Published in final edited form as:

J Subst Abuse Treat. 2019 May ; 100: 1–7. doi:10.1016/j.jsat.2019.02.001.

Cost-Effectiveness of *Individual* versus *Group* Female-Specific Cognitive Behavioral Therapy for Alcohol Use Disorder

Todd A. Olmstead¹, Fiona S. Graff², Alyssa Ames-Sikora³, Barbara S. McCrady^{4,5}, Ayorkor Gaba^{4,6}, and Elizabeth E. Epstein^{4,6}

¹Lyndon B. Johnson School of Public Affairs, The University of Texas at Austin, 2300 Red River Street, Austin, TX 78713

²War Related Illness and Injury Study Center, VA NJ Healthcare System, 385 Tremont Ave., East Orange, NJ 07018

³Department of Psychology, Georgia State University, 140 Decatur Street SE, Atlanta, GA 30302

⁴Center of Alcohol Studies, Rutgers University, 607 Allison Road, Piscataway, NJ 08854

⁵Center on Alcoholism, Substance Abuse, and Addictions, University of New Mexico, 2650 Yale Blvd. SE, Albuquerque, NM 87106

⁶Department of Psychiatry, University of Massachusetts Medical School, 365 Plantation Street, Worcester, Massachusetts, 01605

Abstract

Objective: To determine the relative cost-effectiveness of individual female-specific cognitive behavioral therapy (I-FS-CBT) versus group female-specific cognitive behavioral therapy (G-FS-CBT).

Methods: This cost-effectiveness study is based on a randomized controlled trial in which 155 women seeking treatment for alcohol use disorder at an academic outpatient clinic were randomized to 12 manual-guided sessions of I-FS-CBT (n = 75) or G-FS-CBT (n = 80). The

Corresponding: Todd A. Olmstead, PhD, Lyndon B. Johnson School of Public Affairs, The University of Texas at Austin, 2300 Red River Street, Austin, TX 78713, tolmstead@austin.utexas.edu, 512.471.8456.

Authors' current contact information:

Todd A. Olmstead, PhD, Lyndon B. Johnson School of Public Affairs, The University of Texas at Austin, 2300 Red River Street, Austin, TX 78713, tolmstead@austin.utexas.edu, 512.471.8456

Fiona S. Graff, Psy.D., War Related Illness and Injury Study Center, VA NJ Healthcare System, 385 Tremont Ave., East Orange, NJ 07018, fionagraff@gmail.com, 973.676.1000 x1537

Alyssa Ames-Sikora, MA, Department of Psychology, Georgia State University, 140 Decatur Street SE Atlanta, GA 30302, aames3@gsu.edu

Barbara S. McCrady, Ph.D., Center on Alcoholism, Substance Abuse, and Addictions (CASAA), University of New Mexico, 2650 Yale Blvd. SE, Albuquerque, NM 87106, bmccrady@unm.edu, 505.925.2388

Ayorkor Gaba, Psy.D, Department of Psychiatry, University of Massachusetts Medical School, 365 Plantation Street, Worcester, Massachusetts 01605, ayorkor.gaba@umassmed.edu, 732-309-8152

Elizabeth E. Epstein, Ph.D., Department of Psychiatry, University of Massachusetts Medical School, 365 Plantation Street, Worcester, Massachusetts, 01605., Elizabeth.Epstein@umassmed.edu, 732.995.7315

Publisher's Disclaimer: This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final citable form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

Declarations of Interest: None

primary patient outcomes were the number of drinking days and the number of heavy drinking days during the 12-week treatment and 1-year follow-up periods. All cost data (including resource utilizations) were collected prospectively alongside the trial. Incremental cost-effectiveness ratios and cost-effectiveness acceptability curves were used to determine the cost-effectiveness of I-FS-CBT relative to G-FS-CBT. Results are presented from the provider perspective.

Results: During the 12-week treatment period, G-FS-CBT is likely to be cost-effective when the threshold value to decision-makers of one fewer drinking day (or one fewer day of heavy drinking) is less than \$141 (or \$258), and I-FS-CBT is likely to be cost-effective if the threshold is greater than \$141 (or \$258). During the 1-year follow-up period, G-FS-CBT is likely to be cost-effective when the threshold value to decision-makers of one fewer drinking day (or one fewer day of heavy drinking) is less than \$54 (or \$169), and I-FS-CBT is likely to be cost-effective if the threshold is greater than \$54 (or \$169). The results are robust to sensitivity analyses on several key cost parameters.

Conclusions: Compared to I-FS-CBT, G-FS-CBT holds promise as a cost-effective approach, in both the short run and the long run, for improving drinking outcomes of women with alcohol use disorder.

Keywords

cost effectiveness; alcohol use disorder; women; female specific therapy; cognitive behavioral therapy

1. Introduction

Cost-effectiveness represents a frequently unexplored outcome in alcohol-use disorder (AUD) treatment research (NIAAA, 2000; Popovici & French, 2013), yet this line of inquiry is particularly relevant in a climate of considerable change in national healthcare policy. In clinical practice, information regarding the cost-effectiveness of AUD treatment is invaluable to programmatic decision-making and justification of resource allocation. An economic analysis is also essential to make findings from clinical research relevant and accessible to community clinics and treatment centers.

