

## PSYCHOTHERAPY OUTCOME AND SATISFACTION METHODS APPLIED TO INTENSIVE OUTPATIENT PROGRAMMING IN A PRIVATE PRACTICE SETTING

EDWARD A. WISE

*Mental Health Resources*

*This study evaluated an Intensive Outpatient Program (IOP) in a private practice setting. Patients received the Symptom Checklist–90—Revised pre- and posttreatment and the Client Satisfaction Questionnaire–8 posttreatment, which were used to measure symptom change and client satisfaction. Reliable Change Indices and Clinically Significant change scores were calculated for each individual (N = 225) using several variables from therapist and clinician perspectives. Highly significant improvement rates on client symptom ratings were also validated by clinician global assessment of functioning ratings, client satisfaction, and discharge rates to lower levels of care. This IOP was demonstrated to be more effective than psychotherapy in clinically representative studies, randomized clinical trials, and controlled clinical trials as well as comparable to many widely used medical interventions.*

Psychotherapy outcome studies have developed and matured since the classic Smith, Glass, and Miller (1980) meta-analytic review. For example, quasi-experimental pre- and posttest de-

signs, randomized clinical trials, manualized therapy studies, and clinically representative studies have attempted to evaluate therapy outcomes. Relatively new methodological developments, such as Jacobson and Truax's (1991) Reliable Clinical Index (RCI) statistic, which accounts for measurement error, and their addition of Clinically Significant change (CS), requiring that the patient move from a dysfunctional to functional range on the normative continuum, have refined the definitions of "treatment outcomes." These developments have given rise to what Howard, Moras, Brill, Martinovich, and Lutz (1996) have called "patient focused research," in which the emphasis is not just on the outcomes for treated groups but on knowing how a particular individual in that group responded to a specific treatment or provider.

In the interim, as a result of increased financial pressures and economies of scale, managed care organizations have facilitated the development of partial hospitalization programs (PHPs) and other alternative service delivery models (e.g., at home interventions and mobile crisis triage). A relatively new development is an Intensive Outpatient Program (IOP), defined as one that offers 3 to 4 hr of structured programming three to five times a week (e.g., American Association of Community Psychiatrists [AACCP], 1998; American Society of Addiction Medicine [ASAM], 1991; Kiser, Lefkovitz, Kennedy, & Knight, 1996; Kiser et al., 1998). Obviously, the costs for 9–15 hr of programming per week are considerably less than the typical 30 hr per week partial program and are far less costly than 24-hr care. Like the PHPs before them, however, IOPs are being underutilized as a result of inequitable funding, lack of insurer arrangements, and clinician bias against utilizing new services (Piper, Rosie, Joyce, & Azim, 1996; Washton, 1997).

---

Correspondence regarding this article should be addressed to Edward A. Wise, Mental Health Resources, 1027 South Yates Road, Memphis, Tennessee 38119. E-mail: eawmhr@aol.com

Thus, access to alternative outpatient programs has been limited because of restricted outpatient benefits, efforts by stakeholders to maintain the status quo, and an absence of empirical evidence supporting alternative levels of care that have been successfully implemented in the public sector to limit expenses (Goldstein & Horgan, 1988).

This program is designed for people who are in acute psychological distress, such as those experiencing a suicidal crisis, unrelenting panic attacks, or other symptoms that render them unable to work or carry out activities of daily living but who can be treated on an outpatient basis. This approach is well suited for others in acute distress requiring crisis intervention in a therapeutic milieu combined with active psychotherapy, problem-solving skills, and support. The IOP is also appropriate for patients who have been recently discharged from short-term hospitalization or partial hospitalization and who continue to require daily therapeutic contact but not daily medical monitoring. Individuals who require more than 3 hr of care per day or who are imminently suicidal, homicidal, or psychotic are not appropriate for this level of care.

The IOP reported in this study is pragmatically eclectic in orientation and operates within a private practice setting. The IOP is run by two doctoral-level psychologists and a master's level counselor. A consulting psychiatrist is in the office one half day a week for medication management. The IOP has three core components that target affect, cognition, and psychophysiological symptoms. The typical group size is about 6–8 patients. As in inpatient and PHP programs, new patients are added and discharged on a continuous basis.

The affective component consists of learning to identify and express feelings, finding safe and appropriate outlets for avoided feelings, and tolerating high levels of emotion without becoming overwhelmed. As in traditional group psychotherapy, resistance to affect or change is confronted and interpersonal as well as family-of-origin problems are identified. Repetitive relationship themes that adversely affect current daily functioning are addressed. Patients are assisted in labeling their emotional experience and expressing it in meaningful ways. Difficulties with anger, fear, or sadness are addressed in this context.

The cognitive aspect of the program consists of teaching problem-solving skills based on the

principles of rational behavior therapy. Crisis situations or conflictual relationships are anticipated, action plans are developed, cognitive errors are identified, and rational behavioral strategies are renewed and strengthened. Assertiveness, problem solving, anger management, anxiety management, and similar skills training take place by examining one's self-talk, disputing irrational ideas, and teaching coping skills.

The psychophysiological aspects of the program are concerned with bodily processes that may exacerbate symptoms or be misinterpreted by patients. Relaxation exercises, skills for coping with chronic illness, and pain management are some of the typical subject areas addressed.

The psychoeducational groups focus on different content areas depending on the individual needs of the group participants. Some of the areas that may be addressed are loss, out-of-control teens, workplace problems, chronic illness, conflict, and self-defeating behavioral patterns. These groups are primarily designed to teach and rehearse new coping skills that provide new strategies or solutions to current problems. Approaches such as daily homework assignments, self-monitoring, and diaries are frequently used to develop and maintain coping strategies. (More descriptive information about the program may also be found in Wise, 2003.)

The goals of the current study were to present data on a relatively large sample of IOP patients in a clinically representative, private practice setting, using standardized symptom and satisfaction measures in a pre- and posttest design, accompanied by RCI and CS analyses. The study demonstrates that sophisticated outcome research and IOP treatment programming can be delivered and studied in private practice settings.

