Routledge
Taylor & Francis Group

Evidence-Based Effectiveness of a Private Practice Intensive Outpatient Program With Dual Diagnosis Patients

EDWARD A. WISE

Mental Health Resources, PLLC, Memphis, Tennessee, USA

The purpose of this study is to demonstrate the effectiveness of a freestanding intensive outpatient program (IOP) in a private practice setting for the treatment of dual diagnosis substance-abusing patients. Pre- and post-treatment Symptom Checklist 90–Revised, Global Assessment of Functioning, and patient functional rating scales were analyzed. Reliable change indices and clinically significant change measures were utilized. Trajectories of change for depression and number of days substances were used were analyzed, based on weekly Brief Symptom Inventory and substance use data reported by patients. Client satisfaction was also assessed at the end of treatment. Although patients started treatment with psychiatric symptoms comparable to those found in inpatient settings and 56% presented with suicidal or homicidal ideation, all symptom scales, functional impairments, and number of days used were significantly reduced by the end of treatment. Effect size statistics, reliable change indices, and statistically significant results indicated that 56% to 74% of patients treated in this program may be expected to improve, depending on the stringency of the criteria utilized. It is clear that dual diagnosis substance-abusing patients can be safely and effectively treated in a private practice IOP setting.

KEYWORDS Outcomes, evaluation, treatment, program, comorbid, dual diagnosis, intensive outpatient, evidence-based, effectiveness

Fewer than 2% of intensive outpatient programs (IOPs) or partial hospital-ization programs are provided in private practice settings (Barry & Lefkovitz, 2006). Yet, it has been repeatedly demonstrated that psychiatric patients in acute distress, typified by suicidal ideation, can be safely and effectively treated in a multidisciplinary outpatient private practice setting (e.g., Wise, 2003a, 2003b, 2005). The adoption of such treatment models represent op-portunities for collaborative, multidisciplinary care and practice expansion. While this research has demonstrated that IOPs can be cost-effective and empirically validated with severely depressed, suicidal, commercially insured psychiatric patients in a multidisciplinary practice, there is no evidence that such a program, in a similar setting, could be effective with more complex dual diagnosis substance-abusing patients.

A recent review of controlled studies of interventions for individuals with co-occurring disorders (Drake, O'Neal, & Wallach, 2008), found only two IOP studies, and while both of these assessed substance use outcomes, neither studied mental health outcomes. Thus, this review declared that in-tensive outpatient, integrated treatment was an understudied level of care with this population. In another review of dual diagnosis studies, Tiet and Mausbach (2007) found only two psychosocial treatment studies related to depression and substance abuse. They noted that most studies failed to re-port on both substance use and psychiatric symptoms, utilized small sample sizes, and did not occur in real-world settings. Similarly, Hesse (2009) was only able to identify five clinical trials that studied depressed patients and four that examined anxious patients in addition to their alcohol or other drug (AOD) use. These were primarily conducted in inpatient and partial hospitalization programs, but none were conducted in IOPs. On the other hand, Timko, Chen, Sempel, and Barnett (2006) demonstrated that dually diagnosed patients achieved significant cost savings when randomized into hospital or community care. The current author was unable to locate any effectiveness studies of interdisciplinary IOP private practice, dual diagno-sis models. There are numerous obstacles to providing this level of care to this complex group of patients in a private practice setting, not to mention the logistical problems associated with conducting naturalistic effectiveness studies, which may explain this lack of research.

In fact, the data that are available related to IOP outcomes with dual diagnosis patients tend to be exclusively focused on the severely men-tally ill, such as patients with schizophrenia (e.g., Drake, Mercer-McFadden, Mueser, McHugo, & Bond, 1998), as opposed to the severely depressed dually diagnosed patients with functional role impairments (e.g., Drake et al., 2008). Nonetheless, Granholm, Anthenelli, Monteiro, Sevcik, and Stoler (2003) demonstrated a significant reduction in hospital days following inte-grated dual diagnosis outpatient treatment. Recent data (Substance Abuse and Mental Health Services Administration, 2007) has also demonstrated that people with major depression were two to four times more likely to abuse

or be dependent on alcohol or drugs, respectively. This same study reported that only 8.5% of those with dual diagnoses received integrated treatment for both their mental and substance use problems. There is clearly a documented need for IOP outcome data, measuring both psychiatric symptoms and AOD use, on integrated dual diagnosis services delivered in private practice settings.

At a time when hospitalization days and reimbursement rates are closely managed, outpatient practices are well-positioned to care for more acute and complex patients if multidisciplinary treatment teams can be organized in office-based settings. A demonstration in which acute, dually diagnosed, complex patients could be effectively treated in private practice settings while averting costly hospitalization could represent a practice innovation, with the potential to expand multidisciplinary practice opportunities while providing needed services in office based settings.

METHODS

This study represents an attempt to measure the treatment effectiveness of an integrated dual diagnosis IOP, based on pre- and post-treatment client-rated symptom and functional measures, weekly symptom measures, clinician functional ratings, and client satisfaction in a naturalistic, private practice setting. This practice is not affiliated with any hospital and operates as a traditional, freestanding, multidisciplinary private practice, with the exception of providing IOP services.

Program Description

The dual diagnosis IOP consists of three hours of group treatment per day, 3 to 5 days per week, and utilizes two treatment modules that provide the content to be covered in each group. The term modules here is used to denote "an evidence-based approach to treatment that focuses on finding the common elements among standard treatment manuals and applying them according to a decision making process that accounts for pace, timing, or selection of techniques and is guided by client specific variables" (Borntrager, Chorpita, Higa-McMillan, & Weisz, 2009). The actual number of days attended is driven by phase of treatment (initial, middle, end), acuity, and safety-related issues. One of these treatment modules was previously used with a psychiatric IOP (Wise, 2003a, 2003b, 2005) and addresses coping skills designed to assess specific and typical skill deficits (e.g., cognitive behavior therapy for depression, assertiveness training, anxiety management). Each coping skill module is composed of specific content designed to be covered on a session-by-session basis. The second substance abuse treatment

module consists of similarly organized session content related to alcohol and drug education, stages of change, motivational enhancement, relapse prevention, effective alternatives, developing social support, etc., delivered in a motivational interviewing (MI) framework (Miller, 2004a, 2004b).

In addition to these educational groups, these patients were also treated in a traditional process group in which repetitive relationship problems were identified and addressed. Each of the three group sessions were approximately 45 to 50 minutes in duration. Although not subject to the rigors of a treatment manual per se (i.e., interrater reliability, fidelity adherence, etc.), these programming methods are structured and operationally function as treatment manuals tailored to specific client needs in accord with Borntrager et al. (2009). Similarly, the process groups are conducted in a structured format. Master's-level licensed therapists were trained in the delivery, process, and structure of each treatment module. They observed skilled therapists facilitating the group sessions, then were observed conducting the groups, and eventually were allowed to function independently. The treatment modules and training process provide for continuity of care and quality control and ensure that each patient will be exposed to a core set of coping skills.