Group cognitive-behavioral therapy (CBT) may be less expensive than individual CBT, given that services can be provided to multiple patients simultaneously; however, few formal, systematic cost-effectiveness analyses exist and those that do have generated mixed findings across disorders (Tucker & Oei, 2007). For instance, researchers have speculated that group treatment for substance use disorder (SUD) is cost-effective compared to individual treatment based on comparable outcomes (Graham, Annis, Brett & Venesoen, 1996; Marques & Formigoni, 2001; Schmitz et al., 1997; Sobell, Sobell & Agrawal, 2009; Weiss, Jaffee, de Menil, & Cogley, 2004), but this speculation is based on favorable effectiveness/therapist time ratios rather than formal microcosting-based economic analyses. When differential effects are non-significant between group and individual conditions, an incremental cost-effectiveness analysis is still warranted because cost-effectiveness depends on the *joint* density of cost and effect differences, as opposed to *individual* differences in either cost or effect (Briggs & O'Brien, 2001; Drummond et al., 2005; Glick et al. 2015).

Formal examination of cost-effectiveness assumptions is important, given that a group format is the most prevalent treatment modality for alcohol and drug use disorders in clinical settings (Weiss et al., 2004).

The current study is a prospective economic analysis of individual versus group CBT designed specifically for women with AUDs. Development of female-specific treatment is supported by substantial gender differences in the development and course of AUDs (Epstein, McCrady, Hallgren, Cook, et al., 2018; Epstein & Menges, 2013; Mann et al., 2005; Nolen-Hoeksema, 2004; Roerecke & Rehm, 2013) as well as drinking trajectory and response to treatment (Abulseoud et al., 2013; Schneider, Kviz, Isola & Filstead, 1995; Zweig, McCrady & Epstein, 2009). Evidence suggests that women tend to remain in and experience more positive outcomes from treatment targeted specifically to their needs (Ashley, Marsden & Brady, 2003), and among women with AUDs, social support may be an important mechanism of change (McCrady et al., 2009). To this end, Epstein, McCrady, Hallgren, Gaba, et al. (2018) developed and compared group female-specific CBT (G-FS-CBT) to individual female-specific CBT (I-FS-CBT) and found that both types of treatment were associated with significant positive changes in drinking during treatment that were sustained during follow-up, with no significant difference in effects between treatment conditions. As noted above, a formal cost-effectiveness analysis of the Epstein, McCrady, Hallgren, Gaba, et al. trial is warranted to help policy- and decision-makers decide which modality to invest in and implement to treat women with AUDs.

1.1 The Current Study

The goal of the present study was to conduct a cost-effectiveness analysis of group versus individual FS-CBT for treating women with AUDs. We computed incremental cost-effectiveness ratios (Drummond et al., 2005; Gold et al., 1996) and cost-effectiveness acceptability curves (Fenwick, Claxton, & Schulpher, 2001) that define the range of values over which each intervention would likely be considered cost-effective for improving each of two patient outcomes measured during both the treatment and follow-up periods: (a) the number of drinking days (which relates to the frequency of drinking) and (b) the number of heavy drinking days (which relates to the intensity of drinking). We also checked the robustness of our cost-effectiveness results by conducting sensitivity analyses on several key cost parameters that would probably change had the trial been implemented under alternative realistic conditions. In addition to being one of the few economic analyses of a gender-specific treatment for either AUD or SUD (French et al., 2002; Ruger & Lazar, 2012; Svikis et al., 1997), this study adds to the general literature on the cost-effectiveness of individual versus group CBT interventions and specifically for AUD or SUD treatments (Holder et al., 1991; Sobell et al., 2009; Tucker & Oei, 2007).

2. Methods

Methods and results of the main trial are described in the main report (Epstein, McCrady, Hallgren, Gaba, et al., 2018). The study design and main outcomes are summarized briefly below, followed by the methods used for the cost-effectiveness analysis. Outcome and

resource utilization data for these analyses are taken from the main trial and combined with cost data obtained prospectively from study personnel.

2.1 Description of Main Trial

Participants were 155 women enrolled in a randomized controlled trial of 12 manual-guided sessions of individual (I-FS-CBT) or group (G-FS-CBT) female-specific cognitive behavioral therapy for AUD; 138 women attended at least one treatment session ($n=73$ for I-FS-CBT; $n=65$ for G-FS-CBT). The provider setting was an outpatient clinic at the Rutgers Center of Alcohol Studies on the Rutgers University campus. Women were recruited from newspaper and social media advertisements and community outreach efforts. Initial screening was done by phone, then eligibility was established via an in-person clinical screen. In addition, a baseline interview assessed for recent alcohol use, psychosocial functioning, and the presence of psychiatric disorders. Inclusion criteria were age 18 years of age or older, DSM-IV TR (APA, 2000) current alcohol abuse or dependence disorder, and alcohol consumption within the past 30 days. Exclusion criteria included evidence of recent (past 6 months) psychotic symptoms or significant cognitive impairment, physiological dependence on a substance other than alcohol or nicotine, and concurrent treatment for an AUD. There were no differences between treatment conditions on baseline demographic, psychosocial functioning, or alcohol use variables. All participants met criteria for Alcohol Dependence and on average reported 58.7 (SD=27.4) drinking days and 51.3 (SD=28.4) heavy drinking days during the 90 days prior to the start of the first treatment session. Mean sample age was 49.0 ($range=19-75$), and the majority of women were Caucasian (87%). Women had on average 15.7 years of education ($SD=2.9$), and 56.5% were employed full or part-time; see Epstein, McCrady, Hallgren, Gaba, et al. (2018) for additional characteristics. All participants provided informed consent, and all procedures were approved by the site Institutional Review Board.