## Method

All patients went through an intake screening process. At that time, a mental status exam and psychosocial history were obtained, along with the completion of numerous standard admission forms (e.g., insurance, consent, and releases). Patients were specifically assessed to be sure that they met predefined admission criteria (Wise, 2000). All patients completed a Symptom Checklist-90—Revised (SCL-90—R; Derogatis, 1983) prior to intake. The SCL-90—R is a self-administered, 90-item, 5-point rating scale (0–4) that has been shown to be a valid and reliable

measure of nine psychological symptom dimensions (Somatization, Obsessive–Compulsive, Interpersonal Sensitivity, Depression, Anxiety, Hostility, Phobic Anxiety, Paranoid Ideation, and Psychoticism subscales). There are also three global indices that assess broad domains of psychological distress. Positive Symptom Total (PST) is a measure of the total number of symptoms endorsed, Positive Symptom Distress Index (PSDI) reflects the average intensity of symptoms endorsed, and the Global Severity Index (GSI) is often used as a global measure of psychological distress as it utilizes data from both the total number of symptoms endorsed and the intensity level of distress. The SCL–90—R has been repeatedly demonstrated to be a sensitive measure of therapy outcomes in a wide variety of settings with a diversity of patient populations and numerous types of treatments, and has been recommended for assessing outcomes in clinical settings. IOP pretreatment SCL–90—R comparisons were made with Derogatis’s inpatient normative sample and with a more contemporary local sample of 100 consecutive psychiatric inpatient referrals for psychological testing, to establish baseline severity compared with hospitalized patients. Several studies have examined the relationship between RCI and client satisfaction and have reported inconsistent or incomplete results (e.g., Ankuta & Abeles, 1993; Lunnen & Ogles, 1998; Pekarik & Wolff, 1996). In an attempt to assess an additional dimension of CS from the client’s perspective, the Client Satisfaction Questionnaire–8 (CSQ–8; Attkisson & Zwick, 1982) was included with the rationale that patients would not be satisfied with a program that did not significantly help them solve their problems. During the last week of treatment, patients again took the SCL–90—R along with the CSQ–8. All patients who completed treatment and pre- and post-SCL–90—Rs were included in the study. Because the CSQ–8 was not used from the beginning of the study, the number of completed surveys ( $n = 161$ ) is lower than the total sample ( $N = 225$ ). Other data regarding total number of visits, pre- and post-Global Assessment of Functioning (GAF), diagnoses, and so forth were gleaned from a review of the charts.

RCIs were calculated on pre- and posttreatment measurements. RCIs were introduced by Jacobson and Truax (1991) to determine if the magnitude of statistically significant differences in psychotherapy outcome studies reflected mea-

surement error or significant statistical change beyond that accounted for by such error. CS, determined by establishing a normative continuum and crossing the cutoff from one level of severity to another, was assessed utilizing formulas from Tingey, Lambert, Burlingame, and Hansen (1996). Because of the nature of the population, the severity of their symptoms and the fact that this is a higher level of care than once-a-week individual psychotherapy, it was not reasonable to utilize the “return-to-normal” CS criterion. Instead, CS change was assessed by movement from one level of the normative continuum to another, consistent with the Tingey et al. methodology.

These RCI and CS criteria are very stringent and difficult to meet. Consequently, this study made several adjustments compared with previous studies. For example, instead of using the GSI, a global measure of distress calculated by averaging the scores of all completed items, the Depression scale was used. With a population that is primarily depressed, the GSI scale minimizes specific symptom severity, as well as subsequent improvement. As an extreme example, a male patient who endorses all SCL–90—R Depression scale items with a “4,” the highest score allowable, but enters a “0” for all other symptoms, would obtain a GSI raw score of 0.58 and a  $T$  score of 62. On the other hand, this same individual would obtain a Depression raw score of 4.00 and a  $T$  score of 81. For a predominantly depressed population, the Depression scale would appear to be a more accurate measurement of symptom improvement or deterioration. Similarly, Chorpita (2001) stated that “the statistical power to detect change will be best afforded by those measures with the greatest conceptual specificity. In other words, to maximize group differences, a depression study should measure depression” (p. 452) and not a global construct. Consequently, the RCI and CS procedures were applied to the Depression subscale. Because Tingey et al. (1996) provided only normative data for the GSI scale, data from Tingey (1989) were obtained to establish normative continuum cutoffs for the Depression scale using the formula  $(SD 1)(Mx 2) + (SD 2)(Mx 1)/SD 1 + SD 2$ . The resulting cutoff points were as follows: Asymptomatic = .23, Mild = .67, Moderate = 1.47, and Severe = 2.13. While Tingey and Tingey et al. provided some very important data for calculating cutoff points along an SCL–90—R norma-

tive continuum, data remain lacking at the upper levels of severity. Data on severely distressed individuals are necessary to build on and clarify these more severe or dysfunctional upper ranges of the normative continuum with the SCL-90—R so that individuals who demonstrate CS can be more accurately and reliably identified. For this reason, the severe range in this study was calculated using the formula cited along with data from Tingey's severe range and the average IOP Depression score and standard deviation ( $Mx = 2.46$ ;  $SD = .92$ ; cutoff = 2.13). The criterion that adjacent samples be distinct was met by using the  $t$  and  $d$  tests, as advised by Tingey et al. Nonetheless, it should be understood that those patients who are the most distressed have to make disproportionately larger symptom reductions to meet the CS criteria, compared with those who are less symptomatic. Thus, these most distressed individuals may pass the RCI criteria but not the CS criteria. Others (e.g., Kazdin, 1999; Wise, in press-b) have recommended the use of operationalized "real-world" variables to determine CS. This study also evaluated the CS criteria using discharge to a lower level of care, discharge GAF scores, and CSQ-8 scores that exceeded the two standard deviation cutoff recommended by Jacobson and Truax (1991; i.e., pretreatment GAF  $\pm$  two standard deviations; normative CSQ-8  $\pm$  two standard deviations).

In an effort to correct the alpha level for the relatively large number of statistical tests conducted, Bonferroni or Hochberg post hoc analyses were made on all significant findings to reduce the likelihood of Type I errors (Norman & Streiner, 2000). All findings reported as significant surpassed the alpha levels determined by at least one of these corrections. Meta-analytic statistics were also calculated to facilitate the comparison of these findings to other studies.