Outcome Measurement

Patients were assessed at intake and discharge using pre- and post-treatment Symptom Checklist 90–Revised (SCL-90-R) scores, weekly Brief Symptom Inventory (BSI) scores, client and clinician functional ratings, and client satisfaction measures. (It should be noted that the SCL-90-R is the parent scale from which the BSI scales were derived.) Wise (2004) demonstrated that reliable change indices (RCIs) and clinically significant (CS) change variables may include symptom reduction, client satisfaction, level of functional improvement, and discharge status and that various *SD* units could be used to assess CS change variables along a continuum. Similarly, Tingey, Lambert, Burlingame, and Hansen (1996) developed normative continuum cutoffs. They further required that adjacent samples (i.e., severe, moderate, mild, asymptomatic) be statistically distinct. Wise (2003b) expanded the Tingey et al. (1996) SCL-90-R normative continuum by adding IOP data, representing a greater level of severity in more acutely distressed patients. The cutoffs previously reported (Wise, 2003b) were used in the present study to define CS change variables with the SCL-90-R Global Severity Index (GSI).

CS criteria have proven to be problematic in previous studies measuring substance abuse because normative data from various populations (community, outpatient, inpatient, etc.) are lacking and tend to be highly skewed, with relatively large *SD*s (e.g., see Cisler, Kowalchuk, Saunders, Zweben, & Trinh, 2005; Roberts, Neal, Kivlahan, Baer, & Marlatt, 2000). Jacobson and Truax (1991) assumed CS variables would be normally distributed and hence

movement toward normality could be easily and objectively quantified. Due to the highly skewed nature of this type of count data, Cisler et al. (2005) found that to meet the 95% confidence interval (CI) for percent days abstinent (PDA) as a CS variable, individuals had to achieve 92% PDA to be classified as functional. Similarly, Roberts et al. (2000) attempted to "control for the non-normal distribution" of their data by calculating percentiles and then performing sophisticated bootstrap methods to derive a reliable CS cut point. In the present case, where an MI and harm reduction philosophy are adopted, abstinence is not always a client's goal, so this criterion alone is insufficient to assess CS. However, a reduction in use is a crucial treatment variable that requires some method of assessment that is easily obtainable in a private practice setting.

Since PDA is a typical benchmark, it was adapted as a CS measure with a reduction in days used as indicative of movement towards normality. Each additional day of abstinence in the preceding week was assigned a proportional percentage to establish a continuum of days abstinent. Thus, an individual who increased PDA by 1 day at pre- and post-treatment would show a 16%, or 1 day, CS improvement. In order to obtain 100% change, an individual would have used 7 days in the week prior to admission and 0 days in the last week of treatment. Due to low frequency counts, data cells were collapsed, so that 2 to 3 days (3%–44%) and 4 to 6 days (58%–86%) of improved PDA were combined. While this method clearly underestimates the achievement of abstinence as a state, it does allow for the measurement of change along a normative continuum. This also allowed for a pre- and post-treatment of number of days used/abstinent to determine CS in a methodologically sound way that was consistent with our treatment philosophy. This is the first study the author could locate that used both substance reduction measured along a continuum as the CS variable and RCI to assess psychiatric symptoms to measure both reliable and CS change in dual diagnosis patients in a private practice outpatient setting.

In the previous studies related to our behavioral health IOP, clinician-rated global assessment of functioning (GAF) scores were assigned at pre- and post-treatment and consistently showed significant treatment improvements (e.g., Wise, 2003b, 2005). However, because functional capacities are typically not restored until the last phase of treatment (Howard, Lueger, Maling, & Martinovich, 1993), and in an attempt to further assess the validity of these clinician ratings, an 8-item, 5-point rating scale that was developed and previously used with the psychiatric IOP was also used in the current study. Hence, to supplement the clinician-rated pre- and post-treatment GAF scores, clients also "rate(d) the extent to which your problems interfere with or are a source of discomfort or concern to you" in eight specific functional domains of their life (e.g., job, family, social) (Wise, 2005).

This study also used a hierarchical linear model (HLM) to determine the trajectory of change as a result of IOP treatment with respect to both

psychological and substance abuse symptoms. Because HLM requires at least three measures during the course of treatment (Singer & Willett, 2003) it was necessary to build measures into clinical programming in a convenient way that allowed repeated measures that were valid and reliable. As part of the weekly treatment planning process, clients were given the BSI items comprising the Depression and Anxiety scales (Derogatis, 2001; Wise, 2005). Each patient also answered the items related to frequency of use from the Substance Use subscale of the Maudsley Addiction Profile self-completion form (MAP-sc) (Luty, Perry, Umoh, & Gormer, 2006). The Maudsley Addiction Profile (Marsden et al., 1998) is a public domain instrument that was adapted into a self-administered format and demonstrates acceptable psychometric properties (Luty et al., 2006).

MAP-sc Substance Use subscale items were incorporated into our weekly treatment plan and used to provide ongoing feedback to patients as well as to guide interventions. However, several alterations were necessary to adapt this instrument to our practice. Because we were interested in monitoring weekly changes, we altered the time frame from "month" to "week"; due to the infrequency with which it was anticipated that heroin users would be treated, instead of using the item related to the number of days heroin was used, we inserted "marijuana." Similarly, to ease respondent burden, the items related to cocaine, crack cocaine, benzodiazepines, amphetamines, methadone, hallucinogens, inhalants, or other drugs that we believed to be infrequent in our sample were combined into a single composite item for the number of days any of these drugs were used. Recent epidemiological data (Grant et al., 2004) confirm that alcohol and marijuana abuse are the most frequent substance use disorders, "far exceeding the rates of other drug-specific use disorders" (p. 115). Finally, each patient was also requested at intake to have his/her significant other complete a collateral questionnaire that contained the same MAP-sc items to obtain external ratings of substance abuse.

Hypotheses

It is hypothesized that the dual diagnosis patients will demonstrate psychopathology similar to inpatient SCL-90-R norms and demonstrate marked functional impairments across a variety of domains as measured by the client functional rating scale and GAF scores. It is hypothesized that significant treatment gains will be made with respect to SCL-90-R/BSI scales, particularly on the Depression and GSI scales, and show a significant decline during treatment. Additionally, it is believed that frequency of substance abuse will be significantly reduced and show a significant decline during treatment. It is expected that functional impairments will improve at the end of treatment as measured by the client functional rating scale and clinician-generated GAF scores. Finally, it is hypothesized that these consumers will be satisfied on

the Client Satisfaction Scale-8 (CSQ-8), as evidenced by an average score greater than 3.0.

