Simultaneous modeling of the impact of treatments on alcohol consumption and quality of life in the COMBINE study: a coupled hidden Markov analysis

James J. Prisciandaro, PhD[1], **Stacia M. DeSantis, PhD**[2], and **Dipankar Bandyopadhyay, PhD**[3]

[1]Medical University of South Carolina; Department of Psychiatry and Behavioral Sciences; Clinical Neuroscience Division; 67 President Street, MSC861; Charleston, SC 29425

[2]Medical University of South Carolina; Department of Medicine; Division of Biostatistics and Epidemiology; 135 Cannon Street, MSC835; Charleston, SC 29425

[3]University of Minnesota School of Public Health; Division of Biostatistics; A452 Mayo Building, MMC 303; 420 Delaware Street SE; Minneapolis, MN 55455

Abstract

Background—Clinical trials for alcoholism have historically regarded alcohol consumption as the primary outcome. In a subset of trials, quality of life has been considered as a secondary outcome. Joint latent-variable modeling techniques may provide a more accurate and powerful simultaneous analysis of primary and secondary outcomes in clinical trials. The goal of the present study was to evaluate longitudinal associations between treatment status, alcohol consumption, and quality of life in the Combined Pharmacotherapies and Behavioral Interventions for Alcohol Dependence (COMBINE) study.

Methods—1,383 alcohol-dependent patients were randomized to nine treatment groups. Percent heavy drinking days (PHDD) and health-related QOL from the 30 days preceding baseline, week 16, and week 52 were calculated using the Form 90 and the Medical Outcomes Study Health Survey Short Form-12 (SF-12), respectively. Latent profile analysis (LPA) was conducted to determine an appropriate number of latent states to represent PHDD and QOL. Subsequently, univariate and coupled Hidden Markov Model (HMM)s (for PHDD f& SF-12 mental health and PHDD & SF-12 physical) were fit to the data.

Results—LPA suggested that PHDD should be represented by 3 latent states and that each SF-12 scale should be represented by 2 states. Joint modeling results suggested that: (i) naltrexone significantly predicted decreased PHDD (p<0.05), and marginally predicted improved mental health QOL via decreased PHDD (p<0.10), and (ii) that the combinations of naltrexone and combined behavioral intervention (CBI), and acamprosate and CBI, each predicted significantly improved physical QOL (p<0.05), and marginally predicted decreased PHDD via improved physical QOL (p<0.10).

Conclusions—The present study illustrates a powerful and novel statistical approach for simultaneously evaluating the impact of treatments on primary and secondary outcomes in clinical trials. The present study also suggests that behavioral interventions may impact drinking behavior through their ameliorative effects on QOL.

Corresponding author: James J. Prisciandaro; Medical University of South Carolina; Department of Psychiatry and Behavioral Sciences; Clinical Neuroscience Division; 67 President Street, MSC861; Charleston, SC 29425; telephone: (843) 792-1433; fax: (843) 792-0528; prisian@musc.edu.

Keywords

hidden Markov model; COMBINE; joint analysis; alcoholism; alcohol dependence; naltrexone; acamprosate; behavioral; latent state

---

INTRODUCTION

Clinical trials for alcoholism have traditionally focused on symptomatic improvement as the primary, and often sole, outcome (Donovan et al., 2005). More recently, clinical trials have begun evaluating the impact of treatments on the broader welfare of patients, as measured by psychosocial functioning and/or quality of life (QOL). This change reflects acknowledgment that psychosocial functioning/QOL are important outcomes in their own right, and that clinically significant symptom improvement does not always translate into clinically significant improvements in general welfare (Donovan et al., 2005).

Clinical trials that have evaluated patients' general welfare have focused on health-related QOL (Turner-Bowker et al., 2002), which reflects patients' subjective evaluations of their physical and mental health (CDC, 2000). Research has demonstrated that heavy drinking is associated with decreased health-related QOL (Kraemer et al., 2002; Okoro et al., 2004), that individuals with alcohol dependence report levels of health-related QOL that are equivalent to individuals with chronic diseases or psychiatric disorders (Ware, 1999), and that detoxification, abstinence, and treatment may be associated with increased QOL (Morgan et al., 2003; Morgan et al., 2004).

The Combined Pharmacotherapies and Behavioral Interventions for Alcohol Dependence (COMBINE) study was the largest, controlled pharmacotherapy study for the treatment of alcohol dependence (Anton et al., 2006). Nearly 1,400 alcohol-dependent patients, across 11 sites, were randomized to 9 treatment groups representing combinations of pharmacological, behavioral, and placebo interventions. Treatments were delivered over 16 weeks and participant follow-up continued for 52 weeks post-treatment. Study outcomes included both drinking and quality of life. Regarding alcohol consumption, the COMBINE study found that naltrexone, CBI, or both significantly improved drinking outcomes relative to placebo, and that combined naltrexone and CBI was not superior to either naltrexone or CBI alone in improving drinking outcomes. A secondary analysis of non-drinking outcomes in the COMBINE study (LoCastro et al., 2009) found that baseline and week 52 percent heavy drinking days (PHDD) were negatively correlated with physical QOL, and that week 16 PHDD was negatively correlated with both physical and mental health QOL. Both physical and mental health QOL improved from baseline to week 16 across treatment groups. However, whereas mental health QOL remained elevated relative to baseline at week 52, physical QOL fell below baseline at week 52. Interestingly, only the combination of naltrexone and CBI predicted improved physical QOL in adjusted models; no treatments, or treatment combinations, significantly predicted improved mental health QOL.

Two studies have re-analyzed the drinking data from the COMBINE study using latent-variable modeling (Gueorguieva et al., 2010; Witkiewitz et al., 2010). They argued that because drinking behavior in clinical trials is erratic and particularly prone to measurement error and outliers, latent-variable techniques may result in improved representation of the distribution of drinking behavior as they reduce the impact of measurement error and outliers on parameter estimates (Shirley et al., 2010). Gueorguieva and colleagues (2010) conducted latent-trajectory analyses on the COMBINE data and found that naltrexone, CBI, and combined naltrexone and CBI each significantly impacted different trajectories of alcohol consumption (Gueorguieva et al., 2010). Witkiewitz and colleagues (2010) similarly

evaluated the relative goodness-of-fit of trajectory-based models and hidden Markov models (HMM; a first-order Markov process with latent categorical states; Shirley et al., 2010) to drinking behavior in the COMBINE study.

To date, no study has examined the impact of treatment on the multivariate joint outcome of drinking behavior and health-related QOL in the COMBINE study. The present study constructed coupled HMM models to directly estimate the impact of treatment assignment on drinking and QOL while explicitly modeling the dynamic relationship between drinking and QOL over time. In addition to the statistical benefits afforded by latent-variable models that were discussed above, the present joint analytic approach has the added benefits of potentially increased accuracy, and potentially improved statistical power, of model parameter estimates (Holmes et al., 2011).