## Results

This sample ( $N = 225$ ) may be described as predominantly White (69%) women (81%) whose average age was 41 years. They had an average of 13 years of education, 27% had 16 or more years of schooling, and 77% were employed. Fifty percent were married, 24% single, and 20% divorced; there were 5 widows or widowers and 10 checked "other" marital status. The majority of the sample (80%) was diagnosed with depressive disorders, anxiety was present in 32%

of the cases, and 49% had an additional comorbid Axis I diagnosis. Seventy-one percent ( $n = 159$ ) of the sample also received a comorbid personality disorder diagnosis, most frequently with mixed features. Fifty-five percent ( $n = 123$ ) reported current Axis III disorders. Eighty-one percent ( $n = 182$ ) of the sample was admitted directly to the IOP, and only 19% ( $n = 43$ ) was stepped down from a higher level of care. Eighty percent ( $n = 180$ ) of these patients were already receiving some type of psychotropic medication at the time of admission. Thirty-two percent ( $n = 72$ ), or nearly one third of the IOP sample, had a total of 131 previous psychiatric inpatient admissions among them, or an average of 1.8 admissions. Twenty-five percent ( $n = 18$ ) of those with previous inpatient admissions had three or more previous hospitalizations. The average number of IOP sessions was 17, occurring over an average of 53 days, for an average of 3 IOP sessions per week. Of the 225 IOP admissions, 7% ( $n = 15$ ) were readmitted to the IOP. Of the 15 IOP readmissions, 13 had a total of 35 previous inpatient admissions among them, for an average admission rate of 2.7 per patient. During the course of treatment in the IOP, 4% ( $n = 9$ ) were referred to a higher level of care. Eight percent ( $n = 20$ ) attended 3 or fewer sessions, terminated unilaterally, and were considered dropouts.

Figure 1 shows that the pretreatment IOP group was significantly more symptomatic than both the national or the local hospitalized psychiatric samples on all SCL-90—R scales, including the global distress scales, with the exception of the Psychoticism scale. (The PST scale is not presented in Figure 1 as it is based on a different metric.) Overall, it must be concluded that the IOP pretreatment group was significantly more distressed than local and national psychiatric inpatients. These patients endorsed a broad range of severe symptoms.

Figure 2 depicts the IOP pretreatment and posttreatment SCL-90—R scale scores. There, it can be seen that following treatment, the IOP sample demonstrated significant symptom reductions on all 9 of the SCL-90—R symptom scales, but most important, on the Depression, Anxiety, Obsessive-Compulsive, and Interpersonal Sensitivity scales. The posttreatment significant reductions in the GSI, PST scale, and PSDI provide additional evidence of significant decreases in the number and intensity of symptoms across all

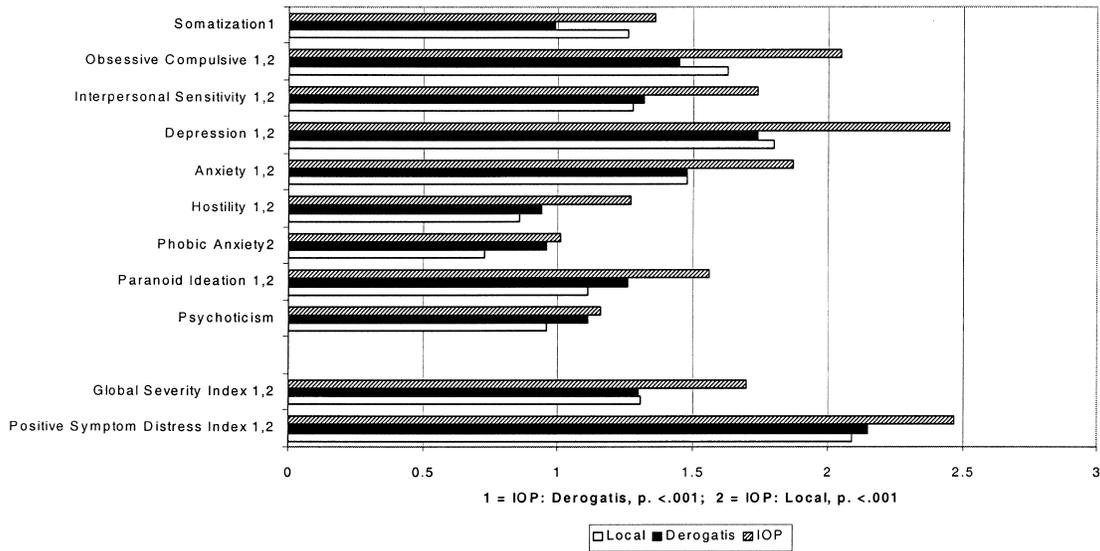


Figure 1. Intensive Outpatient Program (IOP) pretreatment Symptom Check List-90—Revised scores.

three of these broader measures of psychological distress. Furthermore, the significant increase in clinician rated pretreatment GAF ( $Mx = 39$ ) to posttreatment GAF ( $Mx = 57$ ) provides observer confirmation of significant functional improvements for the treatment group,  $t(224) = 12.89$ ,  $p < .0001$ .

To assess the relative efficacy of this IOP treatment to other forms of interventions, Cohen's  $d$  for correlated measures or one-group pre- and postdesigns (Dunlap, Cortina, Vaslow, & Burke,

1996) was conducted using the pre- and posttreatment Depression scale. This resulted in  $d = 1.02$  for the Depression scale, further indicating comparatively large treatment effects as defined by Cohen (1988). Similarly, the effect size correlation between pre- and posttreatment Depression scale scores was .45.