Treatment was manual-guided, and all therapy was delivered by masters' or doctoral-level clinicians with specialized training in AUD treatment. Both treatments were weekly outpatient and stand-alone (i.e., they were not provided in the context of another treatment program); individual therapy duration was 90 minutes for session 1 and 60 minutes for sessions 2–12; group therapy duration was 120 minutes for session 1 and 90 minutes for sessions 2–12. Therapists were cross-trained to deliver both treatment conditions. In addition to conducting therapy sessions and completing administrative tasks such as therapy preparation and note keeping, therapists participated in weekly clinical supervision meetings with two doctoral-level licensed psychologists and an addiction psychiatrist to discuss active cases. Telephone contact with patients was maintained for therapy reminder calls and, less frequently, for detoxification or psychiatrist referrals or in emergency situations.

Both treatments included coping skills training, alcohol psychoeducation, motivational enhancement, functional analysis, stimulus control, cognitive restructuring, and relapse prevention topics, as well as topics that are particularly salient among women drinkers such as social network support, assertiveness training, empowerment, self-care, and mood and anxiety management. The G-FS-CBT manual included material identical to I-FS-CBT, but the session organization was modified for a group format.

2.2 Cost-Effectiveness Analysis

2.2.1 Cost measures—The costs of the labor, space, and materials associated with implementing both interventions were calculated from the perspective of the provider and adjusted to 2017 US\$ using the Consumer Price Index (Bureau of Labor Statistics, 2018). All labor costs included fringe benefits and overhead. Research-specific costs (e.g., incentive payments for study participation) were excluded from the analysis, as they were not considered a clinical cost and were also consistent across treatment conditions.

Cost data were collected prospectively alongside the efficacy trial. Study personnel kept detailed daily diaries in which they logged both the duration (for labor costs) and location (for space costs) of all treatment-related activities, including patient assessments, therapy sessions, scheduling/reminder activities, and weekly meetings. Material costs (e.g., workbooks) were obtained from the parent trial. Diaries were collected from all staff and entered weekly by a designated cost-effectiveness coordinator to ensure timely review of data, and a portion of diaries were double-entered and double entry data were compared to ensure accuracy. Prior to providing any clinical services, all staff participated in an initial series of training sessions to ensure accurate and consistent coding, and periodic “booster” training sessions were provided to prevent drift. The daily diary standardized format and data collection procedures were developed, piloted via a trial cost-effectiveness analysis, and revised during the initial pilot phase of the study, thus ensuring that all costs were accounted for.

Although most items presented in the cost analysis are straight-forward, two require explanation. First, the cost of a participant who failed to attend a scheduled therapy session (i.e., a “no-show”) was assumed to be zero in the G-FS-CBT condition on the grounds that the full cost of that session was prorated over those subjects who did attend, while the corresponding cost of a “no-show” in the I-FS-CBT condition was assumed to be 25% of the full cost of that session on the grounds that the clinician could make productive use of his/her time once he/she realized the participant was not coming. Second, the administrative cost of clinical interactions (e.g., screens, scheduling/reminder calls) with participants who dropped out of the study after randomization but prior to attending at least one treatment session was prorated over the participants who were randomized and completed at least one session.

2.2.2 Effectiveness measures—Effectiveness in the present study was assessed in terms of (a) the number of drinking days (to assess the frequency of drinking), and (b) the number of heavy drinking days (to assess the intensity of drinking) during both the 12-week treatment and 1-year follow-up periods. During treatment, participants completed daily diary cards noting quantity and type of alcohol consumed, which were then reviewed during session and used as both a clinical and research tool. During follow-up, drinking data were collected via the timeline followback (TLFB) method, a reliable, valid clinical instrument in which participants use a calendar method to recall daily drinking (Sobell & Sobell, 1992).¹

¹Follow-up TLFB data substituted for missing daily diary card data, with daily diary card data being the primary data source to compute within-treatment drinking variables. Daily diary card data correlate significantly with retrospective TLFB data at end of treatment (McCrady, Epstein, & Hirsh, 1999).

Because the weeks required to complete the 12-session treatment protocol varied across participants due to the timing of holidays and makeup sessions, the effectiveness measures were normalized assuming a 90-day treatment period. For example, if a given participant experienced 50 drinking days during a 78-day treatment period, then her normalized effectiveness measure would be $50 \times (90/78) = 57.7$ drinking days. Similarly, if another participant experienced 50 drinking days during a 100-day treatment period, then her normalized effectiveness measure would be $50 \times (90/100) = 45$ drinking days. Because the actual length of the planned 1-year follow-up period varied slightly across participants (339 days to 380 days), the follow-up effectiveness measures were normalized assuming a 365-day follow-up period.