MATERIALS AND METHODS

Participants and procedures

The COMBINE study methods have been detailed previously (Anton et al., 2006; COMBINE, 2003). Briefly, 4,965 individuals, recruited from 11 sites, were screened for eligibility. Inclusion criteria included: (i) DSM-IV diagnosis of alcohol dependence via the Structured Clinical Interview for DSM-IV (First et al., 1996), (ii) 4–21 days of abstinence, and (iii) 14 (women) or 21 (men) drinks per week, with 2 heavy drinking days (4 drinks/day [women], 5 drinks/day [men]), during a consecutive 30 period within the 90 days prior to baseline. Exclusion criteria included: (i) recent history of substance abuse (other than alcohol, nicotine, and marijuana), (ii) psychiatric illness requiring medication, and (iii) unstable medical condition (Anton et al., 2006).

1,383 participants (31% women, 69% men) met study criteria and were randomly assigned to one of nine treatment groups following baseline assessment. Eight of the treatment groups included active medication or placebo plus medical management; the ninth group received a combined behavioral intervention (CBI) only. Groups receiving medical management included: (i) naltrexone, (ii) acamprosate, (iii) naltrexone + acamprosate, (iv) placebo, (v) naltrexone + CBI, (vi) acamprosate + CBI, (vii) naltrexone + acamprosate + CBI, and (viii) placebo + CBI. Participants received treatment for 16 weeks. Medical management was offered at 9 visits; CBI was offered for 20 sessions. Following the end of active treatment (week 16), participants were followed for 52 weeks (at weeks 26, 52, and 68).

Measures

Alcohol consumption data were collected using the Form 90 structured interview (Miller and Del Boca, 1994). Daily data were available from the 90 days preceding the baseline assessment to the end of the follow-up. The drinking summary variable, percent heavy drinking days (PHDD), where a heavy drinking day is defined as 4 drinks/day for women and 5 drinks/day for men, was used for the present analysis because it captures both frequency and quantity of alcohol use; a number of previous COMBINE studies have examined PHDD (Anton et al., 2006; Gueorguieva et al., 2010; LoCastro et al., 2009; Witkiewitz et al., 2010).

QOL was assessed using the Medical Outcome Study (MOS) Health Survey, Short Form-12, version 2 (SF-12v2; Ware et al., 1996) as it is the most commonly used quality of life measure in alcoholism research (Donovan et al., 2005). The SF-12 (12-item) reproduces at least 90% of the variance contained in longer (36-item) versions, and provides norm-based standardized scoring (z-scores) for two summary scales: physical and mental health (higher scale scores represent better quality of life). The physical scale encompasses physical functioning, physical role impairment, and pain; the mental health scale encompasses

vitality, social functioning, emotional role impairment, and mental health symptoms (Ware et al., 1994, 1996). The COMBINE study used the standard 4-week recall version of the SF-12. Participants completed the measure at baseline, week 16, and week 52. Given that SF-12 data covered three, 4-week retrospective recall periods, our alcohol consumption variables were constructed to cover the same three periods.

Data analysis

The present study used hidden Markov modeling (HMM) to evaluate longitudinal associations between treatment status, alcohol consumption, and quality of life (MacDonald and Zucchini, 1997). Univariate HMM allows for estimation of the prospective impact of covariates (e.g., treatment group) on transitions between latent states on a given outcome of interest (e.g., light vs. heavy alcohol use). HMM is preferable to non-latent-variable techniques because it is robust to measurement error and outliers. HMM has been prominent in recent alcohol (Shirley et al., 2010; Wall and Li, 2009; Prisciandaro et al., in press; Witkiewitz et al., 2010) and cocaine (DeSantis et al., 2009; DeSantis and Bandyopadhyay, 2011) research, because it readily accommodates the chaotic nature of alcohol/drug use behavior in clinical trials. Witkiewitz and colleagues (2010) demonstrated that HMM provided a superior representation of the COMBINE data relative to other statistical modeling techniques. In the present study, we extended univariate implementations of HMM by estimating coupled (joint) HMMs (Brand et al., 1996). The coupled HMM is analogous to an autoregressive cross-lagged panel model, but with each occasion of each construct represented by a latent state variable. The present study estimated two coupled hidden Markov models (PHDD and SF-12 mental health, PHDD and SF-12 physical) in Mplus 6.1 software (Muthen and Muthen, 2011). All analyses included study site as a clustering variable (Anton et al., 2006).

Step 1. Latent profile analyses—First, we conducted separate latent profile analyses (LPA; latent mixture analyses with continuous indicators) for each construct (PHDD, SF-12 mental health, SF-12 physical) at each time point (baseline, week 16, week 52) to determine the number of latent states (normally-distributed subpopulations) that should be used to represent each outcome in HMM analyses (Muthen and Muthen, 2011). We used the Lo-Mendell-Rubin likelihood ratio test (LMR-LRT; Lo et al., 2001), Bayesian Information Criterion (Schwartz, 1978), and the sample-size adjusted BIC (Sclove, 1987) to choose an optimal number of states for each construct; significant LMR-LRT values indicate that a given "n" (e.g., 3) state solution fits the data significantly better than a "n-1" (e.g., 2) state solution, and lower BIC values indicate better parsimony-adjusted fit among models within the same assessment occasion. To ensure that HMMs would be estimable, interpretable, and reliable, we imposed two additional constraints on model selection: (i) for each construct, the selected number of latent states, and the mean of each latent state, had to be the same for every time-point, and (ii) each state within a given LPA solution had to contain at least 5% of the sample. LPA models were estimated using maximum likelihood with robust standard errors (MLR; Muthen and Muthen, 2011). To avoid local maxima, 100 automatically generated sets of initial stage random starting values were evaluated for each model.

Step 2. Univariate hidden Markov modeling—After determining the optimal number of latent states for each construct, we constructed three univariate HMMs (PHDD, SF-12 mental health, SF-12 physical). Model parameterization of HMM is described in detail in Langeheine and van de Pol (2002; pp. 323–329) as well as the Mplus user's guide (Example 8.12; Muthen and Muthen, 2011). In the present study, the MLR estimator was used and 100 automatically generated sets of initial stage random starting values were evaluated for each model. For each construct, we first estimated HMMs with transition matrices (autoregressive paths and latent state intercepts) constrained to be equal across measurement occasions.