While Figure 1, Figure 2, and the meta-analytic statistics provide important information about the significant treatment effects at the group level, they do not portray the story of

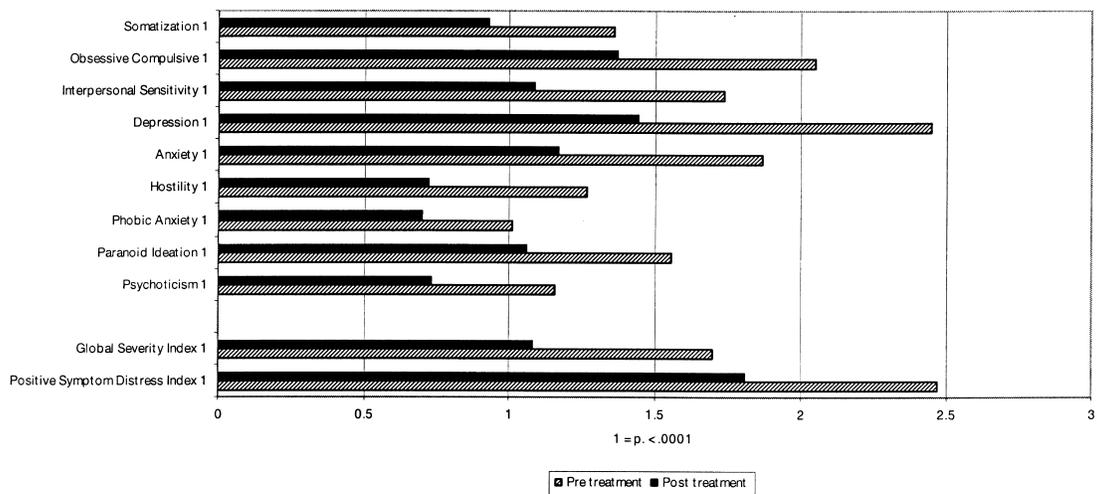


Figure 2. Intensive Outpatient Program pre- and posttreatment Symptom Check List-90—Revised  $Mx$  scores.

change at the individual level. Table 1 shows that applying the most stringent criteria of both RCI and CS change on the Depression scale, 56% ( $n = 125$ ) of the treatment group were categorized as Improved, 2% ( $n = 4$ ) as Deteriorated, and 42% ( $n = 96$ ) as Indeterminant. Table 1 also shows the effects of lowering the RCI criterion from 1.96 (95%) to 1.28 (90%) and 0.84 (80%) and requiring CS change on the Depression scale. Lowering the RCI criteria had only a slight effect. However, RCI and CS criteria can be altered to demonstrate reliable and clinically significant change using other variables. For example, examining those patients classified as Improved using various RCI depression score cutoffs, and who were discharged to a lower level of care, reflecting a real-world operationalized clinically significant change variable, the results change considerably (Table 2). That is, many of those previously classified as Indeterminant using the original RCI and CS criteria together were classified as Improved. Table 3 shows those patients who met the various RCI improvement criteria were discharged to a lower level of care and obtained posttreatment clinician-rated GAFs greater than two standard deviations of the pretreatment  $Mx$  GAF ( $Mx = 38.56$ ,  $SD = 7.46$ ). These findings approximate those observed in Table 1 and provide some real-world variables that appear to correspond to the results obtained by the stringent traditional RCI plus CS criteria. In an attempt to assess a different dimension of the client's perspective on CS, the RCI formula was applied to the Depression scale, and the normative CSQ-8 total score ( $Mx = 27.09$ ,  $SD = 4.01$ ) was utilized by applying Jacobson and Truax's (1991) two standard deviation rule (Table 4). It is observed that the improvement rates in Table 4 increased to a level consistent with those observed in Table 2. Table 5 shows that when discharge to

TABLE 1. RCI and CS Improvement Rates on SCL-90—R Depression Scale

RCI	Improved	Deteriorated	Indeterminant	Total
1.96 (95)	125 (56)	4 (2)	96 (42)	225 (100)
1.28 (90)	132 (59)	4 (2)	89 (40)	
0.84 (80)	135 (60)	4 (2)	86 (38)	

*Note.* Tingey et al. (1996) normative continuum was used. Percentages are given in parentheses. RCI = Reliable Change Index; CS = clinically significant; SCL-90—R = Symptom Check List-90—Revised.

TABLE 2. RCI Improvement Rates on SCL-90—R Depression Scale and Discharge to a Lower Level of Care

RCI	Improved	Deteriorated	Indeterminant	Total
1.96 (95)	137 (61)	7 (3)	81 (36)	225 (100)
1.28 (90)	155 (69)	18 (8)	52 (23)	225 (100)
0.84 (80)	164 (73)	25 (11)	36 (16)	225 (100)

*Note.* Percentages are given in parentheses. RCI = Reliable Change Index; SCL-90—R = Symptom Check List-90—Revised.

a lower level of care is used with clinician rated GAF scores, again using the two standard deviation rule, the Improvement rates increase considerably. Finally, Table 6 shows that the use of three CS criteria variables—GAF, CSQ-8, and discharge—to a lower level of care reduces the Improvement rates to 73% and 81%, still comparatively high.

Because the use of the traditional RCI and CS criteria (Table 1) resulted in so few in the Deteriorated group ( $n = 4$ ), statistical comparisons including this group could not be conducted. Instead, two sets of analyses were performed. In the first, those patients from the Improved and Indeterminant groups who met both RCI and CS criteria were compared. The second set of analyses included those who obtained an RCI greater than or equal to 1.96 and who were discharged to a lower level of care, thereby increasing the Deteriorated group to a large enough size to make reliable statistical comparisons.

When the Improved and Indeterminant groups that met both traditional RCI and CS criteria were analyzed, there were no significant differences between groups with respect to average age or years of education. Proportionately, these groups were not significantly different with respect to their race, employment status, presence versus

TABLE 3. RCI Improvement Rates on SCL-90—R Depression Scale, Discharge GAF Greater Than 53 and Discharge to Lower Level of Care

RCI	Improved	Deteriorated	Indeterminant	Total
1.96 (95)	117 (52)	7 (3)	101 (45)	225 (100)
1.28 (90)	129 (57)	10 (4)	86 (38)	225 (100)
0.84 (80)	134 (60)	28 (8)	63 (28)	225 (100)

*Note.* Percentages are given in parentheses. RCI = Reliable Change Index; SCL-90—R = Symptom Check List-90—Revised; GAF = Global Assessment of Functioning.