Of the 138 participants in the main trial who attended at least one treatment session, five were missing effectiveness data during the treatment period ($n = 1$ for I-FS-CBT; $n = 4$ for G-FS-CBT) and 19 were missing effectiveness data during the follow-up period ($n = 9$ for I-FS-CBT; $n = 10$ for G-FS-CBT). Missing effectiveness data were imputed following Briggs et al. (2002) in which standard regression analysis is used to provide estimates of the missing data conditional on complete variables in the analysis. Specifically, predictive equations were obtained from regressing drinking outcomes on a set of demographic characteristics (age, race, years of education, marital status, employment status, household income, any Axis I disorder, any Axis II disorder), pre-treatment drinking behavior, and the number of treatment sessions attended. Predictive equations from the regressions had R^2 values ranging from .20 (number of heavy drinking days during follow-up) to .46 (number of drinking days during treatment).

2.2.3. Incremental cost-effectiveness analysis—Incremental cost-effectiveness analysis is the appropriate approach to use in this study because I-FS-CBT is expected to be costlier than G-FS-CBT on a per-patient basis (i.e., the cost of a therapy session in I-FS-CBT is borne entirely by the individual patient, while the cost of a therapy session in G-FS-CBT is pro-rated over all attending group members). For each of the effectiveness measures, we calculated incremental cost-effectiveness ratios (ICERs). Following standard economic theory (Drummond et al., 2005; Gold et al., 1996), ICERs are defined as the incremental cost of using I-FS-CBT, compared to GFS-CBT, to produce one fewer drinking day or one fewer day of heavy drinking (i.e., the incremental cost divided by the incremental effect). Economic theory then suggests that the cost-effective intervention is the one with the highest ICER that is less than the decision maker's willingness to pay for one fewer drinking day or one fewer day of heavy drinking (Drummond et al., 2005; Gold et al., 1996). Both incremental costs and incremental effects used to calculate the ICERs were obtained from the study as described above.

Cost-effectiveness acceptability curves (CEACs) are presented to illustrate the statistical uncertainty in the study due to the specific sample (Briggs, 2001; Fenwick, Claxton, & Schulpher, 2001). CEACs show the probability that each intervention would be cost-effective, given the observed data, under different assumptions about the value of one fewer drinking day or one fewer day of heavy drinking. Following Drummond et al., (2015, p.300), costs and effects for each intervention were bootstrapped together (with 2000 replicates) to produce the CEACs.

Finally, we checked the robustness of our cost-effectiveness results by conducting sensitivity analyses on several key cost parameters to assess how the ICERs would probably change had the trial been implemented under alternative realistic conditions.

3. Results

3.1 Cost

Table 1 shows the average cost (2017 US\$) per participant in each treatment arm, as well as the incremental cost of using I-FS-CBT compared to G-FS-CBT. As is typical with psychotherapies for AUDs and other SUDs, labor accounts for the vast majority of the average cost per participant in both interventions, while the average cost of space and materials is trivial (Olmstead et al., 2007; Olmstead et al., 2010). As expected, due to the ability to pro-rate therapy costs over multiple participants in the group condition, the per participant cost of therapy-related activities was significantly lower for G-FS-CBT than for I-FS-CBT (\$304 vs. \$858; p -value < .001). Overall, the incremental cost of using I-FS-CBT rather than G-FS-CBT was \$594 per participant (p -value < .001).

3.2 Effectiveness

The main trial found no statistically significant differences by treatment condition on any of the participant demographic, psychosocial functioning, or alcohol use variables measured at baseline. During the 12-week treatment period, the mean number of drinking days was 23.37 (SD = 23.96) and 27.58 (SD = 25.82) among women assigned to I-FS-CBT and G-FS-CBT, respectively, while the mean number of heavy drinking days was 14.28 (SD = 21.83) and 16.58 (SD = 19.77) among women assigned to I-FS-CBT and G-FS-CBT, respectively; these differences were not statistically significant (p -value = 0.322 for number of drinking days during treatment; p -value = 0.520 for number of heavy drinking days during treatment).

During the 1-year follow-up period, the average number of drinking days was 113.43 (SD = 111.49) and 124.45 (SD = 112.63) among women assigned to I-FS-CBT and G-FS-CBT, respectively, while the average number of heavy drinking days was 71.04 (SD = 94.57) and 74.55 (SD = 89.10) among women assigned to I-FS-CBT and G-FS-CBT, respectively; these differences were not statistically significant (p -value = 0.565 for number of drinking days during follow-up; p -value = 0.824 for number of heavy drinking days during follow-up).

As noted in the Introduction, although incremental effects were non-significant between the two conditions for each outcome during both time periods, incremental cost-effectiveness is still warranted.²

²A common misperception when differential effects are non-significant between two treatments is that the cheaper treatment is cost-effective (i.e., cost-minimization analysis). However, “unless a study has been specifically designed to show the equivalence of treatments (in terms of costs or effects), it would be inappropriate to conduct cost-minimization or outcome-maximization type analysis on the basis of an observed lack of significance in either the effect or cost differences between treatments” (Briggs & O’Brien, 2001). Moreover, “equivalence trials are rare because they require a much larger sample size than those designed to test for differences” (Briggs & O’Brien, 2001). Said differently, “absence of evidence of a difference” is not “evidence of absence of a difference” (Glick et al., 2015). That is, “a focus on hypothesis testing leads to an overemphasis on type I errors (the rejection of the null hypothesis of no difference when there is, in fact, no difference) at the expense of type II errors (the failure to reject the null hypothesis of no difference when in fact a difference does exist)” (Briggs & O’Brien, 2001).