Characteristics of the Treatment Sample

The sample consists of 100 consecutive admissions who had attended three or more IOP days. The group was 57% male, 77% Caucasian, and 23% African American, with an average age of 38.7 ($SD = 12.30$) years and an average education of 14.1 ($SD = 2.29$) years. Ninety-three percent of these patients were insured with managed care benefits, and the remaining 7% were self-pay. The average number of IOP days attended was 12.9 ($SD = 4.3$; range $= 4$–25) over approximately 5 weeks ($M = 4.9$; $SD = 1.75$, range 1–9), representing an average attendance rate of 80% of scheduled IOP days. The average age of first substance use was 16.2 ($SD = 3.87$) years. In 54% of the sample, clinicians diagnosed two Axis I disorders, and in 41% they diagnosed three or more Axis I disorders. The most frequent Axis I disorders were related to depression, alcohol abuse, anxiety, and marijuana abuse (74%, 55%, 27%, and 22%, respectively). These most prevalent diagnoses are consistent with those found in the most recent co-occurring substance use, mood, and anxiety disorders (Grant et al., 2004). Additionally, cocaine, opioid, and polysubstance abuse occurred in 13%, 10%, and 7% of the treatment sample, respectively. One hundred percent of the sample had a diagnosis of at least one psychiatric and one AOD abuse/dependence diagnosis; 56% of the sample presented with suicidal (52%), homicidal (1%), or suicidal and homicidal ideation (3%) at intake. Further, 32% of the sample had previous formal treatment for alcohol and/or drugs and 40% had previously attended Alcoholics Anonymous or Narcotics Anonymous. Fifteen had previous hospital or residential inpatient admissions for psychiatric and/or AOD treatment ($M = 1.4$; $SD = 0.88$). Of these, only two were step-downs into our IOP.

The remaining 98 patients were direct admissions. Sixty-seven percent of the sample was already receiving psychotropic medications from a prescriber not affiliated with the IOP prior to admission, most frequently an antidepressant (51%), anxiolytic (16%), or both (15%). All patients not receiving medication at the initiation of treatment ($n = 33$) were offered a referral to a psychiatrist, and 16 (51%) accepted the referral. Of these, all received prescriptions, most frequently for antidepressants (44%), anxiolytics (e.g., buspirone, hydroxyzine) (25%), or both (13%). Clinicians diagnosed an Axis II disorder in 74% and at least one Axis III disorder in 62%. Axis IV problems were primarily related to arrests for driving under the influence (DUI) (28%), and 51% had been arrested on other charges. Despite these levels of co-morbidity and acuity, only 4% of the sample was referred to any higher level of care during the course of their treatment. All treatment providers associated with each patient received routine coordination of care forms informing them that the patient had started IOP, Axis I through V diagnoses, a

brief clinical formulation, and a request for any relevant information. At the end of treatment, 73% of the study participants continued taking psychotropic medications. Patients were discharged based on numerous criteria, including self-reported use and symptom ratings; progress toward treatment goals; compliance and participation; and insurance authorizations and benefits.

There were eight patients (7%) who started the IOP but attended fewer than 3 IOP days and were considered dropouts. There were no significant differences between the treatment dropouts and those who remained beyond 3 sessions with respect to SCL-90-R symptom scores, number of days used, GAF scores, number of Axis I disorders, or 10 of 11 functional domains. The dropouts did report significantly more DUIs and prior Alcoholics Anonymous or Narcotics Anonymous attendance ($t(104) = 5.10$; $p < .0001$; $t(104) = 1.95$; $p = .05$, respectively).

RESULTS

Figure 1 shows that at intake, these patients were not significantly different than either the national psychiatric inpatient normative SCL-90-R group (Derogatis, 1994) nor 100 consecutive local psychiatric inpatient referrals (Wise, 2005) across 11 of 12 symptom measures. That is, as far as psychopathology, symptom severity, and global distress, these dual diagnosis patients were not significantly different from both inpatient psychiatric groups, with the exception of the Phobic Anxiety scale. Pre- and post-treatment measures (Figure 2) show that the patients achieved highly significant improvements on every SCL-90-R scale measuring psychological
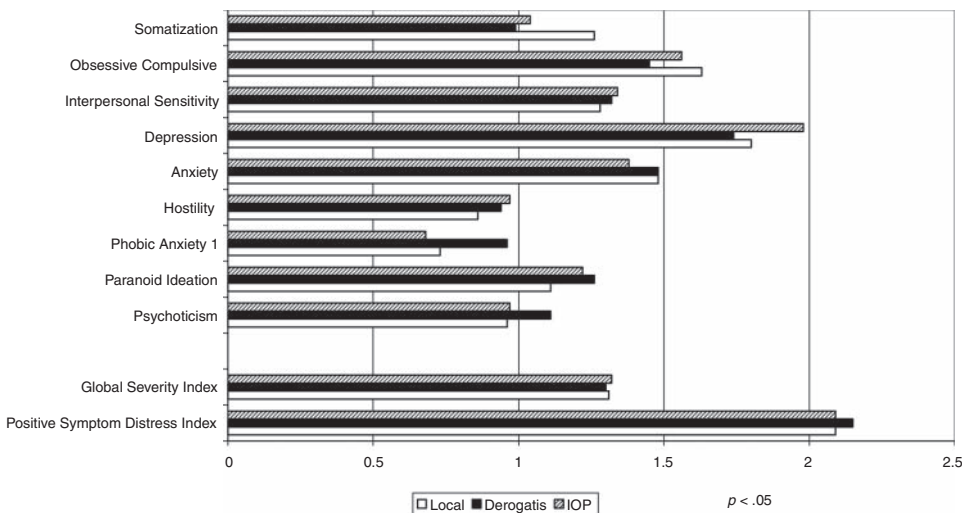
FIGURE 1 Mental Health Resources' (MHR) dual diagnosis IOP vs. psychiatric inpatient SCL-90-R Mx scores.
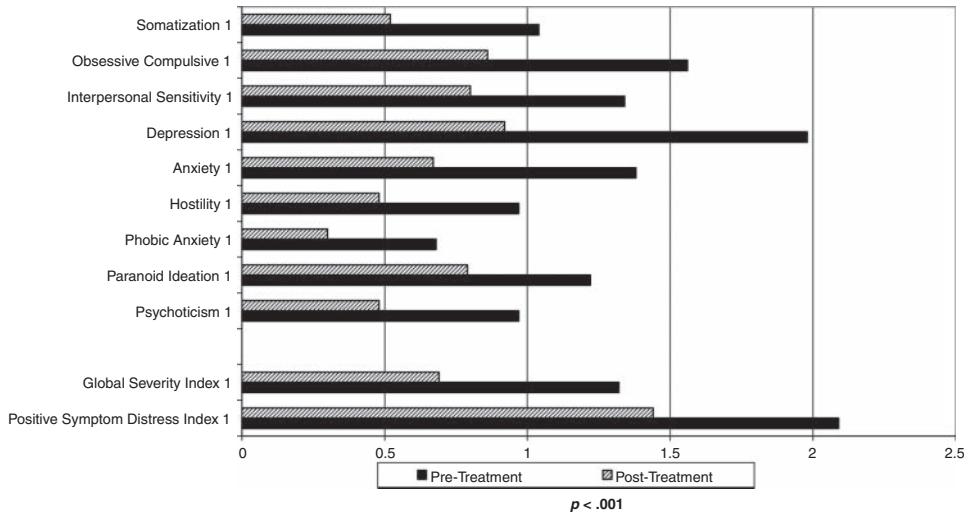
FIGURE 2 MHR's dual diagnosis pre- and post-treatment SCL-90-R average scores.

symptoms and distress. (The IOP Positive Symptom total scale score was not significantly different from either inpatient sample, but is not shown in the graphs as it is based on a different metric.)