Then, we estimated HMMs with transition matrices freed and calculated the difference in fit between the constrained and unconstrained solutions using BIC values and scaled log likelihood ratio tests (Muthen and Muthen, 2011). Equality constraints for transition matrices were statistically evaluated because assessment occasions were not equally spaced and because interventions were manipulated between baseline and week 16 but not between weeks 16 and 52.

Step 3. Coupled (joint) hidden Markov modeling—Coupled HMMs feature two simultaneously estimated HMMs that are connected to one another via cross-lagged autoregressive paths (e.g., baseline PHDD predicting week 16 SF-12, baseline SF-12 predicting week 16 PHDD). As with cross-lagged panel models, lagged associations in coupled HMMs are estimated controlling for the influence of the previous within-construct measurement occasion. Two sets of coupled HMM models were estimated: (i) PHDD and SF-12 mental health, and (ii) PHDD and SF-12 physical. For each set of models, we evaluated the appropriateness of equality constraints for the cross-lagged regression coefficients. Differences in fit between constrained and unconstrained models were evaluated using BIC values and scaled log likelihood ratio tests.

Next, coupled HMMs were re-estimated with treatment covariates predicting PHDD and SF-12 transition matrices. There were seven treatment covariates examined: (i) acamprosate, (ii) naltrexone, (iii) CBI, (iv) acamprosate × naltrexone, (v) acamprosate × CBI, (vi) naltrexone × CBI, (vii) acamprosate × naltrexone × CBI. Consistent with Anton and colleagues (2006) and LoCastro and colleagues (2009), treatment analyses did not include the ninth, combined behavioral intervention only group. Treatment covariate effects were estimated both with and without within-construct (PHDD, SF-12) equality constraints. These constraints were imposed and evaluated separately for each treatment covariate. Figure 1 provides a schematic for the estimated coupled HMMs with treatment covariates. Univariate HMM reports typically include probabilities of transitioning among latent states across time. Mplus does not currently provide these probabilities for coupled HMMs; therefore, we reported associations among latent states using odds ratios. Because transition probabilities and odds ratios are both used purely for description, the substitution of one for the other does not affect model estimation or parameter significance in any way.

Following estimation of the final coupled HMM models, indirect effects from treatment variables to QOL through drinking (and vice versa) were estimated using RMediation (Tofighi and MacKinnon, 2011). As described in Tofighi and MacKinnon, 2011, RMediation uses the product of coefficients method for estimating indirect effects. Using this method, the standard error for any two-path indirect effect can be calculated from the beta and standard error for the coefficient from the predictor to the mediator ("a"), the beta and standard error for the coefficient from the mediator to the outcome ("b"), and the covariance between "a" and "b" estimates. These parameters are adjusted for all relevant effects, as betas and standard errors typically are in structural equation models. The statistical significance of indirect effects was evaluated using both 95% and 90% confidence intervals.

RESULTS

Baseline characteristics and zero-order associations between SF-12 and drinking data in the COMBINE study have been detailed elsewhere (Anton et al., 2006; LoCastro et al., 2009).

Latent profile analyses

As can be seen in Table 1, LMR-LRT and BIC results favored a four-state model for PHDD at baseline and week 16. However, this model could not be estimated in the week 52 data.

Because a three-state model converged at each assessment occasion and because there is prior support for a three-state model of heavy alcohol consumption in the literature (Witkiewitz et al., 2010; Shirley et al, 2010), we chose to pursue the three-state model for HMM analyses. As can also be seen in Table 1, although a three-state solution was supported for SF-12 physical at all three measurement occasions, one of the latent states in this solution had an unacceptably low estimated prevalence in the data (3% for weeks 16 and 52, 6% for baseline). As such, we chose to pursue a two-state model of SF-12 physical. Finally, a two-state solution was supported for SF-12 mental health at baseline and week 16; a three-state solution was favored at week 52. Because the majority of measurement occasions supported a two-state solution, we chose to pursue a two-state model of SF-12 mental health.

Hidden Markov modeling

All univariate HMMs converged to the same admissible solutions for a variety of starting values, thus the models were well identified. For all three evaluated constructs, comparisons of nested models suggested that transition matrices should be freed (see Table 2, paths "a*" and "b*" in Figure 1). Thus, in coupled HMM models, autoregressive coefficients and latent state intercepts were freely estimated across measurement occasions within each construct. Conversely, comparisons of nested models suggested that cross-lagged regression coefficients and treatment covariate effects should remain constrained in the two estimated coupled HMM models (PHDD and SF-12 mental health, PHDD and SF-12 physical; see Table 2, paths "c" through "f" in Figure 1).

Figure 2 displays results from the coupled HMM involving PHDD and SF-12 mental health. Model estimated means for PHDD were 3.08 (state 1, "low"), 40.26 (state 2, "moderate"), and 89.47 (state 3, "high," reference) at each measurement occasion. Estimated counts at each occasion were: occasion 1-state 1=40, state 2=502, state 3=683; occasion 2-state 1=953, state 2=158, state 3=115; occasion 3-state 1=786, state 2=204, state 3=236. Although each latent state contained at least 5% of the sample at each measurement occasion in LPA analyses, in HMM analyses, state 1/occasion 1 contained less than 5% of the sample. This occurred because HMM models determined state compositions based on data from all time points; low drinking states were common in the post-treatment phase, but were not common at baseline (participants were required to have a substantial amount of recent alcohol use at baseline to be included in the study). Given the large sample size of the COMBINE study, and the fact that the HMMs estimated latent class structural parameters using all multivariate data, we did not experience any estimation problems related to low class membership. Because PHDD was represented by three latent states, associations involving PHDD were estimated in a multinomial regression framework, with state 3 as the reference. Model implied means for SF-12 mental health were 0.24 (state 1, "above average") and −1.52 (state 2, "below average," reference). Estimated counts at each occasion were: occasion 1-state 1=519, state 2=707; occasion 2-state 1=970, state 2=256; occasion 3-state 1=911, state 2=315. As can be seen in Figure 2, most autoregressive paths were statistically significant, implying that current drinking state predicted drinking state at the following visit and that current mental health state predicted mental health state at the following visit. Bidirectional cross-lagged regression coefficients between PHDD and SF-12 mental health were statistically significant; the interpretation of the multinomial model is such that improved SF-12 predicted low vs. high PHDD (β=0.73, p=0.01) and low vs. high PHDD predicted improved SF-12 (β=0.71, p=0.04). Of all evaluated treatment covariates, only naltrexone significantly predicted transitional drinking behavior. Specifically, naltrexone was significantly associated with being in the low vs. the high PHDD group (OR=1.27, β=0.24, p=0.01), but was not significantly associated with being in the moderate vs. the high PHDD group (OR=0.97, β=−0.03, p=0.87). None of the treatment covariates significantly predicted

transitional SF-12 mental health scores. However, mediation analyses suggested that naltrexone indirectly predicted transitional SF-12 mental health at a 90% confidence level (mean product of coefficients=0.24, CI 90%=0.02, 0.37). In other words, naltrexone predicted improvements in PHDD, which in turn predicted improvements in SF-12 mental health. No other indirect effects were significant at 90% or 95% confidence levels.