3.3 Cost effectiveness

The ICERs were calculated using incremental costs from Table 1 and incremental effects as described above. Specifically, during the 12-week treatment period, the incremental cost of using I-FS-CBT, compared to G-FS-CBT, to obtain one fewer drinking day was \$141/day (i.e., $(\$1154 - \$560)/(27.58 \text{ days} - 23.37 \text{ days})$) and to obtain one fewer day of heavy drinking was \$258/day (i.e., $(\$1154 - \$560)/(16.58 \text{ days} - 14.28 \text{ days})$). During the 1-year follow-up period, the incremental cost of using I-FS-CBT, compared to G-FS-CBT, to obtain one fewer drinking day was \$54/day (i.e., $(\$1154 - \$560)/(124.45 \text{ days} - 113.43 \text{ days})$) and to obtain one fewer day of heavy drinking was \$169/day (i.e., $(\$1154 - \$560)/(74.55 \text{ days} - 71.04 \text{ days})$).

Figure 1 shows the CEACs associated with the number of drinking days and number of heavy drinking days during both the treatment and follow-up periods. Each CEAC shows the probability that IFS-CBT is cost-effective compared to G-FS-CBT given the observed data (Fenwick, Claxton, & Schulpher, 2001). Note that CEACs are a function of the threshold willingness-to-pay of the decision maker for one fewer unit of outcome. For example, as the decision maker's threshold willingness-to-pay for one fewer drinking day during the 12-week treatment period increases from \$50 to \$100 to \$500, Figure 1a shows that the probability that I-FS-CBT is cost-effective increases from 4% to 34% to 75%, respectively. Similarly, as the decision maker's threshold willingness-to-pay for one fewer heavy drinking day during the 12-week treatment period increases from \$50 to \$100 to \$500, Figure 1a shows that the probability that I-FS-CBT is cost-effective increases from 0% to 16% to 61%, respectively.

3.4 Sensitivity analyses

To determine how our cost-effectiveness results would likely change had the main trial been implemented under alternative realistic conditions, we conducted sensitivity analyses in which we considered two alternative scenarios – one favorable to I-FS-CBT and one unfavorable to I-FS-CBT – in which we made different assumptions about (1) the clinic overhead rate, (2) base wages of key personnel, and (3) the cost of a participant who failed to attend a scheduled therapy session (i.e., a “no-show”) in the I-FSCBT condition. Specifically, in the “favorable” scenario, the clinic overhead rate was assumed to be 20% (compared to 28% in the base scenario) to reflect potential variation in clinic setting and size, all base wages were reduced by 15% to reflect potential geographic variation in labor costs should the treatments be implemented elsewhere, and the cost of a “no-show” in the I-FS-CBT condition was assumed to be zero (compared to 25% of the full cost of that session in the base scenario). In the “unfavorable” scenario, the clinic overhead rate was assumed to be 35%, all base wages were increased by 15%, and the cost of a “no-show” in the I-FS-CBT condition was assumed to be 50% of the full cost of that session. Table 2 presents the results of these sensitivity analyses. For example, as shown in Table 2, during the 12-week treatment period the incremental cost of using I-FSCBT, compared to G-FSCBT, to obtain one fewer drinking day ranges from \$108/day in the favorable scenario to \$179/day in the unfavorable scenario, while the incremental cost of using I-FS-CBT, compared to G-FS-CBT, to obtain one fewer day of heavy drinking ranges from \$197/day in the favorable scenario to \$327/day in the unfavorable scenario.

Discussion

This study examined the relative cost-effectiveness from the provider perspective of two strategies for treating women with AUDs: *group* female-specific CBT and *individual* female-specific CBT. Which intervention (i.e., I-FS-CBT or G-FS-CBT) is likely to be cost-effective for improving alcohol use outcomes during the treatment and follow-up periods depends on the value that providers place on an additional unit of effect (i.e., one fewer drinking day, one fewer day of heavy drinking). Such valuations likely vary widely across providers as a function of a variety of factors, including the reimbursement structures (e.g., pay-for-performance, pay-for-patient-performance, fee-for-service) inherent in the payer mix of the patient populations served by each provider. In the absence of consensus threshold values across providers for alcohol use outcomes, we present ranges of values, defined by the ICERs and CEACs for both patient outcomes (number of drinking days and number of heavy drinking days) during both time periods (treatment and follow-up), over which each treatment is likely to be cost-effective compared to the other. Specifically, during the treatment period, we found that G-FS-CBT is likely to be cost-effective as long as the threshold value to providers of one fewer drinking day (or one fewer day of heavy drinking) is *less* than \$141 (or \$258), and I-FS-CBT is likely to be cost-effective if the threshold is *greater* than \$141 (or \$258). During the follow-up period, we found that G-FS-CBT is likely to be cost-effective as long as the threshold value to providers of one fewer drinking day (or one fewer day of heavy drinking) is *less* than \$54 (or \$169), and I-FS-CBT is likely to be cost-effective if the threshold is *greater* than \$54 (or \$169). Providers can use this information in combination with their own evaluation of the value of one fewer drinking day or one fewer day of heavy drinking during treatment and follow-up to make policy decisions.