Figure 3 shows a significant ($p < .00001$) dose-response relationship for depressive symptoms in these AOD users, as measured by the BSI Depression scale (Derogatis, 2001), across time. When pre- and post-treatment RCIs and CS cutoff criteria (Tingey et al., 1996) were applied to the SCL-90-R GSI to calculate RCIs and CS scores, 49%, 17%, and 4% met the respective 95%, 90%, and 80% CIs for both RCI and CS criteria of improvement, for a
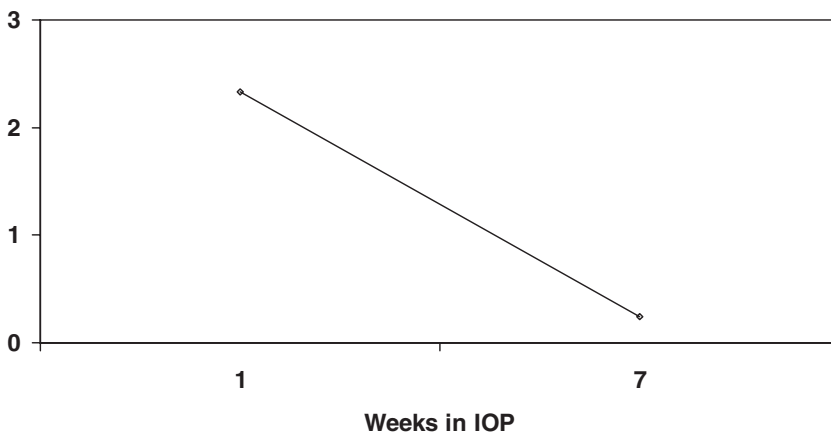

FIGURE 3 MHR's dual diagnosis average depressive symptom improvement from admission to discharge.

total improvement rate of 70%. Of those remaining, 27% were classified as Indeterminant and 3% Deteriorated based on these RCI and CS criteria.

Cohen's *d* (Cohen, 1988) for correlated measures in a pre-/post-treatment design (Lipsey & Wilson, 2001) was conducted using the SCL-90-R Depression scale scores ($Mx_1 = 1.98$, $SD_1 = 1.06$; $Mx_2 = 0.91$, $SD_2 = 0.86$; $r = .46$, $t = 9.09$, $n = 77$). This resulted in $d = 1.08$ for the SCL-90-R Depression scale, further indicating large effects in terms of symptom reduction. Cohen's *d* was also calculated using the SCL-90-R GSI scale scores ($Mx_1 = 1.32$, $SD_1 = 0.78$; $Mx_2 = 0.69$, $SD_2 = 0.61$; $r = .37$) and this resulted in $d = .89$, also a large effect according to Cohen's (1988) classification. Taken together, these effect size statistics confirm that approximately 71% to 74% of patients would be expected to show significant symptom improvement.

In addition to the wide variety of significant psychopathology and co-morbidity exhibited on the SCL-90-R and Axes I through IV, these patients also reported using AOD. The 66 patients who used AOD in the week prior to admission used them an average of 4.5 ($SD = 3.85$) days in the week prior to admission. Because of the curvilinear nature of the change over time, a log transformation of the time variable was necessary to meet the HLM assumptions (Howard, Moras, Brill, Martinovich, & Lutz 1996). These patients showed a significant and steady decline throughout treatment (Figure 4). 82% reduced the number of days used and Cohen's d = -2.33, indicating a very large reduction in the number of days used. Of the 33 patients who had not used AOD in the week prior to treatment, the average number of days used during treatment was .26, and 20 patients reported that they remained abstinent. Consistent with this, the HLM was not significant (*d"1 "*) +*Ł"·"ˆ
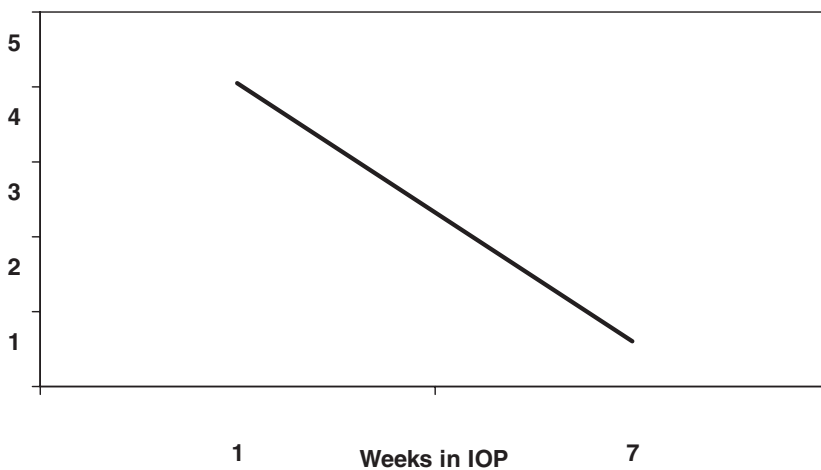
FIGURE 4 MHR's dual diagnosis IOP average number of days using of those who were actively using in the week prior to admission.

TABLE 1 GSI RCI + PDA Changed for Those Who Entered IOP Using ≥1 Day

|	PDA Changed					
	100%	58–86%	30–44%	16%	0%	≥–16%
GSI						
1.96		4	7	3	7	
1.28		2	3	3	1	
0.84	2	2	1			
<0.84 to <–0.83	1	3	4	2	2	1

Note. N = 48.

In an effort to assess the treatment effects on the combined outcome variables of psychiatric symptom severity index (GSI) and the PDA, the pre- and post-treatment RCIs for the GSI scale were combined with the PDA pre- and post-treatment measure as the CS variable. In order to be classified as improving on PDA, an individual had to improve PDA by at least 1 day (16%). Table 1 shows that 56% ($n = 27$) of the sample achieved both reliable and clinically significant improvements with respect to both severity of psychiatric symptoms and PDA. Only one individual (2%) increased use and did not meet any of the RCI criteria. Twenty-one percent ($n = 10$) increased PDA but did not significantly reduce severity of symptoms. Of the 21% who did not change the PDA, eight (80%) showed reliable symptom improvements. While these measures portray conservative estimates of CS, it should also be noted that at the completion of treatment, 53% ($n = 26$) had achieved 100% abstinence and 16% ($n = 8$) reported having used only 1 day in the previous week. Of these 34, 79% ($n = 27$) achieved reliable GSI improvements. However, rather than treat abstinence as a static outcome end point, it is believed that measuring change, particularly when abstinence is not necessarily the patient's goal, is a more meaningful measure.

Fifty-two percent of the patients returned collateral contact questionnaires, primarily completed by friends (43%) or family members (55%). While 20% of the total treatment sample reported that they did not use any substances in the week preceding treatment, 52% of the collateral contacts reported that the individual had not used in that week ($\chi^2(2) = 7.71, p < .05$). Collateral contacts reported that the patient used substances an average of 1.8 ($SD = 2.60$) days in the previous week, whereas these patients reported using an average of 2.7 ($SD = 3.86$) days ($t(51) = 2.05, p < .05$). When the matching collateral information from patients who admitted use in the week prior to treatment were analyzed with their cohorts ($n = 32$), the significant difference increased slightly ($t(31) = 2.55, p < .02$).