Figure 3 displays results from the coupled HMM involving PHDD and SF-12 physical. Model estimated means for PHDD were 3.10 (state 1, "low"), 40.38 (state 2, "moderate"), and 89.48 (state 3, "high," reference). Estimated counts at each measurement occasion were: occasion 1 - state 1=43, state 2=500, state 3=683; occasion 2 - state 1=954, state 2=157, state 3= 114; occasion 3 - state 1=789, state 2=202, state 3=236. As noted above, state 1/ occasion 1 contained less than 5% of the sample because low drinking states were common at post-treatment but not at baseline. Model implied means for SF-12 physical were 0.56 (state 1, "above average") and −1.01 (state 2, "below average," reference). Estimated counts at each occasion were: occasion 1-state 1=998, state 2=228; occasion 2-state 1=1,045, state 2=181; occasion 3-state 1=992, state 2=234. As can be seen in Figure 3, most autoregressive paths were statistically significant, implying that current drinking state predicted drinking state at the following visit and that current physical health state predicted physical health state at the following visit. Cross-lagged regression coefficients between PHDD and SF-12 physical were marginally significant for improvements in SF-12 predicting low vs. high PHDD (β=0.39, p=0.06), and were not significant for PHDD predicting SF-12 (β=−0.05, p=0.91). Again, only naltrexone significantly predicted transitional drinking behavior, and naltrexone was significantly associated with being in the low vs. the high PHDD state (OR=1.28, β=0.25, p=0.02), but was not significantly associated with being in the moderate vs. the high PHDD state (OR=0.96, β=−0.04, p=0.83). Additionally, in the current analysis, CBI in combination with acamprosate (OR=1.91, β=0.65, p=0.02) or naltrexone (OR=3.00, β=1.10, p=0.01) significantly predicted improvements in SF-12 physical scores. Mediation analyses suggested that interactions between CBI and acamprosate or naltrexone indirectly predicted transitional drinking behavior (low vs. high PHDD) at a 90% confidence level (CBI × acamprosate: mean product of coefficients=0.65, CI 90%=0.01, 0.58; CBI × naltrexone: mean product of coefficients=1.1, CI 90%=0.02, 1.00). No other indirect effects were significant at 90% or 95% confidence levels.

DISCUSSION

The present study investigated the impact of treatment for alcohol dependence on drinking behavior and health-related QOL using a novel, joint latent-variable modeling approach, coupled hidden Markov modeling, in the COMBINE study. Our results suggested that: (i) naltrexone significantly predicted decreased heavy drinking days, and marginally predicted improved mental health QOL through its effect on decreased heavy drinking days, and (ii) that the combinations of naltrexone and CBI, and acamprosate and CBI, each predicted significantly improved physical QOL, and marginally predicted decreased heavy drinking days through their effects on improved physical QOL.

Although this pattern of results is consistent with findings from earlier COMBINE reports (Anton et al., 2006; LoCastro et al., 2009), there were also some notable points of divergence. Regarding consistencies, both the present study and earlier COMBINE reports found that naltrexone alone predicted decreased drinking behavior (Anton et al., 2006) and that the combination of naltrexone and CBI predicted improved physical QOL (LoCastro et al., 2009). Conversely, whereas earlier COMBINE reports found that the combination of CBI and naltrexone, as well as the combination of CBI and placebo, directly predicted decreased drinking behavior (Anton et al., 2006), the present study did not find support for a direct effect of CBI, neither in combination with medications nor in combination with

placebo, on drinking behavior. This discrepancy may be explained by our simultaneous modeling of drinking behavior and QOL along with our investigation of both direct and indirect effects of treatment on drinking behavior. Because earlier reports did not take QOL into account when examining the impact of treatments on drinking behavior, their findings of a direct effect of CBI on drinking behavior may have been an artifact of the unmodeled association between QOL and drinking behavior. In the present study, we found evidence for an effect of CBI, in combination with either naltrexone or acamprosate, on drinking behavior; however, this effect was not direct, it was mediated by physical QOL. To our knowledge, mediation via secondary outcomes was not considered in other COMBINE studies and is not general practice. As such, our findings may be interpreted not as contradicting earlier COMBINE reports, but instead as providing a potential explanation as to why CBI predicted decreased drinking behavior. Also worthy of note, the present study found that the combination of CBI and acamprosate predicted improved physical QOL, whereas earlier COMBINE reports did not find this effect (LoCastro et al., 2009). This discrepancy may be explained by the potentially increased statistical power afforded by the present study's latent-variable modeling approach and consideration of joint outcomes from which information for estimation of treatment effects is borrowed (Teixeira-Pinto, Normand et al., 2009; Teixeira-Pinto, Siddique et al., 2009). Regarding why treatment combinations impacted physical QOL in the present study, participants receiving combined treatments may have become more active and productive, which in turn may have decreased their heavy drinking (perhaps through an increase in mood or a decrease in available occasions for use). Another noteworthy discrepancy between the present study and past COMBINE reports is that earlier reports have found that the effects of treatment on drinking behavior dissipated over the course of the follow-up period (Anton et al., 2006), whereas we did not find evidence for differential effects of treatment on outcomes across assessment occasions. Finally, in the present study, most associations between heavy drinking and other variables were only significant for the contrast between "high" and "low" heavy drinking states. This pattern of findings (null coefficients involving moderate vs. high PHDD, but significant coefficients involving low vs. high PHDD) is not entirely surprising given that predictions of extreme contrasts (e.g., low vs. high PHDD) have greater statistical power than predictions of moderate contrasts. However, given the large sample size of the COMBINE study, this pattern of findings suggests that treatments evaluated in the COMBINE study may lack sensitivity for moderate outcomes. Together, the results from the present study provide further insight into the impact of pharmacological and behavioral interventions on drinking behavior and quality of life in alcohol-dependent patients.

The present study is one of several recent articles arguing for additional analyses of clinical trials data in alcoholism research (Gueorguieva et al., 2010; Witkiewitz et al., 2010). As researchers agree that both drinking and quality of life measures are important markers of treatment efficacy, primary and secondary outcomes may be equally important and should be considered on equal footing using one of a variety of available joint modeling approaches. As noted earlier, the present study's coupled hidden Markov modeling approach, or alternatively other latent-variable and/or joint analysis techniques, may lead to improvements in our ability to detect genuine, significant treatment effects in clinical trials. They may also provide us with more accurate estimates of the impact of treatments on primary and secondary clinical outcomes, and may help us to elucidate the mechanisms through which treatments impact outcomes (via mediation analyses). Given the amount of time, energy, and money that goes into conducting clinical trials, it is important that we use our best available statistical technologies to analyze the data that result from those efforts.