The present study has several strengths. First, it is based on a randomized controlled trial. Second, all cost data were collected prospectively alongside the trial. Third, effectiveness was measured through outcomes of frequency and intensity during both treatment and follow-up periods. Fourth, sensitivity analyses supported our main findings. Finally, in the absence of consensus threshold values for one fewer drinking day or one fewer day of heavy drinking, we present ICERs and CEACs to define ranges over which G-FS-CBT and I-FS-CBT would likely be cost-effective from the provider perspective.

Several limitations warrant acknowledgement. First, the effectiveness trial did not include a “treatment as usual” condition or a no treatment control, so the present study is able to shed light only on the relative cost-effectiveness of I-FS-CBT vs. G-FS-CBT. However, in a previous study, I-FS-CBT and standard individual CBT for AUDs were shown to have comparable improvements in drinking behavior among women with AUDs (Epstein, McCrady, Hallgren, Cook, et al., 2018). Second, the economic analysis did not include the cost of training the clinicians for the study interventions. However, including training costs in the incremental cost-effectiveness analysis would likely have had little effect on findings because (a) training costs between the two interventions were similar and thus would be expected to net out of the incremental cost-effective analysis, and (b) in general, training costs are minimal when prorated over time on a per-participant basis (Olmstead et al. 2011). Third, the economic analysis relied on imputed missing effectiveness data for 3.6% and

13.8% of participants during the treatment and follow-up periods, respectively. Listwise deletion (i.e., excluding participants with missing data) produced ICERs and CEACs that are slightly more favorable to G-FS-CBT for both effectiveness measures during both periods (results not shown). Fourth, the absence of QALYs as a patient outcome measure limits the comparability of our findings. Fifth, inasmuch as all therapy in the effectiveness trial was delivered by masters' or doctoral-level clinicians in an academic outpatient clinic, it is unclear the extent to which these cost-effectiveness results would generalize to settings that rely on clinicians with less formal education. Finally, although the Second Panel on Cost-Effectiveness in Health and Medicine (Neumann, et. al, 2017) recommends adopting a healthcare-sector and a societal perspective alongside any other relevant perspective, the data requirements necessitated by both of these perspectives are beyond the scope of this study.

In conclusion, incremental cost-effectiveness ratios and cost-effectiveness acceptability curves were used to show that, compared to I-FS-CBT, G-FS-CBT holds promise from the provider perspective as a cost-effective approach during both the short run and the long run for improving drinking outcomes of women with alcohol use disorder. As such, it adds to the growing literature on the cost-effectiveness of treatments for AUDs by providing the first cost-effectiveness study of individual versus group modalities for women with AUDs as well as one of the few cost-effectiveness studies of gender-specific treatments for AUDs.

Acknowledgments

We thank Jean Shellhorn, Dr. Kevin Hallgren, Dr. Anthony Tobia, Dr. Mark Litt, Dr. Amy Cohn, Dr. Dorian Alaine, Sharon Cook, Noelle Jensen, and the research assistants, therapists, and students who worked on this project.

Funding/Support: This study was funded by the U.S. National Institutes of Health (R01 AA017163).

Role of the Funder/Sponsor: The funding source had no role in the design and conduct of the study; collection, management, analysis and interpretation of the data; preparation, review, or approval of the manuscript; and decision to submit the manuscript for publication.

References

- Abulseoud OA, Karpyak VM, Schneekloth T, Hall-Flavin DK, Loukianova LL, Geske JR, & Frye MA (2013). A retrospective study of gender differences in depressive symptoms and risk of relapse in patients with alcohol dependence. *The American Journal on Addictions*, 22, 437–442. [PubMed: 23952888]
- APA. (2000). *Diagnostic and statistical manual of mental disorders*, fourth edition, text revision. Washington, DC: American Psychiatric Association.
- Ashley OS, Marsden ME, & Brady TM (2003). Effectiveness of substance abuse treatment programming for women: A review. *The American Journal of Drug and Alcohol Abuse*, 29, 19–53. [PubMed: 12731680]
- Briggs A (2001). Handling uncertainty in economic evaluation and presenting the results In Drummond M, & McGuire A. (Eds.) *Economic evaluation in health care: Merging theory with practice* (pp.172–214). Oxford, UK: Oxford University Press.
- Briggs A, & O'Brien B (2001). The death of cost-minimization analysis? *Health Economics*, 10, 179–184. [PubMed: 11252048]
- Briggs A, Clark T, Wolstenholme J, & Clarke P (2002). Missing...presumed at random: cost-analysis of incomplete data. *Health Economics*, 12, 377–392.
- Bureau of Labor Statistics, US Department of Labor. Historical consumer price index for all urban consumers (CPI-U): U.S. city average, all items. <https://www.bls.gov/cpi/tables/supplemental-files/historical-cpi-u-201805.pdf> Accessed 19 June 2018.