Figure 5 demonstrates highly significant client-rated improvements in all areas of functioning. Additionally, these findings were further supported by the clinicians' GAF ratings at admission and discharge. The average GAF score was 39.8 on admission and 54.41 at discharge ($t(97) = 15.02; p < .001$).
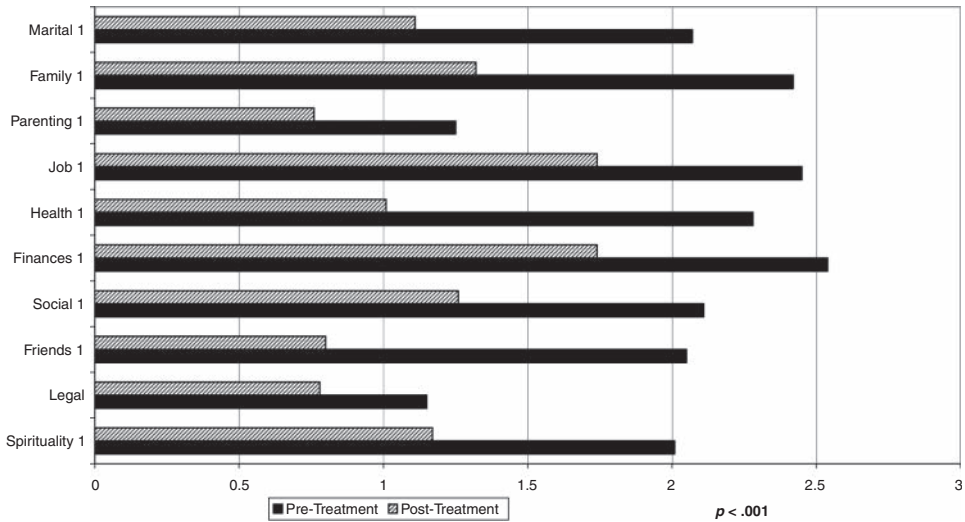
FIGURE 5 MHR's dual diagnosis pre- and post-treatment IOP functional ratings.

Finally, Figure 6 depicts a very high degree of client satisfaction on the CSQ-8 (Attkisson & Greenfield, 1994). On a scale of 1 to 4, the overall average rating was 3.68. These dual diagnosis IOP patients reported significantly greater overall satisfaction based on the total CSQ-8 score compared to the mental health normative group ($t(3194) = 4.97$, $p < .0001$). The dual diagnosis IOP clients also reported significantly greater satisfaction on each of the individual satisfaction items (Figure 6).
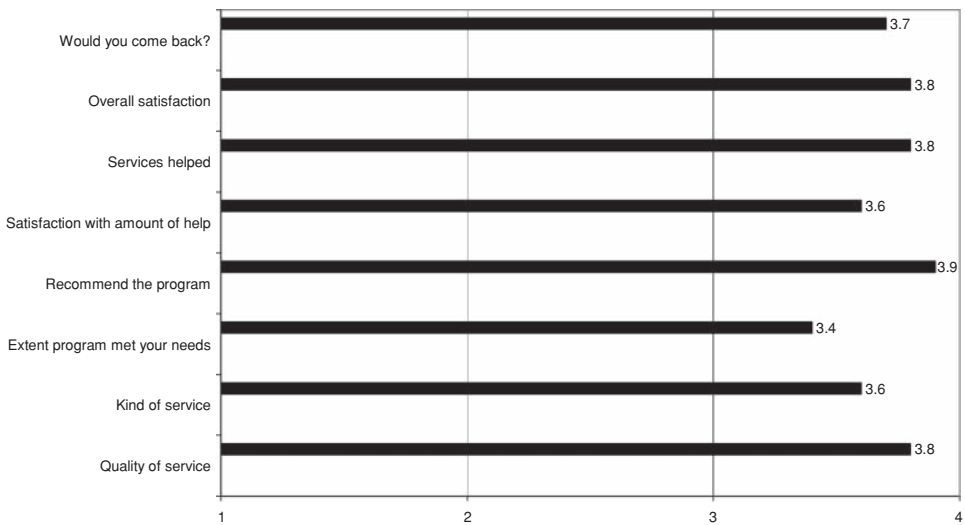
FIGURE 6 MHR's dual diagnosis IOP client satisfaction.

CONCLUSIONS

As hypothesized, these dual diagnosis patients demonstrated significant levels of psychopathology that were comparable to those found on the SCL-90-R inpatient norms and marked functional impairments. Significant treatment gains were shown with respect to psychiatric symptomatology, number of days used, and functional impairments. Similarly, these patients were highly satisfied with their treatment.

These findings demonstrate that medically stable dual diagnosis patients, with psychological symptoms comparable to those of psychiatric inpatients, can be effectively treated on an outpatient basis in an IOP. The fact that this models operates as a hospital diversion program is evidenced not only by the case mix complexity, including the presence of suicidal or homicidal ideation in 56% of the sample, but also by the fact that 98% of the patients were direct admissions, 15% had previous inpatient treatment, and only 4% were referred to any higher level of care. These findings are consistent with those of Granholm et al. (2003), who demonstrated that integrated dual diagnosis outpatient programming can significantly reduce hospital days.

Patients who reported using substances an average of 4.5 days per week prior to admission showed significant, predictable declines in use throughout treatment, reporting significant reductions in use that on average could lead to less than 1 day a week if treatment were completed. Additionally, this goal could be accomplished while simultaneously significantly reducing their psychological symptoms. The convergence of the RCI + CS results using the SCL-90-R GSI scores and cutoffs indicated that 70% of these patients showed statistically reliable and clinically significant improvement with respect to the severity of their psychiatric symptoms. As one would expect, when psychiatric symptoms and PDA were combined for the RCI + CS analysis, more conservative findings were obtained. Under these conditions, 56% demonstrated reliable symptom reductions and clinically significant increases in days abstinent. Viewed another way, 53% of these patients achieved abstinence and an additional 16% reported using 1 day in the week prior to discharge, whereas the average days used at admission was 4.5 ($SD = 3.85$).

Similarly, the effect sizes of $d = 1.08$ for Depression and $d = .89$ for GSI further indicates that approximately 71% to 74% of treated patients can be expected to have a successful outcome if they complete treatment. In light of the absence of published IOP outcome data (Drake et al., 2008), it is difficult to compare these findings to other IOPs. By way of comparison, Dutra et al. (2008) reviewed controlled clinical trials for individual psychosocial interventions for AOD disorders and found an average abstinent rate of 31% and an aggregate $d = .45$. Hesse (2009) reviewed integrated treatment programs for depression and substance abuse in outpatient, inpatient, and partial hospital dual diagnosis treatment programs and found $d = .58$ for depression and an average reduction of 14.13% PDA. Clearly, this IOP was

superior to these recently published controlled clinical trials with outpatient psychosocial treatments, partial hospitalization, and inpatient dual diagnosis treatment programs.