TABLE 4. RCI Improvement Rates on SCL-90-R Depression Scale and CSQ-8 Greater Than 19 (Two Standard Deviations)

RCI	Improved	Deteriorated	Indeterminant	Total
1.96 (95)	98 (61)	5 (3)	58 (36)	161 (100)
1.28 (90)	109 (68)	12 (7)	40 (25)	161 (100)
0.84 (80)	116 (72)	17 (11)	28 (17)	161 (100)

Note. Percentages are given in parentheses. RCI = Reliable Change Index; SCL-90-R = Symptom Check List-90-Revised; CSQ-8 = Client Satisfaction Questionnaire.

absence of depression, psychotropic medication prior to admission, personality disorders, medical conditions, hospital admissions, IOP readmissions, or referral to a higher level of care. The Indeterminant group had significantly higher posttreatment Depression,  $t(218) = 14.59, p < .0001$ ; GSI,  $t(218) = 10.44, p < .0001$ ; and lower GAFs,  $t(218) = 5.67, p < .0001$ , compared with the Improved group. Regarding client satisfaction, using the CSQ-8 total score, the Improved group was also significantly more satisfied than the Indeterminant group,  $t(158) = 5.17, p < .0001$ , although the Indeterminant score was not significantly different from the average CSQ-8 total score in the normative sample,  $t(3187) = 0.27, p = .80$  (Attkisson & Greenfield, 1994).

The following analyses included individuals who passed the criterion of RCI greater than or equal to 1.96 and who were discharged to a lower level of care (Table 2). Using these criteria, those in the RCI Deteriorated group were not significantly different from the Improved or Indeterminant groups with respect to age, sex, race, education, Axis I diagnoses, or number of treatment sessions. Although lessening the criteria to RCI alone increased all cell sizes and generally allowed for reliable statistical analyses of all three groups, only 1 individual from the Deteriorated group was readmitted to the IOP, so the Deteriorated group was not included in the IOP readmission comparisons. When the Improved group was

TABLE 5. GAF and Discharge to Lower Level of Care

CI	Improved	Deteriorated	Indeterminant	Total
GAF $\pm 2 SD$	174 (77)	1 (1)	48 (22)	223 (100)
GAF $\pm 1 SD$	196 (88)	1 (1)	26 (1)	223 (100)

Note. Percentages are given in parentheses. GAF = Global Assessment of Functioning; CI = confidence interval.

TABLE 6. GAF, CSQ-8, and Discharge to Lower Level of Care

CI	Improved	Deteriorated	Indeterminant	Total
GAF/CSQ-8 $\pm 2 SD$	117 (73)	1 (1)	43 (27)	161 (100)
GAF/CSQ-8 $\pm 1 SD$	131 (81)	3 (2)	27 (17)	161 (100)

Note. Percentages are given in parentheses. GAF = Global Assessment of Functioning; CSQ-8 = Client Satisfaction Questionnaire; CI = confidence interval.

compared with the Indeterminant group, however, there were no significant differences with respect to the proportion of IOP readmissions. Prior to treatment, the Improved group was more depressed,  $t(216) = 3.75, p < .001$ , than the Deteriorated group, but not the Indeterminant group ( $p > .05$ ). As one would expect, those in the Deteriorated group were worse than those in the Improved and Indeterminant groups on the measures of depression,  $t(216) = 9.52, p < .0001$ ,  $t(142) = 6.34, p < .001$ ; global distress,  $t(142) = 6.00, p < .0001$ ,  $t(216) = 8.81, p < .001$ ; and GAFs,  $t(142) = 2.91, p = .005$ ,  $t(216) = 5.63, p < .001$ , at discharge. What is most interesting, however, is that approximately half of the Deteriorated group (56%) and Indeterminant group (50%) had comorbid Axis II conditions, whereas 77% of the Improved group received Axis II diagnoses,  $\chi^2(2, N = 225) = 12.89, p < .01$ . On the other hand, 78% and 88% of the Deteriorated group and Indeterminant group had Axis III diagnoses, whereas only 45% of the Improved group had medical diagnoses,  $\chi^2(2, N = 225) = 26.56, p < .001$ .

In an effort to further clarify possible issues related to the Deteriorated group, a chart review was conducted on these 25 patients to identify any issues that may have contributed to their deterioration. Because discharge target dates were usually set approximately 2 weeks in advance, the chart review focused on any issues that developed in the 2 weeks prior to discharge. In addition to the usual stressors associated with leaving treatment, 23 of 25 (92%) patients had some unanticipated acute stressors develop in the 2 weeks prior to discharge that likely contributed to their deterioration. These included medical stressors ( $n = 7$ ; e.g., neurological symptoms, shoulder injury, illness, and pain), family stressors ( $n = 6$ ; e.g., receiving divorce petition, death

of family member, and significant family conflict), occupational stressors ( $n = 6$ ; e.g., job termination, denial of insurance benefits, and initiation of formal Equal Employment Opportunity Commission proceedings), and legal stressors (e.g., housing eviction, receiving jail sentence, harassment, and threatened murder). Additionally, 2 of the patients in the Deteriorated group entered treatment “numb,” with massive denial of symptoms related to traumatic incidents, but they became more symptomatic as treatment progressed. The CSQ-8 data on the latter 2 patients indicated that they were “very satisfied” with their treatment, although the RCI categorized them as Deteriorated.

The patients in the IOP treatment group as a whole were very satisfied with the quality of their care. While the normative CSQ-8 average total score was 27.09 (Attkisson & Greenfield, 1994), the IOP total satisfaction score was a significantly higher 28.79,  $t(3281) = 5.29, p < .0001$ . Furthermore, the patients in the Improved group (as defined in Table 1) were significantly more satisfied with their treatment (CSQ = 30.06) than those in either the Indeterminant,  $t(156) = 5.40, p < .0001$ , CSQ = 26.93, or the Deteriorated,  $t(101) = 3.06, p < .005$ , CSQ = 26.2, groups. There were no significant satisfaction differences between the Indeterminant and Deteriorated groups or between either of these groups and the normative CSQ-8 population.