- Drummond MF, O'Brien B, Stoddart GL, Torrance GW (2015). *Methods for the economic evaluation of health care programs*. 3rd ed Oxford, UK: Oxford University Press
- Epstein EE, McCrady BS, Hallgren KA, Cook S, Jensen NK, & Hildebrandt T (2018). A randomized trial of female-specific cognitive behavior therapy for alcohol dependent women. *Psychology of Addictive Behaviors*, 32(1), 1–15. [PubMed: 29154553]
- Epstein EE, McCrady BS, Hallgren KA, Gaba A, Cook S, Jensen N, Hildebrandt T, Holzhauer CG, & Litt MD (2018). Individual versus group female-specific cognitive behavior therapy for alcohol use disorder. *Journal of Substance Abuse Treatment*, 88, 27–43. [PubMed: 29606224]
- Epstein EE, & Menges D (2013). Women and addiction In McCrady BS & Epstein EE (Eds.), *Addictions: A comprehensive guidebook* (pp. 788–818). NY, NY: Oxford University Press.
- Fenwick E, Claxton K, & Schulpher M (2001). Representing uncertainty: The role of cost-effectiveness acceptability curves. *Health Economics*, 10, 779–787. [PubMed: 11747057]
- French MT, McCollister KE, Cacciola J, Durell J, & Stephens RL (2002). Benefit-cost analysis of addiction treatment in Arkansas: Specialty and standard residential programs for pregnant and parenting women. *Substance Abuse*, 23(1), 31–51. [PubMed: 12444359]
- Glick HA, Doshi JA, Sonnad SS, Polsky D (2015). *Economic evaluation in clinical trials* (pp. 206–208). 2nd ed Oxford, UK: Oxford University Press.
- Gold MR, Siegel JE, Russell LB, Weinstein MC (1996) *Cost-effectiveness in health and medicine*. Oxford, UK: Oxford University Press.
- Graham K, Annis HM, Brett PJ & Venesoen P (1996), A controlled field trial of group versus individual cognitive-behavioural training for relapse prevention. *Addiction*, 91, 1127–1140. [PubMed: 8828241]
- Holder H, Longabaugh R, Miller WR, & Rubonis AV (1991). The cost effectiveness of treatment for alcoholism: a first approximation. *Journal of Studies on Alcohol*, 52(6), 517–540. [PubMed: 1661799]
- Mann K, Ackermann K, Croissant B, Mundle G, Nakovics H, & Diehl A (2005). Neuroimaging of gender differences in alcohol dependence: Are women more vulnerable? *Alcoholism, Clinical and Experimental Research*, 29(5), 896–901.
- Marques AC & Formigoni ML (2001). Comparison of individual and group cognitive-behavioral therapy for alcohol and/or drug-dependent patients. *Addiction*, 96, 835–846. [PubMed: 11399215]
- McCrady BS, Epstein EE & Hirsch LS (1999). Maintaining change after conjoint behavioral alcohol treatment for men: outcomes at 6 months. *Addiction*, 94, 1381–1396. [PubMed: 10615723]
- McCrady BS, Epstein EE, Cook S, Jensen N, & Hildebrandt T (2009). A randomized trial of individual and couple behavioral alcohol treatment for women. *Journal of Consulting and Clinical Psychology*, 77(2), 243. [PubMed: 19309184]
- NIAAA (2000). 10th Special report to the U.S. Congress on Alcohol and Health. Washington, DC: US Department of Health and Human Services.
- Neumann PJ, Sanders GD, Russell LB, Siegel JE, & Ganiats TG (2017) *Cost-effectiveness in health and medicine*. 2nd ed Oxford, UK: Oxford University Press.
- Nolen-Hoeksema S (2004). Gender differences in risk factors and consequences for alcohol use and problems. *Clinical Psychology Review*, 24(8), 981–1010. [PubMed: 15533281]
- Olmstead TA, Carroll K, Canning-Ball M, & Martino S (2011). Cost and cost-effectiveness of three strategies for training clinicians in Motivational Interviewing. *Drug and Alcohol Dependence*, 116, 195–202. [PubMed: 21277713]
- Olmstead TA, Ostrow C, & Carroll K (2010). Cost effectiveness of computer-assisted training in cognitive-behavioral therapy as an adjunct to standard care for addiction. *Drug and Alcohol Dependence*, 110, 200–207. [PubMed: 20392575]
- Olmstead TA, Sindelar JL, Easton CJ, & Carroll K (2007). The cost-effectiveness of four treatments for marijuana dependence. *Addiction*, 102(9), 1443–1453. [PubMed: 17645430]
- Popovici J, & French MT (2013). Economic evaluation of substance abuse interventions In McCrady BS & Epstein EE (Eds.). *Addictions: A Comprehensive Guidebook* (pp 882–899). New York: Oxford University Press.
- Roerecke M, & Rehm J (2013). Alcohol use disorders and mortality: a systemic review and meta-analysis. *Addiction*, 108(9), 1562–1578. [PubMed: 23627868]

