Abstract

To be useful in assessing treatment efficacy, an outcome measure must be sensitive to changes in ratings of symptoms and psychosocial functioning over time and to differences in change across consumers and be interpretable as to the extent of clinical improvement. Results are presented regarding the sensitivity of the BASIS-32 to changes in ratings of symptoms and psychosocial functioning among 1188 persons with severe and persistent mental illness after receiving a variety of clinical and psychosocial interventions in one behavioral health center over a 3-year period. Utilizing the Reliable Change Index, evidence was found for the sensitivity of the BASIS-32 to detect statistically reliable and clinically significant changes on 2 subscales of psychosocial functioning (relations with others and daily living/role functioning) and on the Depression/Anxiety subscale. For impulsive/addictive behavior and psychosis, and the overall BASIS-32 score, substantially less reliable change could be detected, although statistically significant differences were present.

Introduction

Outpatient behavioral health programs need to demonstrate treatment efficacy through improvement in their consumers’ symptoms and ability to function effectively in a community setting. To meet this goal, psychometrically sound outcome measures must be used to evaluate changes in a consumer’s symptoms and psychosocial functioning over time or during the course of treatment. Many outcome instruments have demonstrated reliability and validity, but only a few also exhibit adequate sensitivity to change over time, that is, the ability to detect and differentiate positive or negative changes. Furthermore, conventional statistical comparisons between groups of consumers tell very little about the efficacy of any treatment intervention or its clinical significance because (1) the tests provide no information on the variability of response to treatment within the sample, and (2) a statistically significant treatment effect may not clearly relate to the clinical significance of the effect, that is, “its ability to meet standards of efficacy set by consumers, clinicians, and researchers.” A variety of criteria have been suggested as to what these clinical significance standards should be: a high percentage of consumers improving; a level of change that is recognizable by peers and significant others; an elimination of or reduction in the presenting problem; or changes that significantly
improve one’s ability to function effectively in daily living. To be a useful tool in assessing treatment efficacy, then, an outcome measure must be sensitive to change over time and variability in change across consumers and be interpretable in terms of clinical improvement or recovery.

The BASIS-32 is designed to assess a consumer’s perspective on his or her level of difficulty with a broad range of symptoms and problems over the past week. Difficulty is rated on a 5-point scale from 0 (no difficulty) to 4 (extreme difficulty). The BASIS-32 yields 5 subscale scores (relationship to self/others; daily living/role performance; depression/anxiety; impulsive/addictive behavior; psychosis), and an overall average score concerning the consumer’s symptoms and functioning.

Previous investigators have adequately demonstrated most of the psychometric properties of the BASIS-32 instrument, including its reliability and validity, in both inpatient and outpatient settings. In an outpatient sample similar to the one used in this analysis, mean scores of 1.90 on 3 of the subscales (relationship to self/others, daily living/role functioning, and depression/anxiety) suggested that consumers expressed moderate difficulty in these areas, close to the midpoint of the scale, whereas less difficulty was reported for impulsive/addictive behavior (mean = 0.85) and psychosis (mean = 0.68). No ceiling effects (<1% of consumers reporting extreme difficulty) were found on any subscale, but the Impulsive/Addictive and Psychosis subscales evidenced a floor effect, with 15% and 33% (respectively) of consumers reporting no difficulty. At follow-up, mean scores ranged from 1.51 to 1.60 on 3 of the subscales (relationship to self/others, daily living/role functioning, and depression/anxiety), whereas less difficulty was reported for impulsive/addictive behavior (mean = 0.61) and psychosis (mean = 0.50).

However, previous results concerning its sensitivity to change over time consist mainly of statistical comparisons of pre-post intervention BASIS-32 scores. Additional data regarding change over time, variability in change across consumers, and clinical improvement interpretability are needed to address this issue more precisely.

The Reliable Change Index

If clinically significant change is conceptualized as consumers entering therapy as part of a dysfunctional population and departing therapy as no longer part of that population, such change would be inferred if a post-treatment score falls within (or closer to the mean of) the functional population on the variable of interest. In the absence of data from a normative sample, a cutoff point for “recovery” (versus improvement) could be established by using a “2 standard deviations” solution. The Reliable Change (RC) index was developed to determine statistically reliable change that is also clinically significant. That is, the consumer demonstrates better scores on the follow-up assessment, which are 2 standard deviations from the pretest score. This degree of improvement or recovery is defined as “true change.”

A modification and refinement of the RC index was subsequently suggested to improve on the pre-post difference score by taking into account regression to the mean owing to measurement unreliability. This modification resulted in the RC index utilizing Improved pre-post Difference scores (RCID), which is standardized to enhance comparability of research findings across studies with respect to the number of consumers who have actually changed, thus protecting against type II errors. The RCID index for approximating true difference scores is calculated by using the formula:

$$ RC_{ID} = \frac{(x_2 - x_1) r_{DD} + (M_2 - M_1)(1 - r_{DD})}{\sqrt{S_{E(1)}^2 + S_{E(2)}^2}} $$

where RCID = the RC index utilizing Improved Difference (ID) scores (when significant, the absolute value should be >1.96); x1 = pre score; x2 = post score; M1 = mean of pre scores; M2 = mean of post scores; SE(1) = 1 - rxx(1) prevalence of pre scores; SE(2) = 1 - rxx(2) prevalence of post scores; S_{E(1)}^2 = s^2(1 - rxx(1))^2 = standard error of measurement of pre scores; S_{E(2)}^2 = s^2(1 -