The above discussion should be interpreted in light of the present study's limitations. First, the inclusion/exclusion criteria of the COMBINE study (e.g., 4–21 days of abstinence) limit the generalizability of our findings to the population of alcohol-dependent patients as a

whole. Second, although we found some support for statistical mediation, reported indirect effects were only significant at a 90% confidence level. Though the COMBINE study had a large sample size, statistical mediation has very high sample size requirements, particularly for small to medium-sized effects (Fritz and MacKinnon, 2007). Third, only three assessment occasions were utilized in the present analyses because QOL was administered on only three occasions in the COMBINE study. Future research should evaluate QOL, and other general functioning outcomes, at higher resolution so that more accurate estimates of treatment effects on longitudinal outcomes can be obtained. Fourth, HMM and mediation approaches are generally limited to large samples; even with large sample sizes, HMM models can be difficult to estimate due to their complexity. Finally, Mplus does not presently allow for the simultaneous consideration of contemporaneous and cross-lagged associations among latent categorical variables. Because the coupled HMM model, developed by Brand and colleagues (1996), does not contain contemporaneous associations among latent categorical variables, this was not a major concern for the present study.

These limitations notwithstanding, the present study illustrates a novel and potentially powerful statistical approach for simultaneously evaluating the impact of treatments on primary and secondary outcomes in clinical trials. This approach may overcome limitations of traditional statistical techniques, and in doing so, may provide enhanced accuracy and power to detect and estimate treatment effects in clinical trials. As such, we encourage clinical trials researchers to consider employing HMM, statistical mediation, joint-modeling, and other related methodologies in the future. Substantively, the present study suggests that behavioral interventions may impact drinking behavior through their ameliorative effects on general well-being. Together, these findings suggest that a combined naltrexone and behavioral intervention may provide the best available treatment for both drinking and quality of life in individuals with alcohol dependence.

Acknowledgments

REFERENCES

Anton RF, O'Malley SS, Ciraulo DA, Cisler RA, Couper D, Donovan DM, Gastfriend DR, Hosking JD, Johnson BA, LoCastro JS, Longabaugh R, Mason BJ, Mattson ME, Miller WR, Pettinati HM, Randall CL, Swift R, Weiss RD, Williams LD, Zweben A. Combined pharmacotherapies and behavioral interventions for alcohol dependence. The COMBINE Study: A randomized controlled trial. JAMA. 2006; 295:2003–2017. [PubMed: 16670409]

Brand, M.; Oliver, N.; Pentland, A. Coupled Hidden Markov Models for Complex Action Recognition; Proceedings of IEEE Conference on Computer Vision and Pattern Recognition; 1996. p. 994-999.

CDC. Measuring healthy days: population assessment of health-related quality of life. Atlanta, Georgia: US Department of Health and Human Services, CDC; 2000 Nov.

COMBINE Study Group. Testing combined pharmacotherapies and behavioral interventions in alcohol dependence (The COMBINE Study): a pilot feasibility study. Alcohol Clin Exp Res. 2003; 27:1123–1131. [PubMed: 12878918]

DeSantis SM, Bandyopadhyay D. Hidden Markov Models for zero-inflated Poisson counts with an application to substance use. Stat Med. 2011; 30:1678–1694. [PubMed: 21538455]

DeSantis SM, Bandyopadhyay D, Back SE, Brady KT. Laboratory stress- and cue-reactivity studies are associated with decreased substance use among drug-dependent individuals. Drug Alcohol Depend. 2009; 105:227–233. [PubMed: 19717245]

Donovan D, Mattson ME, Cisler RA, Longbaugh R, Zweben A. Quality of life as an outcome measure in alcoholism research. J Stud Alcohol. 2005; (Supplement No. 15):119–139.

First, M.; Spitzer, R.; Gibbon, M.; Williams, J. Structured Clinical Interview for DSM-IV Axis I Disorders, Research Version, Patient Edition. New York: New York State Psychiatric Institute; 1996.

Fritz MS, MacKinnon DP. Required sample size to detect the mediated effect. Psychol Sci. 2007; 18:233–239. [PubMed: 17444920]

Gueorguieva R, Wu R, Donovan D, Rounsaville BJ, Couper D, Krystal JH, O'Malley SS. Naltrexone and combined behavioral intervention effects on trajectories of drinking in the COMBINE study. Drug Alcohol Depen. 2010; 107:221–229.

Holmes TH, Anderson AL, Li SH, Elkashef AM. Advantages of joint modeling of component HIV risk behaviors and non-response: application to randomized trials in cocaine-dependent and methamphetamine-dependent populations. Front Psychiatry. 2011; 2:41. [PubMed: 21779253]

Kraemer KL, Maisto SA, Conigliaro J, McNeil M, Gordon AJ, Kelley ME. Decreased alcohol consumption in outpatient drinkers is associated with improved quality of life and fewer alcohol-related consequences. J Gen Intern Med. 2002; 17:382–386. [PubMed: 12047737]

Langeheine, R.; van de Pol, F. Latent Markov Chains. In: Hagenaars, JA.; McCutcheon, AL., editors. Applied latent class analysis. Cambridge: University Press; 2002. p. 304-341.

Lo Y, Mendell NR, Rubin DB. Testing the number of components in a normal mixture. Biometrika. 2001; 88:767–778.

LoCastro JS, Youngblood M, Cisler RA, Mattson ME, Zweben A, Anton RF, Donovan DM. Alcohol treatment effects on secondary nondrinking outcomes and quality of life: the COMBINE study. J Stud Alcohol Drugs. 2009; 70:186–196. [PubMed: 19261230]

MacDonald, IL.; Zucchini, W. Hidden Markov and Other Methods for Discrete-valued Time Series. New York: Chapman and Hall; 1997.

Miller WR, Del Boca FK. Measurement of drinking behavior using the Form 90 family of instruments. J Stud Alcohol. 1994; (Supplement No. 12):112–118.

Morgan MY, Landron F, Lehert P. for the New European Alcoholism Treatment Study Group. Improvement in quality of life after treatment for alcohol dependence with acamprosate and psychosocial support. Alcohol Clin Exp Res. 2004; 28:64–77. [PubMed: 14745303]

Morgan TJ, Morgenstern J, Blanchard KA, Labouvie E, Bux DA. Health-related quality of life for adults participating in outpatient substance abuse treatment. Amer J Addict. 2003; 12:198–210. [PubMed: 12851016]

Muthen, LK.; Muthen, BO. Mplus User's Guide. 6th ed.. Los Angeles, CA: Muthen & Muthen; 2011.