In addition to significant reductions in daily use and psychological symptoms, these patients also reported significant functional improvements in all spheres, which were corroborated by the significant improvement in clinician GAF ratings. Finally, patients were very satisfied with the quality of service, would recommend the program, would come back if needed, and overall were highly satisfied.

Wise (e.g., 2003b, 2005) previously demonstrated that acute psychiatric patients could be effectively treated in an IOP. These findings serve to illustrate that the previous outcomes obtained with psychiatric IOP patients have important implications for the treatment of substance-abusing dual diagnosis patients. More specifically, about one-third of the dual diagnosis IOP consisted of treatment identical to that previously proven to be effective with psychiatric patients (Wise, 1999, 2000, 2003a, 2003b, 2004, 2005). Thus, although the dual diagnosis population is more complicated to treat, it was not surprising that similar symptom improvements were obtained to that of our psychiatric patients. However, this dual diagnosis IOP also received new, unproven treatment modules designed to specifically address substance abuse issues. This study demonstrates that the use of this new material resulted in significant reductions of substance abuse in a medically stable dual diagnosis population who were using approximately 4 to 5 days per week and whose psychological acuity was comparable to psychiatric inpatients. This naturalistic study demonstrates that it is feasible to effectively treat dual diagnosis patients in an office-based IOP.

The dropout and treatment groups generally reported comparable levels of symptom distress, days used, and functional impairments. The fact that the dropouts reported significantly more DUIs and prior Alcoholics Anonymous or Narcotics Anonymous attendance indicates more experience with the criminal justice system and self-help groups. Perhaps these individuals tended to select themselves out of this treatment because they did not feel they fit in as well with the other patients. Alternatively, perhaps they were not motivated or did not feel they would benefit from treatment. In any event, with such a low number of dropouts ($n = 8$; 7%), it is difficult to draw any firm conclusions regarding this group, but it would seem to be a relatively low proportion of dropouts.

Limitations

Limitations of the study include those related to naturalistic or treatment as usual designs. There was no control group, and patients were not screened out based on co-morbidities or suicidal ideation, unless they were imminently

suicidal, homicidal, psychotic, or medically unstable. We were unable to partial out the effects of medications, although the majority of those who were on medications had already been receiving them prior to treatment. Similarly, in this "treatment as usual," private practice setting, we did not have ready access to urinalysis data or other bioassay data to corroborate patient reports regarding their reported use. Positive urine drug test results, however, are not without limitations, particularly in a harm reduction context. For example, a positive screen result does not indicate the frequency of use but only the presence or absence of a drug for a given time frame. Hence, the infrequent use of alcohol could be undetected, whereas screening for marijuana on a weekly basis could result in false-positives. Those who do not accurately self-report might also use widely available methods to alter or substitute urine specimens (e.g., see Dasgupta, 2007; Jaffee, Trucco, Levy, & Weiss, 2007). Most importantly, however, the use of biological measures are inconsistent with the philosophies of an MI and harm reduction approach if abstinence is not the patients' goal. In the frequent case in which a patient's goal was to reduce use, a positive screen result would not be a useful measure of reduction.

Additionally, Miller and Wilbourne (2002) found that approximately 57% of 361 controlled studies did not use objective verification of self-report. Numerous reviews of self-reported alcohol abuse have generally found self-reports to be consistently reliable (e.g., Amor, Polich, & Stambul, 1978; Maisto & Cooper, 1980; McLellan et al., 1985; Polich, 1982; O'Farrell & Maisto, 1987; Skinner, 1984; Sobell & Sobell, 1978, 1981, 1982) if not optimal measures (Sobell, Sobell, Connors, & Agrawal, 2003). Luty, Perry, Umoh, and Gormer (2006) reported numerous studies that demonstrated the validity of substance use self-report questionnaires in general and the MAP-sc specifically. Alternatively, there are some data with schizophrenic and bipolar patients indicating that cognitive impairment and state of sobriety at the time of the self-report predict the accuracy of the self-report. Consistent with this, other reports indicate that setting and contextual variables (e.g., criminal justice, emergency department, type of drug, psychosis) impact the validity of self-reported drug use (e.g., Hser, Maglione, & Boyle, 1999). The current patients primarily used alcohol and marijuana, were not court-mandated, were nonpsychotic, were in a private practice setting, were not intoxicated at the time of the report, had no cognitive impairments, and reported higher rates of use than their collateral contacts. In fact, they reported using an average of 4.5 days in the week prior to admission, which is not consistent with a response set designed to minimize or deny use. Subsequently, it is believed that the setting and patient variables, convergence of multiple measures, from multiple sources (clients, clinician, and collaterals) supports the validity of these findings.

Similarly, it is of interest to note the discrepancy between the patient and collateral contact reports on the number of days of use in the week

prior to admission. While it is possible that patients selected informants who would underreport number of days using, it seems illogical that the patients themselves would then report more use than their collaterals. In fact, there is a strong body of literature that indicates self- and collateral reports tend to be highly reliable and accurate (e.g., Babor, Steinberg, Anton, & Del Boca, 2000; Babor, Stephens, & Marlatt, 1987; Connors & Maisto, 2003; Marlatt, Stephens, Kivlahan, Buef, & Banaji, 1986; O'Farrell & Maisto, 1987). Together, these reviews demonstrate that there is a high degree of agreement between subjects and collaterals and that when discrepancies occur, typically the subjects' report shows more impairment compared to collaterals. Consistent with the literature, collateral contacts in this study underestimated the frequency of use and were not more reliable informants than the patients themselves. This finding suggests that with this population the need for collateral corroboration of substance use may not be necessary for judging the reliability of self-reported use. On the other hand, it may be helpful to identify and involve collaterals who underreport, in an effort to educate them about the extent of their loved ones' use and enlist their support.

Using cumulative measures of days used/abstinence across all AODs is arguably a limitation. However, this is a condition of treatment as usual. Additionally, in light of the type of drugs primarily used, the fact that most patients used more than one drug, and that the goal of treatment was to reduce use, PDA was deemed a measure that was reliable across AODs to assess harm reduction.

An attempt was made to obtain quantifiable data regarding quantity used and amount spent per week for each substance. However, these data were deemed invalid because quantities were not reported consistently between respondents (e.g., 3 drinks, half a pint, 1 joint, a nickel bag) and amount spent was typically missing. Hence, by focusing on PDA, change may have been underestimated because individuals may have reduced the quantities used per day. However, percentage days using or abstinent are widely accepted benchmarks. The use of PDA as a CS variable might also be criticized along the grounds that a drinker, for example, may reduce the number of days used from 5 to 2, but still drink to intoxication and demonstrate role impairment(s) as a result. While such use may indeed prove problematic, mandating abstinence in a private practice setting is not feasible and is inconsistent with our treatment philosophy. Instead, we choose to view these individuals in the framework of a harm reduction model, having made reliable and clinically significant changes; they may well be "on their way" to abstinence.

