining the relative importance of the mechanisms of behavior change within Alcoholics Anonymous: A multiple mediator analysis. *Addiction, 107,* 289–299.

Kelly, J. F., Stout, R. L., Magill, M., Tonigan, J. S., & Pagano, M. E. (2011).

Spirituality in recovery: A lagged mediational analysis of Alcoholics Anonymous' principal theoretical mechanism of behavior change. *Alcoholism: Clinical and Experimental Research, 35,* 454–463.

Kenny, D. A. (1979). *Correlation and causality.* New York, NY: Wiley.

Litt, M. D., Kadden, R. M., Kabela-Cormier, E., & Petry, N. M. (2009). Changing network support for drinking: Network support project 2-year follow-up. *Journal of Consulting and Clinical Psychology, 77,* 229–242.

MacCallum, R. C., Browne, M. W., & Sugawara, H. M. (1996). Power analysis and determination of sample size for covariance structure modeling. *Psychological Methods, 1,* 130–149.

Magura, S., McKean, J., Kosten, S., & Tonigan, J. S. (2012). A novel application of propensity score matching to estimate Alcoholics Anonymous' effect on drinking outcomes. *Drug and Alcohol Dependence.* Advance online publication. Retrieved from: http://dx.doi.org/10.1016/j.drugalcdep.2012.09.011

McKellar, J., Stewart, E., & Humphreys, K. (2003). 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. *Journal of Consulting and Clinical Psychology, 71,* 302–308.

Miller, W. R., & Del Boca, F. K. (1994). Measurement of drinking behavior using the Form 90 family of instruments. *Journal of Studies on Alcohol, Supplement 12,* 112–118.

Miller, W. R., Tonigan, J. S., & Longabaugh, R. (1995). *The Drinker Inventory of Consequences (DrInC).* (Project MATCH Monograph Series, Vol. 4, NIH Publication No. 95-3911). Rockville, MD: National Institute on Alcohol Abuse and Alcoholism.

Muthén, L.K., & Muthén, B.O. (1998–2010). *Mplus user's guide* (6th ed.). Los Angeles, CA: Authors. p.10

Nowinski, J., & Baker, S. 2003. *The twelve-step facilitation handbook.* Center City, MN: Hazelden.

Project MATCH Research Group. (1993). Project MATCH: Rationale and methods for a multisite clinical trial matching patients to alcoholism treatment. *Alcoholism: Clinical and Experimental Research, 17,* 1130–1145.

Project MATCH Research Group. (1997). Matching Alcoholism Treatments to Client Heterogeneity: Project MATCH posttreatment drinking outcomes. *Journal of Studies on Alcohol, 58,* 7–29.

Pugesek, B. H., Tomer, A., & von Eye, A. (2003). *Structural equation modeling.* Cambridge, England: Cambridge University Press.

Schuckit, M. A., Tipp, J. E., Smith, T. L., & Bucholz, K. K. (1997). Periods of abstinence following the onset of alcohol dependence in 1,853 men and women. *Journal of Studies on Alcohol, 58,* 581–589.

Steiger, J. H. (2007). Understanding the limitations of global fit assessment in structural equation modeling. *Personality and Individual Differences, 42,* 893–898.

Timko, C., & DeBenedetti, A. (2007). A randomized controlled trial of intensive referral to 12-step self-help groups: One-year outcomes. *Drug and Alcohol Dependence, 90,* 270–279.

Tonigan, J. S., Conners, G. J., & Miller, W. R. (2003). Participation and involvement in Alcoholics Anonymous. In T. F. Babor and F. K. Del Boca (Eds.), *Treatment matching in alcoholism* (pp. 184–204). Cambridge, England: Cambridge University Press.

Tonigan, J. S., Toscova, R., & Miller, W. R. (1996). Meta-analysis of the literature on Alcoholics Anonymous: Sample and study characteristics moderate findings. *Journal of Studies on Alcohol, 57,* 65–72.

van Montfort, K., Oud, J., & Satorra, A. (Eds.). (2004). *Recent developments on structural equation models: Theory and applications (Mathematical Modeling: Theory and Applications).* New York, NY: Springer.

Walitzer, K. S., Dermen, K. H., & Barrick, C. (2009). Facilitating involvement in Alcoholics Anonymous during out-patient treatment: A randomized clinical trial. *Addiction, 104,* 391–401.

Walsh, D. C., Hingson, R. W., Merrigan, D. M., Levenson, S. M., Cupples, L. A., Heeren, T., . . . Kelly, C. A. (1991). A randomized trial of treatment options for alcohol-abusing workers. *The New England Journal of Medicine, 325,* 775–782.

Zweben, A., & Cisler, R. A. (2003). Clinical and methodological utility of a composite outcome measure for alcohol treatment research. *Alcoholism: Clinical and Experimental Research, 27,* 1680–1685.