Okoro CA, Brewer RD, Naimi TS, Moriarty DG, Giles WH, Mokdad AH. Binge drinking and health-related quality of life: Do popular perceptions match reality? Amer J Prev Med. 2004; 26:230–233. [PubMed: 15026103]

Prisciandaro JJ, DeSantis SM, Chiuzan C, Brown DG, Brady KT, Tolliver BK. Impact of depressive symptoms on future alcohol use in patients with co-occurring bipolar disorder and alcohol dependence: a prospective analysis in an 8-week randomized controlled trial of acamprosate. Alcohol Clin Exp Res. In press.

Schwarz G. Estimating the dimension of a model. Ann Stat. 1978; 6:461–464.

Sclove L. Application of model-selection criteria to some problems in multivariate analysis. Psychometrika. 1987; 52:333–343.

Shirley KE, Small DS, Lynch KG, Maisto SA, Oslin DW. Hidden Markov models for alcoholism treatment trial data. Ann Appl Stat. 2010; 4:366–395.

Teixeira-Pinto A, Normad SLT. Correlated bivariate continuous and binary outcomes: issues and applications. Stat Med. 2009; 28:1753–1773. [PubMed: 19358234]

Teixeira-Pinto A, Siddique J, Gibbons RD, Normand SLT. Statistical approaches to modeling multiple outcomes in psychiatric studies. Psychiat Ann. 2009; 39:729–735.

Tofighi D, MacKinnon DP. RMediation: an R package for mediation analysis confidence intervals. Behav Res. 2011; 43:692–700.

Turner-Bowker, DM.; Bartley, PJ.; Ware, JE, Jr. SF-36® Health Survey & "SF" Bibliography: Third Edition (1988–2000). Lincoln, RI: Quality Metric Incorporated; 2002.

Wall MM, Li R. Multiple indicator hidden Markov model with an application to medical utilization data. Stat Med. 2009; 28:293–310. [PubMed: 18991318]

Ware, JE. Research Society on Alcoholism. Santa Barbara, CA: 1999 Jul 1. Quality of Life.

Ware, JE.; Kosinski, M.; Keller, SK. SF-36® Physical and Mental Health Summary Scales: a User's Manual. Boston, MA: The Health Institute; 1994.

Ware JE, Kosinski M, Keller SD. A 12-Item Short-Form Health Survey: construction of scales and preliminary tests of reliability and validity. Medical Care. 1996; 34:220–233. [PubMed: 8628042]