Finally, one reviewer raised a question about researcher allegiance. Researcher allegiance effects have been shown to exert a significant positive effect on psychotherapy outcome research when two or more treatments are compared against each other (e.g., Luborsky et al., 1999). However, there

has been considerable debate about what this means (e.g., Beutler, 2002; Luborsky et al., 2002; Thase, 1999), how relevant it is (e.g., Lambert, 1999), and what can be done about it (Jacobson, 1999). Indeed, Shaw (1999) stated that "if allegiance to a treatment affects the outcome of treatment ("Mecca effect") ... then we would wise to *maximize* these results ... " (p. 131; emphasis added). Luborsky et al. (2002) indicated that when comparisons of treatments were studied, corrections for allegiance might be indicated, but " ... may not be entirely preventable ... " (p. 102). Jacobson (1999) has stated that researcher alliance is less of a problem when "the relative efficacy of the treatment relative to others is not at issue" (p. 118). Hence, while researcher allegiance may have an effect, this appears to be less of a problem when treatments are not being compared, as is the case here, according to Jacobson (1999) and Luborsky et al. (2002). Additionally, Luborsky et al. (2002) found that the methodological quality of the treatment comparison research was not associated with allegiance. Consistent with this, the present naturalistic study relied upon standardized, widely accepted instrumentation, with reasonably sound methodological techniques, and five previously published peer-reviewed studies have obtained similar findings, thereby lending support to the reliability, robustness, and now generalization of the findings to dual diagnosis patients.

REFERENCES

Amor, D., Polich, J., & Stambul, H. (1978). *Alcoholism and treatment*. New York: Wiley Interscience.

Attkisson, C., & Greenfield, T. (1994). Client Satisfaction Questionnaire-8 and Service Satisfaction Scale-30. In M. Maruish (Ed.), *The use of psychological testing for treatment planning and outcome assessment*. Hillsdale, NJ: Erlbaum.

Babor, T. F., Steinberg, K., Anton, R., & Del Boca, F. (2000). Talk is cheap: Measuring drinking outcomes in clinical trials. *Journal of Studies on Alcohol*, *61*, 55–63.

Babor, T. F., Stephens, R. S., & Marlatt, G. A. (1987). Verbal report methods in clinical research on alcoholism: Response bias and its minimization. *Journal of Studies on Alcohol*, *48*, 410–424.

Barry, A., & Lefkovitz, P. (2006). *Overview of the partial hospitalization & intensive outpatient industry: 2005*. Portsmouth, VA: Association for Ambulatory Behavioral Healthcare.

Beutler, L. E. (2002). The dodo bird is extinct. *Clinical Psychology: Science and Practice*, *9*(1), 30–34.

Borntrager, C., Chorpita, B., Higa-McMillan, C., & Weisz, J. (2009). Provider attitudes towards evidence-based practices: Are the concerns with the evidence or with the manuals? *Psychiatric Services*, *60*(5), 677–681.

Cisler, R. A., Kowalchuk, R. K., Saunders, S. M., Zweben, A., & Trinh, H. Q. (2005). Applying clinical significance methodology to alcoholism treatment trials: Determining recovery outcome status with individual- and population-based measures. *Alcoholism: Clinical and Experimental Research*, *29*(11), 1991–2000.

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Hillsdale, NJ: Erlbaum.

Connors, G. J., & Maisto, S. A. (2003). Drinking reports from collateral individuals. *Addiction*, *98*(suppl 2), 21–29.

Dasgupta, A. (2007). The effects of adulterants and selected ingested compounds on drugs-of-abuse testing in urine. *American Journal of Clinical Pathology*, *128*(3), 491–503.

Derogatis, L. (1994). *SCL-90-R: Administration, scoring and procedures manual-II*. Minneapolis, MN: NCS, Pearson.

Derogatis, L. (2001). *Brief Symptom Inventory: Administration, scoring and procedures manual*. Minneapolis, MN: NCS, Pearson.

Drake, R. E., Mercer-McFadden, C., Mueser, K. T., McHugo, G. J., & Bond, G. R. (1998). Review of integrated mental health and substance abuse treatment for patients with dual disorders. *Schizophrenia Bulletin*, *24*(4), 589–608.

Drake, R. E., O'Neal, E., & Wallach, M. (2008). A systematic review of psychosocial research on psychosocial interventions for people with co-occurring severe mental and substance abuse disorders. *Journal of Substance Abuse Treatment*, *34*(1), 123–138.

Dutra, L., Stathopoulou, G., Basden, S., Leyro, T., Powers, M., & Otto, M. (2008). A meta-analytic review of psychosocial interventions for substance use disorders. *American Journal of Psychiatry*, *165*(2), 179–187.

Granholm, E., Anthenelli, R., Monteiro, R., Sevcik, J., & Stolar, M. (2003). Brief integrated outpatient dual diagnosis treatment reduces psychiatric hospitalizations. *American Journal of Addictions*, *12*(4), 306–313.

Grant, B., Stinson, F., Dawson, D., Chou, P., Dufour, M., Compton, W., et al. (2004). Prevalence and co-occurrence of substance use disorders and independent mood and anxiety disorders: Results from the national epidemiologic survey on alcohol and related conditions. *Archives of General Psychiatry*, *61*(8), 807–816.

Hesse, M. (2009). Integrated psychological treatment for substance use and co-morbid anxiety or depression vs. treatment for substance use alone: A systematic review of the published literature. *BMC Psychiatry*, *9*(6), 1–8.

Howard, K. I., Lueger, R. J., Maling, M. S., & Martinovich, Z. (1993). A phase model of psychotherapy outcome: Causal mediation of change. *Journal of Consulting and Clinical Psychology*, *61*(4), 678–685.

Howard, K. I., Moras, K., Brill, P. L., Martinovich, Z., & Lutz, W. (1996). Evaluation of psychotherapy: Efficacy, effectiveness, and patient progress. *American Psychologist*, *51*(10), 1059–1064.

Hser, Y., Maglione, M., & Boyle, K. (1999). Validity of self-report of drug use among STD patients, ER patients and arrestees. *American Journal of Drug and Alcohol Abuse*, *25*(1), 81–91.

Jacobson, N. S. (1999). The role of the allegiance effect in psychotherapy research: Controlling and accounting for it. *Clinical Psychology: Science and Practice*, *6*(1), 116–119.

Jacobson, N. S., & Truax, P. (1991). Clinical significance: A statistical approach to defining meaningful change in psychotherapy research. *Journal of Consulting and Clinical Psychology*, *59*(1), 12–19.

Jaffee, W. B., Trucco, E., Levy, S., & Weiss, R. D. (2007). Is this urine really negative? A systematic review of tampering methods in urine drug screening and testing. *Journal of Substance Abuse Treatment*, *33*(1), 33–42.

Lambert, M. (1999). Are differential treatment effects inflated by researcher therapy allegiance? Could Clever Hans count? *Clinical Psychology: Science and Practice*, *6*(1), 127–130.

Lipsey, M. W., & Wilson, D. B. (2001). *Practical meta-analysis*. Thousand Oaks, CA: Sage.