## Discussion

The RCI and CS individual change rates observed in this study are impressive because they are comparable to those found in randomized controlled clinical trials (Hansen, Lambert, & Forman, 2002) and better than those found in the individual psychotherapy outcome literature using these criteria (e.g., Ankuta & Abeles, 1993; Lunnen & Ogles, 1998; Pekarik & Wolff, 1996; Piper, Joyce, Azim, & McCallum, 1998). It should also be noted that in both clinically representative and randomized control contexts, the patients are significantly less acute in that they are typically not suicidal, in need of a higher level of programmed care, and scored significantly lower on the SCL GSI scale (e.g., Elkin et al., 1989),  $t(462) = 4.60, p < .0001$ . While the majority of patients were classified as Improved under the most stringent criteria (Table 1), the majority of the remaining patients appeared to be on

their way to improvement, as evidenced by their classification as Improved when the criteria for change were altered.

Limitations to the present study are the lack of a randomized control group and the inability to partial out differential treatment effects for medication versus group therapies versus no treatment, and so forth. This is also a selected sample in that these are severely depressed individuals, typically with suicidal ideation, who are able to enter a no-harm contract. A randomized control design was beyond the scope of ordinary practice patterns. The study aimed to demonstrate the effectiveness of an IOP as it is actually practiced, including the combined effects of medication and group psychotherapies. Hence, while these issues are recognized as potential shortcomings of the study, they also reflect actual and typical real-world practice patterns. Additionally, in an effort to mitigate threats to the validity of these naturalistic findings, RCIs and CS change scores were used to further control for and measure reliable change and clinically meaningful change.

This study demonstrates that an IOP can be an effective level of care for severely depressed, complex patients involving a broad spectrum of symptoms, including clinician diagnosed personality disorders. The majority of these patients arrived in a suicidal crisis, which is typically the chief criterion utilized for an inpatient admission (Hollon, Thase, & Markowitz, 2002). However, these individuals were able to enter into a firm no-harm contract and had enough family or social support to assist in providing observation when needed. Nearly one third of this sample had an average of 1.8 previous inpatient admissions, indicating a repetitive dysfunctional pattern of behavior that had not been sufficiently altered and resulted in repeated hospitalizations. Even though these patients were more distressed than both national and locally hospitalized psychiatric patients on nearly every measure, failed to respond to medication, and many had been previously treated in the hospital, these patients were effectively treated in an IOP and were generally very satisfied with the treatment they received.

It should be noted, however, that this IOP appears to be somewhat different than the typical IOP. More specifically, in a national IOP survey, Scheifler, Kiser, and Knight (2001) found that the majority of IOPs (92%) were operated in agencies or facilities, and only 4% of IOPs were provided in a private practice setting. Similarly,

while the national average number of treatment days, dropout rates, and readmission rates were 21%, 20%, and 18%, respectively, ours were comparatively lower at 17%, 8%, and 7%, respectively. Additionally, 81% of these patients were direct admits to the IOP, whereas most IOPs tend to be utilized as step-down programs from a higher level of care rather than operating as a hospital diversion model.

An earlier report of post-IOP adjustment (Wise, in press-a) on the first 100 IOP participants found that 77% of survey respondents continued to receive some type of mental health care, 6% had attended a partial hospital program, and there were no posttreatment hospitalizations. Eighty-nine percent indicated that if the need arose, they would return to our IOP for treatment. These follow-up findings provide additional evidence that the IOP patients experienced clinically significant improvements, achieved stabilization, decreased symptoms, were diverted from the hospital, were introduced to a level of care with which they were very satisfied, and would return to the IOP if they became acutely distressed.

The finding that the traditional RCI Improved group contained a significantly higher proportion of individuals diagnosed with personality disorders is of interest. Rothschild and Zimmerman (2002) have demonstrated that personality disorders are associated with chronicity in major depression. On the other hand, Ogrodniczuk, Piper, Joyce, and McCallum (2001) have noted that "the presence of a personality disorder has [been shown to have] a negative effect, a positive effect, or no effect on outcome" (p. 111). As in most clinical settings, the current Axis II diagnoses were made by treating clinicians who did not use objective or standardized measures. Consequently, the meaning of these personality disorder diagnoses is suspect given the conflicting nature of the literature, the difficulties clinicians have in making these diagnoses reliably (e.g., Morey & Ochoa, 1989), and the absence of objective external criteria.

The fact that the traditional RCI Improved group contained significantly fewer individuals diagnosed with medical conditions than both the Indeterminant and Deteriorated groups is an important finding. Medical complications may add an additional stressor that contributes to increased reports of symptom severity and causes premature but mutually agreed upon termination, thereby contributing to stagnation (Indetermi-

nant) or Deterioration. This would seem to be corroborated by the fact that 7 of 25 (28%) patients in the Deteriorated group developed medical complications in the last 2 weeks of treatment. Other unpredictable psychosocial stressors also appeared to be related to Deterioration at discharge and warrant further investigation as potential CS variables that contribute to Improvement or Deterioration.

The debate involving the relationship between client satisfaction and symptom improvement, as well as that between RCIs and CSQ-8 scores, are noteworthy. For example, neither Lunnen and Ogles (1998) nor Pekarik and Wolff (1996) found a relationship between RCI and client satisfaction. In contrast, the traditional RCI Improved group in this study was significantly more satisfied than either the Indeterminant or the Deteriorated groups. Ankuta and Abeles (1993), Holcomb, Parker, Leong, Thiele, and Higdon (1998), and LaSala (1997) also demonstrated significant relationships between clinical improvement and client satisfaction. On the other hand, even the RCI Indeterminant and Deteriorated groups in this study were equally as satisfied as the CSQ-8 normative sample. It is certainly conceivable that client satisfaction is measuring something distinct from symptom improvement. It is equally plausible that client satisfaction may have an additive effect on Improvement and can be a useful CS variable, as indicated in Tables 4 and 6. Client perspectives regarding satisfaction can supplement the RCI criteria to demonstrate a relationship between satisfaction and improvement. In any event, it is clear that the IOP treatment group was very satisfied with the treatment they received.