Witkiewitz K, Maisto SA, Donovan DM. A comparison of methods for estimating change in drinking following alcohol treatment. Alcohol Clin Exp Res. 2010; 34:2116–2125. [PubMed: 20860611]
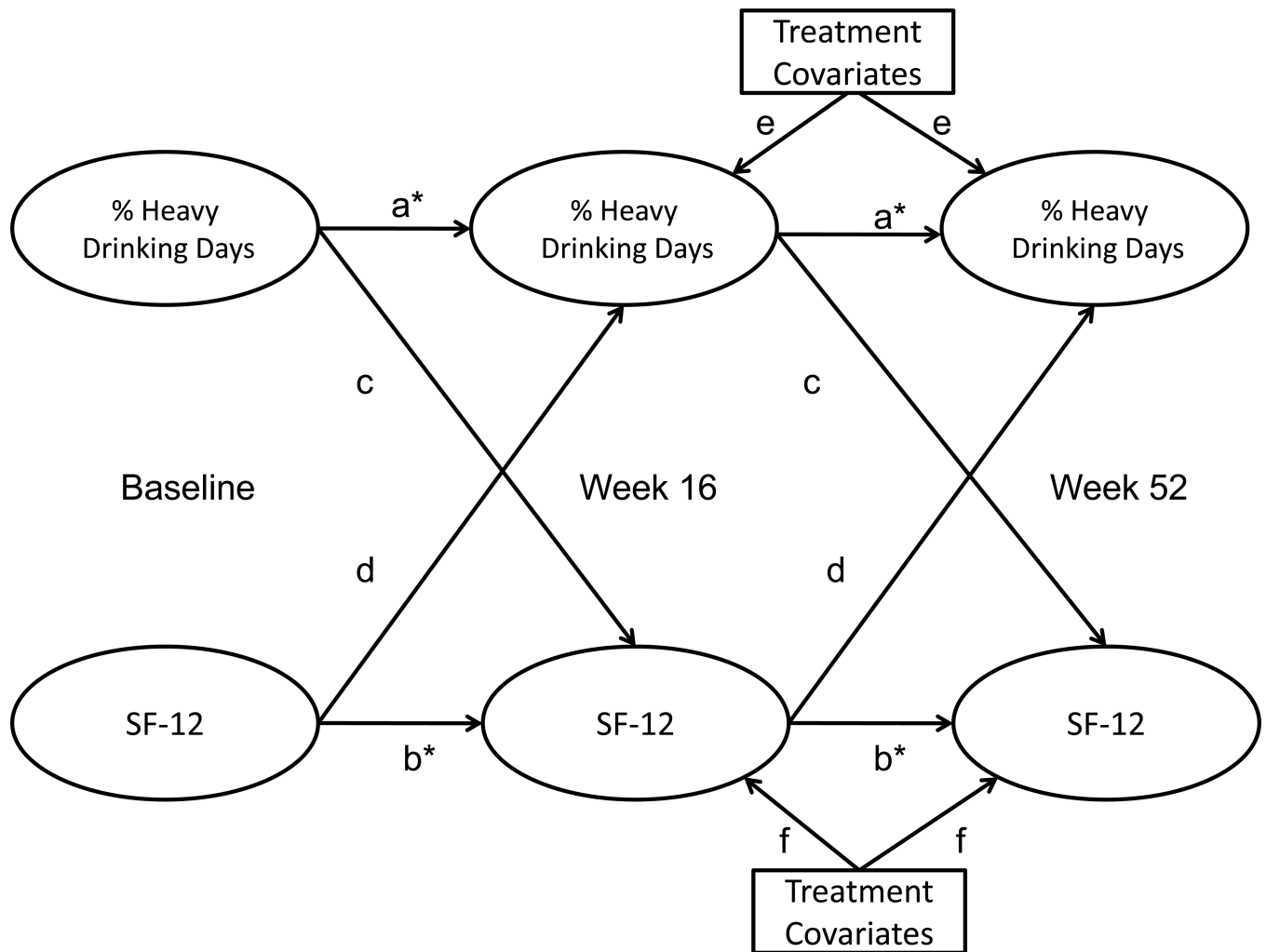
Figure 1.
Simplified model schematic of coupled hidden Markov models. Two coupled HMMs were estimated: % heavy drinking days and SF-12 mental health, and % heavy drinking days and SF-12 physical. For each model, paths denoted by the same letters were constrained to be equal, and constraints were tested via likelihood ratio tests (LRT) and Bayesian Information Criteria (BIC). Starred paths (a, b) were freed following these tests (Table 2). Although constrained paths between treatment covariates and week 16 and 52 outcomes are represented by two letters (one for % heavy drinking days and one for SF-12), each treatment covariate (e.g., naltrexone, acamprosate) and each outcome were separately constrained to one another across assessment occasions.
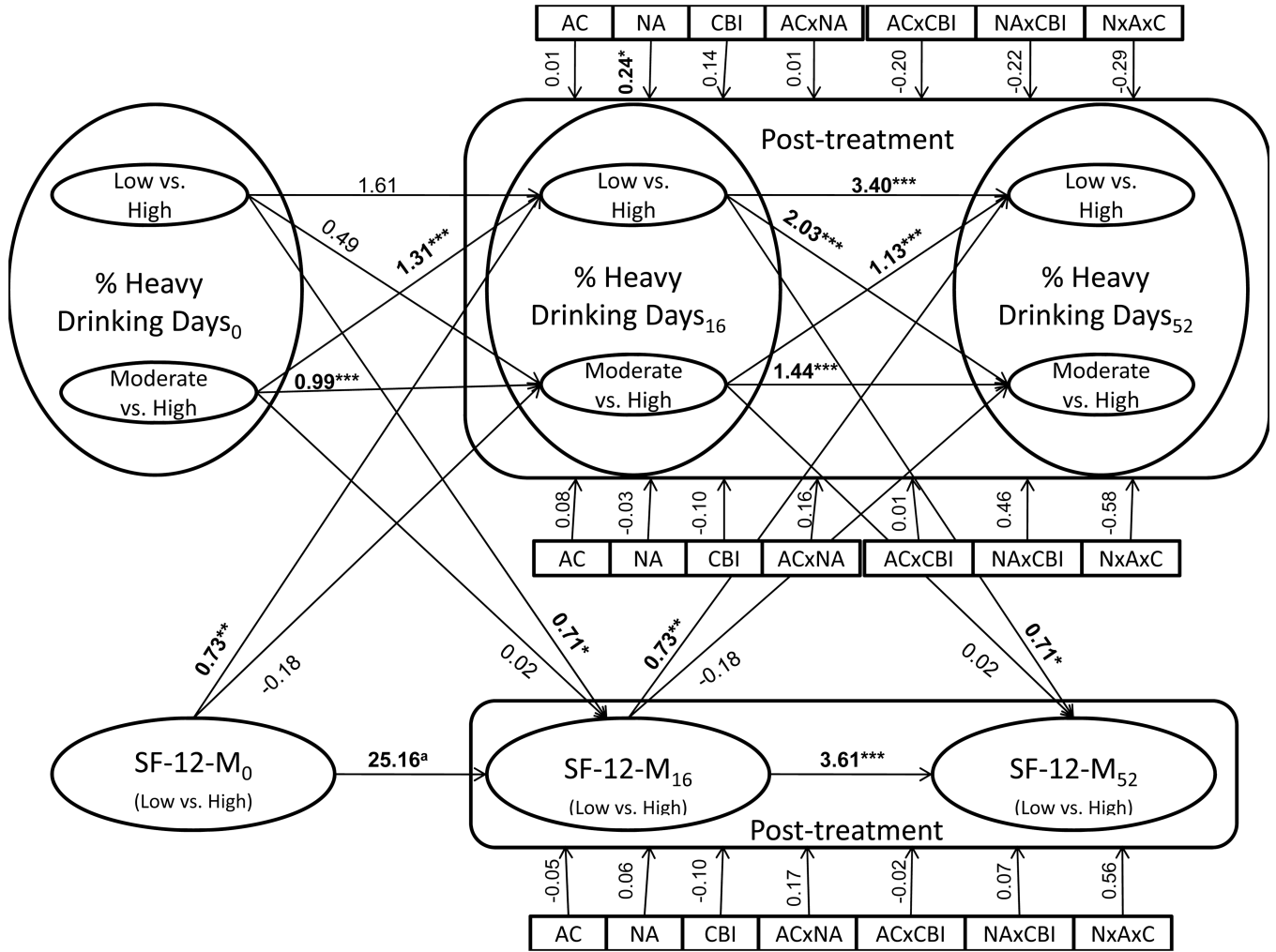
Figure 2.
Results from the coupled hidden Markov model involving % heavy drinking days (PHDD) and SF-12 mental health. Because PHDD was represented by 3 latent states, multinomial logistic regression (low vs. high, moderate vs. high) was used to estimate relationships between PHDD and other variables. Horizontal paths represent autoregressive parameters, diagonal paths represent cross-lagged coefficients, and vertical lines represent the impact of treatment coefficients on post-treatment outcomes. One coefficient is provided for each treatment covariate predicting each outcome because these associations were constrained to be equal across assessment occasions (Table 2). Significant coefficients are in bold.
[a]This path was automatically constrained by Mplus because of empty cells in the joint distribution of the categorical latent variables reflecting a near-perfect stability of SF-12 mental health between baseline and week 16.
* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$
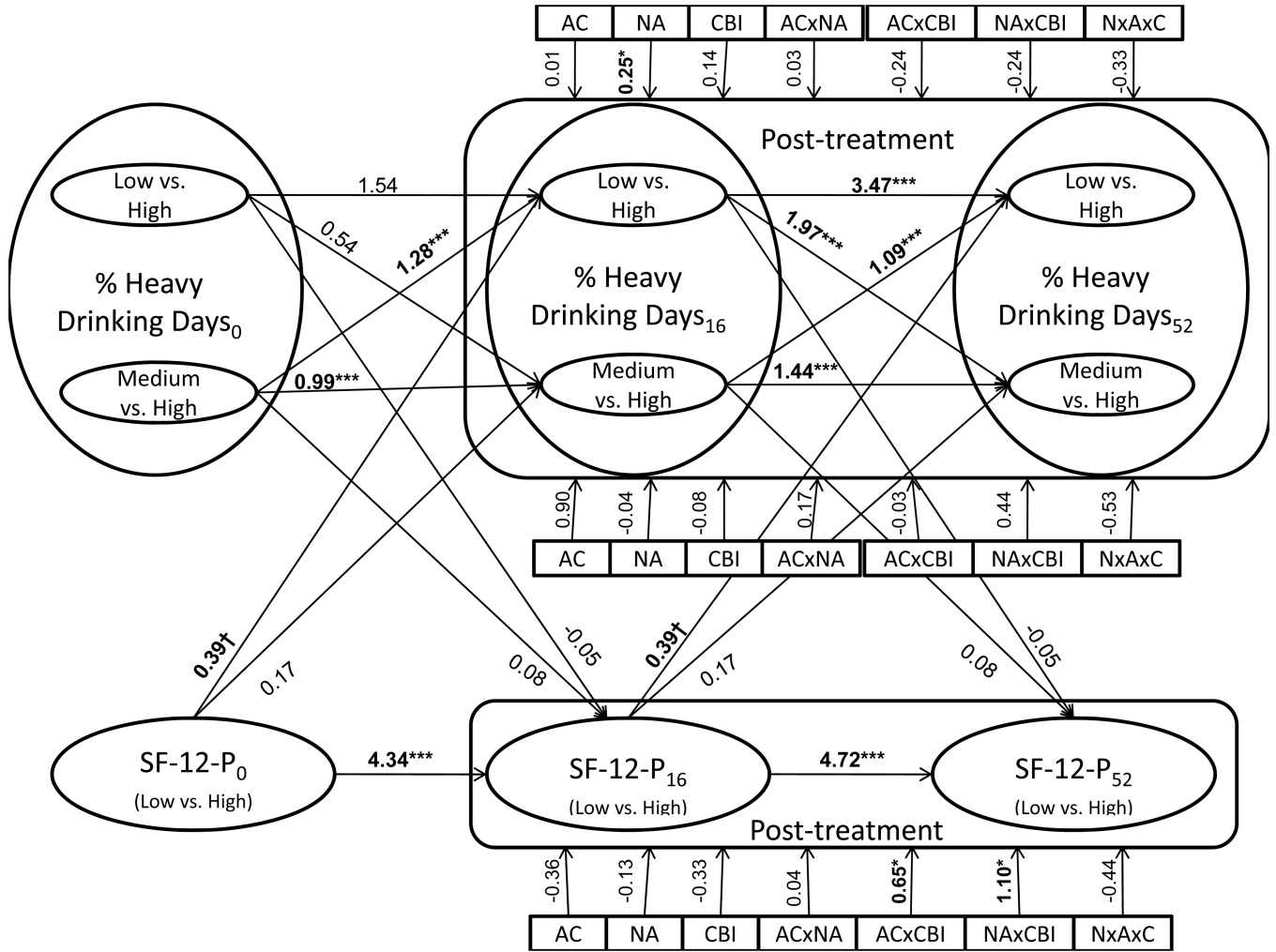
Figure 3.
Results from the coupled hidden Markov model involving % heavy drinking days (PHDD) and SF-12 physical. Because PHDD was represented by 3 latent states, multinomial logistic regression (low vs. high, moderate vs. high) was used to estimate relationships between PHDD and other variables. Horizontal paths represent autoregressive parameters, diagonal paths represent cross-lagged coefficients, and vertical lines represent the impact of treatment coefficients on post-treatment outcomes. One coefficient is provided for each treatment covariate predicting each outcome because these associations were constrained to be equal across assessment occasions (Table 2). Significant coefficients are in bold.
† $p < 0.10$, * $p < 0.05$, *** $p < 0.001$