Luborsky, L., Diguer, L., Seligman, D., Rosenthal, R., Krause, E., Johnson, S., et al. (1999). The researcher's own therapy allegiances: A "wild card" in comparisons of treatment efficacy. *Clinical Psychology: Science and Practice*, *6*(1), 95–106.

Luborsky, L., Rosenthal, R., Diguer, L., Andrusyna, T., Berman, J., Levitt, J., et al. (2002). The dodo bird verdict is alive and well—mostly. *Clinical Psychology: Science and Practice*, *9*(1), 2–12.

Luty, J., Perry, V., Umoh, O., & Gormer, D. (2006). Validation and development of a self-report outcome measures (MAP-sc) in opiate addiction. *Psychiatric Bulletin*, *30*, 134–139.

Maisto, S. A., & Cooper, A. (1980). A historical perspective on alcohol and drug treatment outcome research. In L. C. Sobell, M. B. Sobell, & E. Ward (Eds.), *Evaluating alcohol and drug abuse treatment effectiveness: Recent advances* (pp. 1–14). New York: Pergamon.

Marlatt, G. A., Stephens, R. S., Kivlahan, D., Buef, J. J., & Banaji, M. (1986). *Empirical evidence on the reliability and validity of self·reports of alcohol use and associated behaviors*. Workshop on the Validity of Self-Report in Alcoholism Treatment Research. Washington, DC: National Institute on Alcohol Abuse and Alcoholism.

Marsden, J., Gossop, M., Stewart, D., Best, D., Farrell, M., & Strang, J. (1998). *The Maudsley Addiction Profile: A brief instrument for treatment outcome research. Development and user manual*. London: National Addiction Centre, Institute of Psychiatry.

McLellan, A. T., Luborsky, L., Cacciola, J., Griffith, J., Evans, F., Barr, H., et al. (1985). New data from the Addiction Severity Index: Reliability and validity in three centres. *Journal of Nervous and Mental Disease*, *173*(7), 412–423.

Miller, W. R. (2004a). *Enhancing motivation for change in substance abuse treatment*. Rockville, MD: SAMSHA, Center for Substance Abuse Treatment.

Miller, W. R. (Ed.). (2004b). *COMBINE Monograph Series, Volume 1. Combined behavioral interventions manual: A clinical research guide for therapists treating people with alcohol abuse and dependence*. DHHS Publication No. 04-5288. Bethesda, MD: NIAAA.

Miller, W. R., & Wilbourne, P. (2002). Mesa Grande: A methodological analysis of clinical trials of treatments for alcohol use disorders. *Addiction*, *97*(3), 265–277.

O'Farrell, T., & Maisto, S. A. (1987). The utility of self-report and biological measures of alcohol consumption in alcoholism treatment outcome studies. *Advances in Behavioral Research and Therapy*, *9*(2–3), 91–125.

Polich, J. (1982). The validity of self-reports in alcoholism research. *Addictive Behaviors*, *7*(2), 123–132.

Roberts, L. J., Neal, D. J., Kivlahan, D. R., Baer, J. S., & Marlatt, G. A. (2000). Individual drinking changes following a brief intervention among college students: Clinical significance in an indicated preventive context. *Journal of Consulting and Clinical Psychology*, *68*(3), 500–505.

Shaw, B. (1999). How to use the allegiance effect to maximize competence and therapeutic outcomes. *Clinical Psychology: Science and Practice*, *6*(1), 131–132.

Singer, J. D., & Willett, J. B. (2003). *Applied longitudinal data analysis: Modeling change and event occurrence*. New York: Oxford University Press.

Skinner, H. A. (1984). Assessing alcohol use by patients in treatment. In R. C. Smart, H. D. Cappell, F. B. Glaser, Y. Israel, H. Kalant, R. E. Popham, et al. (Eds.), *Research advances in alcohol and drug problems, 8*, 183–107. New York: Plenum Press.

Sobell, L. C., & Sobell, M. B. (1978). Validity of self-reports in three populations of alcoholics. *Journal of Consulting and Clinical Psychology*, *46*(5), 901–907.

Sobell, L. C., & Sobell, M. B. (1981). Outcome criteria and the assessment of alcohol treatment efficacy. *Evaluation of the alcoholic: Implications for research theory and treatment*. (Research monograph 5, pp. 369–382). Rockville, MD: National Institute on Alcohol Abuse and Alcoholism.

Sobell, L. C., & Sobell, M. B. (1982). *Alcoholism treatment outcome evaluation methodology, prevention, intervention and treatment. Concerns and models*. (Alcohol and health monograph 3, pp. 293–321). Rockville, MD: National Institute on Alcohol Abuse and Alcoholism.

Sobell, L. C., Sobell, M. B., Connors, G., & Agrawal, S. (2003). Assessing drinking outcomes in alcohol treatment efficacy studies: Selecting a yardstick of success. *Alcoholism: Clinical and Experimental Research*, *27*(10), 1661–1666.

Substance Abuse and Mental Health Services Administration. (2007). *Results from the 2006 National Survey on Drug Use and Health: National findings* (Office of Applied Studies, NSDUH Series H-32, DHHS Publication No. SMA 07-4293). Rockville, MD: National Institute on Alcohol Abuse and Alcoholism.

Thase, M. (1999). What is the investigator allegiance effect and what should we do about it? *Clinical Psychology: Science and Practice*, *6*(1), 113–115.

Tiet, Q., & Mausbach, B. (2007). Treatments for patients with dual diagnosis: A review. *Alcoholism: Clinical and Experimental Research*, *31*(4), 513–536.

Timko, C., Chen, S., Sempel, J., & Barnett, P. (2006). Dual diagnosis patients in community or hospital care: One year outcomes and health care utilization and costs. *Journal of Mental Health*, *15*(2), 163–177.

Tingey, R., Lambert, M., Burlingame, G., & Hansen, N. (1996). Assessing clinical significance: Proposed extensions to method. *Psychotherapy Research*, *6*(2), 109–123.

Wise, E. A. (1999). Clinical outcomes, client satisfaction and innovative programming in private practice. In K. M. Coughlin (Ed). *Behavioral outcomes & guidelines sourcebook: A practical guide to measuring, managing and standardizing mental health and substance abuse treatment* (pp. 135–143). New York: Faulkner & Gray.

Wise, E. A. (2000). Mental health intensive outpatient programming: An outcome and satisfaction evaluation of a private practice model. *Professional Psychology: Research and Practice*, *31*(4), 412–417.

Wise, E. A. (2003a). Empirical validation of a mental health intensive outpatient program in a private practice setting. *American Journal of Orthopsychiatry*, *73*(4), 405–410.

Wise, E. A. (2003b). Psychotherapy outcome and satisfaction methods applied to intensive outpatient programming in a private practice setting. *Psychotherapy: Theory, Research, Practice and Training*, *40*(3), 203–214.

Wise, E. A. (2004). Methods for analyzing psychotherapy outcomes: A review of clinical significance, reliable change and recommendations for future directions. *Journal of Personality Assessment*, *82*(1), 50–59.

Wise, E. A. (2005). Effectiveness of intensive outpatient programming in private practice: Integrating practice, outcomes and business. *American Psychologist*, *60*(8), 885–895.