This article highlights many interesting methodological problems with respect to the calculation of psychotherapy improvement rates. For example, compared with the traditional RCI and CS variables, reducing the RCI symptom improvement criterion to .80 appeared to be more consistent with clinical judgments of meaningful change, as operationalized by a transfer to a lower level of care, clinician discharge GAF ratings, and client satisfaction ratings. Additionally, when the two CS variables of clinician discharge GAF ratings and discharge to a lower level of care were used as improvement criteria (Table 5), the Improvement rates increased considerably compared with the traditional RCI plus CS criteria (Table 1). These comparisons suggest that the

use of traditional RCI and CS criteria are the most stringent and conservative measures of change. In fact, in this study using the RCI at the .84 level, in conjunction with real-world variables, appeared to capture meaningful change when other operational measures, such as clinician discharge GAFs, discharge to a lower level of care, CSQ-8, and so forth, were used to confirm the validity of this criterion.

This study also demonstrates that one reason the RCI is so stringent or conservative lies in the fact that it calculates a ceiling (greater than or equal to 1.96) that can never be exceeded. Thus, if only 60% of a sample passes the RCI criteria for symptom improvement, no further CS analyses can increase this improvement rate. In this sense, RCI can only lower CS Improvement rates. Using the two CS variables, clinician's discharge GAF ratings and the discharge of a patient to a lower level of care (Table 5) demonstrated that the use of CS variables without RCI increased the Improvement rates that were previously constrained by the RCI. The most conservative Improvement rates were obtained from the most traditional use of the RCI with two CS criteria variables (52%), and the greatest Improvement rates were obtained from the two CS criteria without RCI (89%).

Because no other IOP, partial hospitalization, or inpatient studies utilizing the SCL-90-R could be found that examined group and individual (RCI and CS) change rates, it is difficult to compare the efficacy of this program to others. Consequently, Cohen's (1988)  $d$  is of particular importance as it allows for the comparison of this treatment to other forms of interventions. It will be recalled that Cohen's  $d = 1.02$ , which would be classified as a large effect size (.2 = small, .5 = medium, and .8 = large). Shadish et al. (1997) found that in clinically representative studies utilizing control groups, the average effect size  $d$  equaled .59. Robinson, Berman, and Neimeyer (1990) found average effect sizes of .83 and .84 for individual and group psychotherapy, respectively, in controlled clinical trials with depressed outpatients. These findings indicate that the IOP treatment effects in this study were greater than those found in the average clinical setting and controlled clinical trials. Again, however, because the majority of studies reviewed were made up of individual psychotherapy studies and only one reviewed group therapy, they are of somewhat limited compara-

tive value. Compared with the present  $d = 1.02$ , selected meta-analytic medical studies have shown that the effect of bypass surgery on angina equals .80, drug treatment on arthritis ranges from .45 to .77, and the effect of Cyclosporine on organ rejection equals .39 (Dunlap et al., 1996). Similarly, the obtained correlation between IOP pre- and posttreatment ( $r = .45$ ) is greater than the relationships between sleeping pills and improvement in insomnia ( $r = .30$ ), accuracy of home pregnancy tests ( $r = .38$ ), Viagra and improved male sexual performance ( $r = .38$ ), and computed tomography imaging results used to detect lymph node metastases in cervical cancer ( $r = .41$ ; Myer et al., 2001). The obtained  $r = .45$  is similar to the relationships between single photon emission computed tomography imaging and identification of coronary artery disease ( $r = .46$ ), bone scan results detecting osteomyelitis ( $r = .48$ ), and clinical exam and routine lab tests to detect metastatic lung cancer ( $r = .48$ ; Myer et al., 2001).

In summary, the patients treated in this IOP started treatment significantly more symptomatic on nearly every symptom measure and significantly more distressed than local and national hospitalized psychiatric patients. Yet, while being treated at home, they demonstrated significant symptom improvement across every symptom scale and overall reported high levels of satisfaction. Follow-up data indicated that these patients maintained their treatment gains and levels of client satisfaction. At the individual level, using the traditional criteria of RCI symptom change scores (1.96) and moving from one functional distribution to another (CS), 56% Improved; whereas using RCI symptom change scores and discharge to a lower level of care, 61%–73% Improved. On the other hand, when the CS variables of GAF scores and discharge to a lower level of care were utilized, 77% and 88% Improved, respectively, and when the CSQ was added to these, 73% and 81% met the Improved criterion. The meta-analytic data provided additional support and clarity for interpreting these findings and indicate that this treatment is better than and comparable to many medical diagnostic tests and interventions. In fact, the meta-analytic data indicate that the average IOP patient would have a better outcome than 84% of untreated individuals, and approximately 72% of those treated in this IOP would be predicted to have successful outcomes ( $d = 1.02$ ; Wampold,

2001). It should be noted that the meta-analytic prediction that approximately 72% of those treated would be expected to have successful outcomes is almost the midpoint between the most extreme Improvement rates found and is also very similar to that observed when RCI = .84 was used (Tables 2 and 4). It should be clear, however, that no single measure of psychotherapy outcome can display the full and complex picture of clinical, functional, client, and clinician views of improvement.