Table 1

Fit Criteria for Latent Profile Models

Variable	# states	LMR LRT	p	BIC	n Adj BIC
% Heavy Drinking Days					
Baseline	2	292.36	<0.001	12912.83	12900.13
	3	81.50	0.022	12840.16	12821.10
	4	75.45	0.015	12773.96	12748.55
Week 16	2	1006.49	<0.001	12704.52	12672.76
	3	638.40	<0.001	11292.84	11280.14
	4	259.20	0.038	10624.22	10605.16
Week 52	2	850.88	<0.001	10361.27	10335.85
	3	501.30	<0.001	10069.37	10037.61
	4[a]	0	0.500	10927.08	10914.37
SF-12 Mental Health					
Baseline	2	43.33	0.014	4063.14	4050.44
	3	25.28	0.201	4050.51	4031.45
Week 16	2	202.20	<0.001	2842.29	2829.59
	3	57.13	0.087	2795.08	2776.03
Week 52	2	117.45	<0.001	2645.87	2633.17
	3	51.37	<0.001	2604.45	2585.39
SF-12 Physical					
Baseline	2	207.74	<0.001	3127.40	3114.70
	3	45.83	0.017	3092.79	3073.73
	4	15.20	0.092	3090.94	3065.53
Week 16	2	287.71	<0.001	2038.40	2025.69
	3	71.06	0.014	1976.26	1957.21
	4	20.39	0.175	1968.42	1943.01
Week 52	2	221.48	0.011	2174.70	2162.00
	3	96.72	<0.001	2084.62	2065.56
	4	30.57	0.150	2065.52	2040.11

Note. The Lo-Mendell-Rubin (LMR) Likelihood Ratio Test (LRT) provides a test of the difference in fit between n and $n-1$ states. A significant LMR LRT suggests that an n state solution is preferred over an $n-1$ state solution.

[a]The 4-state Week 52 % Heavy Drinking Days model did not converge.

BIC = Bayesian Information Criterion; *n* Adj = sample size adjusted; SF-12 = Medical Outcomes Study Health Survey, Short Form-12 version 2.

Table 2

Tests of Model Constraints in Hidden Markov Models

Variable(s) in HMM		Constraints	df	LRT	BIC H0	BIC H1
PHDD		Transition matrices	6	321.46*	33985.85	33712.82
SF-12 MH		Transition matrices	2	823.82*	9443.12	9365.31
SF-12 P		Transition matrices	2	13.19*	7009.34	7002.02
PHDD	SF-12 MH	Cross-lagged coeff	4	1.54	43063.91	43091.62
PHDD	SF-12 P	Cross-lagged coeff	4	1.33	40740.11	40766.81
PHDD	SF-12 MH	Medication coeff	21	32.89	38311.52	38423.00
PHDD	SF-12 P	Medication coeff	21	33.14	36225.82	36338.28

Note. HMM = Hidden Markov model; df = degrees of freedom; LRT = scaled likelihood ratio test; BIC = Bayesian Information Criterion; H0 = constrained model; H1 = unconstrained model; PHDD = percent heavy drinking days; SF-12 = Medical Outcomes Study Health Survey, Short Form-12 version 2; MH = mental health; P = physical; coeff = regression coefficients.

*
$p < 0.01$