- Ruger JP, Lazar CM. Economic evaluation of drug abuse treatment and HIV prevention programs in pregnant women: a systematic review. *Addict Behav.* 2012 1; 37(1):1–10. [PubMed: 21962429]
- Schmitz JM, Oswald LM, Jacks SD, Rustin T, Rhoades HM & Grabowski J (1997). Relapse prevention treatment for cocaine dependence: Group vs. individual format. *Addictive Behaviors*, 22(3), 405–418. [PubMed: 9183510]
- Schneider KM, Kviz FJ, Isola ML, & Filstead WJ (1995). Evaluating multiple outcomes and gender differences in alcoholism treatment. *Addictive Behaviors*, 20(1), 1–21. [PubMed: 7785474]
- Sobell LC, & Sobell MB (1992). Timeline follow-back: A technique for assessing self-reported alcohol consumption In Litten RZ, & Allen JP (Eds.), *Measuring alcohol consumption: Psychosocial and biochemical methods* (pp. 41–72). Totowa, NJ: Humana Press.
- Sobell LC, Sobell MB & Agrawal S, (2009). Randomized controlled trial of a cognitive-behavioral motivational intervention in a group versus individual format for substance use disorders. *Psychology of Addictive Behaviors*, 23(4), 672–683. [PubMed: 20025373]
- Svikis DS; Golden AS; Huggins GR; Pickens RW; McCaul ME; Velez ML; Rosendale CT; Brooner RK; Gazaway PM; Stitzer ML; Ball CE. Cost-effectiveness of treatment for drug-abusing pregnant women. *Drug Alcohol Depend* 1997 4 14;45(1–2):105–13. [PubMed: 9179512]
- Tucker M, & Oei T (2007). Is group more cost effective than individual cognitive behaviour therapy? The evidence is not solid yet. *Behavioural and Cognitive Psychotherapy*, 35(1), 77–91.
- Weiss RD, Jaffee WB, Menil de VP, & Cogley CB (2004). Group therapy for substance use disorders: What do we know? *Harvard Review of Psychiatry*, 12(6), 339–350. [PubMed: 15764469]
- Zweig RD, McCrady BS, & Epstein EE (2009). Investigation of the psychometric properties of the drinking patterns questionnaire. *Addictive Disorders & their Treatment*, 8(1), 39–51.

Highlights

- Economic evaluation of two female-specific (FS) strategies for treating women with AUD
- Group FS-CBT was likely to be cost effective compared to individual FS-CBT
- Results robust to sensitivity analyses on several key cost parameters
- Results robust to drinking outcomes during treatment and 1-year follow-up periods

Figure 1.

Cost-effectiveness acceptability curves showing the probability that I-FS-CBT is cost-effective, compared to G-FS-CBT, for 'number drinking days' and 'number of heavy drinking days' during the (a) 12-week treatment period, and (b) 1-year follow-up period. For a given threshold value, the probability that I-FS-CBT is cost-effective compared to G-FS-CBT is equivalent to the proportion of the 2000 bootstrapped replicates for which I-FS-CBT had the highest net benefit (Fenwick, Claxton, & Schulpher, 2001)

Table 1

Average Cost per Participant in Each Treatment Arm (2017 US\$)

Component	I-FS-CBT (n = 73)		G-FS-CBT (n = 65)		I-G	
	Mean (\$)	SD(\$)	Mean (\$)	SD(\$)	Mean (\$)	p-value ^a
Labor						
<u>Assessments</u>						
Telephone screen	11	4	9	6	2	.026
Clinical screen (in person)	102	41	88	52	14	.083
No-shows	3	10	1	4	2	.064
Other ^b	18	48	20	52	-2	.800
<i>Assessments subtotal</i>	<i>134</i>	<i>60</i>	<i>118</i>	<i>70</i>	<i>16</i>	<i>.149</i>
<u>Therapy</u>						
Preparation	259	152	87	48	172	<.001
Sessions	486	235	200	132	286	<.001
Notes (post session)	71	61	14	18	57	<.001
Phone contacts ^c	9	19	3	7	6	.010
No-shows	31	57	0	0	31	<.001
Other	2	14	0	2	2	.250
<i>Therapy subtotal</i>	<i>858</i>	<i>395</i>	<i>304</i>	<i>156</i>	<i>554</i>	<i><.001</i>
<u>Miscellaneous</u>						
Scheduling/reminders	17	25	4	10	13	<.001
Weekly meetings	29	13	27	11	3	.219
Dropouts before treatment ^d	63	–	63	–	0	-
<i>Miscellaneous subtotal</i>	<i>110</i>	<i>28</i>	<i>94</i>	<i>75</i>	<i>16</i>	<i><.001</i>
<i>Labor Subtotal</i>	<i>1102</i>	<i>406</i>	<i>516</i>	<i>163</i>	<i>586</i>	<i><.001</i>
Space	14	5	6	2	8	<.001
Materials	38	5	38	4	0	.676
TOTAL	1154	410	560	165	594	<.001

SD = standard deviation

^aP-values for treatment arm differences determined by t-tests.^bMiscellaneous administrative tasks associated with conducting the telephone and clinical screens, including preparation and post-assessment paperwork.^cTreatment-related phone contacts with participant, family member, or other treatment provider.^dLabor (e.g., telephone screens, clinical screens, scheduling/reminders) for participants who dropped out after randomization but prior to attending at least one treatment session.

Table 2

Incremental cost-effectiveness ratios – base case and sensitivity analyses (2017 US\$)

	Base Case ^a	Favorable scenario ^b	Unfavorable scenario ^c
12-Week Treatment Period			
Cost of one fewer drinking day (\$)	141	108	179
Cost of one fewer day of heavy drinking (\$)	258	197	327
1-Year Follow-Up Period			
Cost of one fewer drinking day (\$)	54	41	68
Cost of one fewer day of heavy drinking (\$)	169	129	214

^aBase case corresponds to actual implementation of the efficacy trial.

^bFavorable scenario assumes (1) clinic overhead rate = 20%, (2) all base wages reduced by 15%, and (3) the cost of a “no-show” in the I-FS-CBT condition is zero.

^cUnfavorable scenario assumes (1) clinic overhead rate = 35%, (2) all base wages increased by 15%, and (3) the cost of a “no-show” in the I-FS-CBT condition = 50% of the full cost of that session.