## References

- American Association of Community Psychiatrists. (1998). *Level of care utilization system for psychiatric and addiction services* (Version 2.0). Pittsburgh, PA: Author.
- American Society of Addiction Medicine. (1991). *ASAM patient placement criteria for the treatment of psychoactive substance use disorders*. New York: Author.
- ANKUTA, G., & ABELES, N. (1993). Client satisfaction, clinical significance and meaningful change in psychotherapy. *Professional Psychology, 24*, 70–74.
- 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* (pp. 402–422). Hillsdale, NJ: Erlbaum.
- ATTKISSON, C., & ZWICK, R. (1982). The Client Satisfaction Questionnaire: Psychometric properties and correlations with service utilizations and psychotherapy outcome. *Evaluation and Program Planning, 5*, 233–237.
- CHORPITA, B. F. (2001). Reflections on clinical significance: What do our best treatments accomplish and how can we best find out? *Clinical Psychology: Science and Practice, 8*(4), 451–454.
- COHEN, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Hillsdale, NJ: Erlbaum.
- DEROGATIS, L. (1983). *SCL-90–R: Administration, scoring and procedures manual—II*. Towson, MD: Clinical Psychometric Research.
- DUNLAP, W. P., CORTINA, J. M., VASLOW, J. B., & BURKE, M. J. (1996). Meta-analysis of experiments with matched groups or repeated measures. *Psychological Methods, 1*, 170–177.
- ELKIN, I., SHEA, T., WATKINS, J., IMBER, S., SOTSKY, ST., COLLINS, J., et al. (1989). National Institute of Mental Health treatment of depression collaborative research program. *Archives of General Psychiatry, 46*, 962–971.
- GOLDSTEIN, J., & HORGAN, C. (1988). Inpatient and outpatient psychiatric services: Substitutes or complements? *Hospital and Community Psychiatry, 39*, 632–636.
- HANSEN, N., LAMBERT, M., & FORMAN, E. (2002). The psychotherapy dose–response effect and its implications for treatment. *Clinical Psychology: Science and Practice, 9*(3), 329–343.
- HOLCOMB, W., PARKER, J., LEONG, G., THIELE, J., & HIGDON, J. (1998). Customer satisfaction and self-reported treatment outcomes among psychiatric inpatients. *Psychiatric Services, 49*(7), 929–934.
- HOLLON, S., THASE, M., & MARKOWITZ, J. (2002). Treatment and prevention of depression. *Psychological Science in the Public Interest, 3*(2), 39–77.
- HOWARD, K., MORAS, K., BRILL, P., MARTINOVICH, Z., & LUTZ, W. (1996). Evaluation of psychotherapy: Efficacy, effectiveness and patient progress. *American Psychologist, 51*, 1059–1064.
- 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*, 12–19.
- KAZDIN, A. E. (1999). The meanings and measurement of clinical significance. *Journal of Consulting and Clinical Psychology, 67*, 332–339.
- KISER, L., LEFKOVITZ, P., KENNEDY, L., & KNIGHT, M. (1996). *The continuum of ambulatory behavioral healthcare services: A position paper from the Association for Ambulatory Behavioral Healthcare*. Washington, DC: Association for Ambulatory Healthcare.
- KISER, L., TRACHTA, A., BRAGMAN, J., CURLEY-SPADARO, K., COOKE, J., RAMSTAND, W., & FITZHUGH, K. (1998). *Standards and guidelines for Level II services*. Washington, DC: Association for Ambulatory Healthcare.
- LASALA, M. (1997). Client satisfaction: Consideration of correlates and response bias. *Families in Society, 78*(1), 54–64.
- LUNNEN, K., & OGLE, B. (1998). A multiperspective, multi-variable evaluation of reliable change. *Journal of Consulting and Clinical Psychology, 66*, 400–410.
- MOREY, L., & OCHOA, E. (1989). An investigation into the adherence to diagnostic criteria: Clinical diagnosis of the DSM III personality disorders. *Journal of Personality Disorders, 3*, 180–192.
- MYER, G. J., FINN, S. E., EYDE, L. D., KAY, G. G., MORELAND, K. L., DIES, R. R., et al. (2001). Psychological testing and psychological assessment: A review of evidence and issues. *American Psychologist, 56*, 128–165.
- NORMAN, G., & STREINER, D. (2000). *Biostatistics: The bare essentials*. Hamilton, Ontario, Canada: B. C. Decker.
- OGRODNICZUK, J., PIPER, W., JOYCE, A., & MCCALLUM, M. (2001). Using axis II information to predict outcome in short-term individual psychotherapy. *Journal of Personality Disorders, 15*(2), 110–122.
- PEKARIK, G., & WOLFF, C. (1996). Relationship of satisfaction to symptom change, follow-up adjustment, and clinical significance. *Professional Psychology: Research and Practice, 27*, 202–208.
- PIPER, W., JOYCE, A., AZIM, H., & MCCALLUM, M. (1998). Interpretive and supportive forms of psychotherapy and patient personality variables. *Journal of Clinical and Consulting Psychology, 66*, 558–567.
- PIPER, W., ROSIE, J., JOYCE, A., & AZIM, H. (1996). *Time limited day treatment for personality disorders: Integration of research and practice in a group program*. Washington, DC: American Psychological Association.
- ROBINSON, L., BERMAN, J., & NEIMEYER, R. (1990). Psychotherapy for the treatment of depression: A comprehensive review of controlled outcome research. *Psychological Bulletin, 108*, 30–49.
- ROTHSCHILD, L., & ZIMMERMAN, M. (2002). Personality disorders and the duration of depressive episode: A retrospective study. *Journal of Personality Disorders, 16*(4), 293–303.
- SCHEIFLER, P., KISER, L., & KNIGHT, M. (2001). *Benchmarks and critical indicators for intensive outpatient programs:*

- Results of a national survey of IOPs.* Alexandria, VA: Association for Ambulatory Behavioral Healthcare.
- SHADISH, W., MATT, G., NAVARRO, A., SIEGLE, G., CRITS-CHRISTOPH, P., HAZELRIGG, M., et al. (1997). Evidence that therapy works in clinically representative conditions. *Journal of Consulting and Clinical Psychology, 65*, 355–365.
- SMITH, M., GLASS, G., & MILLER, T. (1980). *The benefits of psychotherapy.* Baltimore, MD: Johns Hopkins University Press.
- TINGEY, R. C. (1989). Assessing clinical significance: Extension in methods and application to the SCL-90—R. *Dissertation Abstracts International, 50*, 04B.
- TINGEY, R., LAMBERT, M., BURLINGAME, G., & HANSEN, N. (1996). Assessing clinical significance: Proposed extensions to method. *Psychotherapy Research, 6*(2), 109–123.
- WAMPOLD, B. E. (2001). *The great psychotherapy debate: Models, methods and findings.* Hillsdale, NJ: Erlbaum.
- WASHTON, A. (1997). Evolution of intensive outpatient treatment as a “legitimate” treatment modality. *Journal of Addictive Diseases, 16*(2), xxi–xxvii.
- 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*, 412–417.
- WISE, E. A. (2003). A private practice model for intensive outpatient services. *Psychiatric Services, 54*(4), 572–573.
- WISE, E. A. (in press-a). Empirical validation of a mental health intensive outpatient program in a private practice setting. *American Journal of Orthopsychiatry.*
- WISE, E. A. (in press-b). A review of clinical significance, reliable change and recommendations for future directions. *Journal of Personality Assessment.*