# Is Symptomatic Improvement in Clinical Trials of Cognitive-Behavioral Therapy for Psychosis Clinically Significant?

**BRANDON A. GAUDIANO, PhD**

**Although cognitive-behavioral therapy (CBT) is becoming increasingly popular as an adjunctive treatment for psychosis, few studies to date have examined the clinical (in contrast to statistical) significance of treatment gains using standardized methods. The aim of the current study was to investigate the clinical significance of symptomatic reductions reported in trials of CBT for schizophrenia and related disorders using standardized group methods of analysis. An electronic literature search identified 12 studies that met the inclusion criteria of being randomized, controlled trials that compared CBT to routine care alone or to another comparison treatment. The analysis involved the following steps. First, reliable change on symptom measures was examined. Next, the proportion of patients in each study estimated to show clinically significant symptomatic reductions (i.e., two standard deviations) was calculated. When both post-treatment and follow-up data were considered, 42% of CBT conditions compared with only 25% of comparison conditions demonstrated reliable change on at least one psychotic symptom measure per study. Proportions of clinically significant symptomatic improvement in studies showing reliable change were similar between CBT and comparison conditions. Due to the adjunctive nature of CBT for schizophrenia and the limits imposed by the evaluation of group datasets, results of the current study are considered promising but preliminary. Future trials should examine clinical significance using similar standardized methods within studies, as well as broader functional outcome measures, to provide a clearer picture of the benefits derived from this type of intervention.** *(Journal of Psychiatric Practice* **2006;12:11–23**)

KEY WORDS: cognitive-behavioral therapy, cognitive therapy, behavior therapy, clinical significance, outcomes, schizophrenia, psychotic disorders, psychosis, randomized controlled trials, measurement

Schizophrenia and other psychotic disorders are chronic and debilitating conditions associated with high levels of psychiatric and medical comorbidity.[1,2] Even with the latest advances in pharmacologic treatment, many individuals continue to experience residual symptoms and poor quality of life.[3,4] Furthermore, rates of non-adherence to medication regimens have been noted to be as high as 89% in some samples.[5] Traditionally, psychotherapy has not been a primary intervention for this population as early attempts resulted in reports of null or iatrogenic effects.[6–8] However, there has been a renewed interest in psychosocial approaches for schizophrenia and related conditions in recent years that, in contrast to their historical counterparts, tend to be coping-focused, goal-oriented, time-limited, and symptom-related.[9] It is important to emphasize that the goal of cognitive-behavioral therapy (CBT) is not full remission of severe mental illness. Adjunctive CBT endeavors to produce clinically meaningful improvements in symptoms and functioning beyond those achieved through pharmacotherapy alone.

In recent years, CBT has become increasingly popular as an adjunctive treatment for psychotic disorders, and

GAUDIANO: Brown Medical School & Butler Hospital.



Please send correspondence and reprint requests to: Brandon A. Gaudiano, PhD, Psychosocial Research Program, Butler Hospital, 345 Blackstone Boulevard, Providence, RI 02906.

Acknowledgements: I thank Dr. Peter Trower for providing additional data for analyses. I also thank Dr. Joseph McGlinchey and the anonymous reviewers for providing helpful feedback on an earlier draft of this manuscript.

randomized controlled trials (RCTs) have been published documenting its efficacy. CBT for psychosis is typically an integrated approach that contains various components, including psychoeducation about illness, goal setting, symptom monitoring, cognitive restructuring, skills training, and homework assignments.[10] CBT has been adapted to be delivered in inpatient and outpatient settings, in group and individual formats, as briefer and longer-term interventions, in the acute and residual phases of illness, in the prevention of psychosis in high-risk groups, and by less experienced clinicians.[11,12]

Several groups have conducted meta-analytic reviews of the research in this area, all suggesting the safety and efficacy of CBT for psychosis.[13–16] For example, in an analysis of seven randomized trials, Gould et al. reported effect sizes in the large range (*ES* = 0.65) for CBT relative to comparison conditions on psychotic symptom measures.[14] Analyses of follow-up data showed that patients receiving CBT continued to improve after treatment (*ES* = 0.93). However, these early trials have been plagued by methodological shortcomings, such as small sample sizes, unblinded assessors, poor treatment fidelity, limited sample demographics, selective reporting of outcome measures, and inadequate comparison groups.[13]

Not only are there questions about the methodological rigor of early trials of CBT for psychosis, but Gould et al.[14] also argue specifically for the need for clinical significance analyses of treatment gains observed in outcome trials. In general, clinical significance refers to treatment effects that are both reliable and meaningful,[17–18] in contrast to statistical significance, which determines whether or not an observed difference is likely to be due to chance. Statistical significance is influenced by many variables in addition to the magnitude of the difference, so that small and clinically insignificant gains can be statistically *significant* if sample sizes are large enough, and vice versa.

Traditionally, clinical significance has been defined as representing a return to "normal functioning" following treatment.[17] Jacobson and Truax[19] proposed three possible criteria to measure clinical significance: a) scores following treatment that fall outside the range of the dysfunctional population in the direction of normality; b) scores following treatment that fall within the range of the normal population; or c) scores following treatment that fall closer to the mean of the normal than of the dysfunctional population. Normative comparison based on equivalency testing has been developed for use with group datasets.[18] However, a return to normal functioning is not an appropriate or possible criterion to apply to some clinical populations. For example, one would not expect most individuals with severe, chronic mental disorders, such as schizophrenia or autism, to return to normal functioning post-treatment.[17–20] In these cases, it is deemed more appropriate to define clinically significant change according to a standardized level of improvement (e.g., two standard deviations change) following treatment (i.e., cutoff score *a*[17]). Such criteria ensure that symptomatic reductions are large enough to show recognizable and obvious improvements compared to baseline functioning that would be similar to normative comparisons in more tractable conditions.

However, examining *absolute* change over the course of treatment does not take into account the degree to which the gains can be attributed to the intervention itself, in contrast to factors such as regression to the mean and measurement error. Jacobson et al.[17] attempted to redress this problem by developing a criterion called the Reliable Change Index (RCI). The RCI is a formula that takes into account the reliability of an assessment instrument to determine if treatment gains exceed the error attributable to measurement. If the RCI exceeds a statistical critical value, then scores are deemed to reflect "real" differences. Although there have been many modifications to the RCI formula over the years (see Hafkenscheid[21]), these changes have been shown to produce only minimal improvements in the measure and the original formula typically is preferred to ensure comparability between studies.[22] Although originally developed for evaluating the gains of individual patients, clinical significance methods have been used to examine group data from an aggregated sample of studies, in a manner similar to what is done in meta-analysis.[18,23,24]

To date, few trials of CBT for psychosis have reported the clinical significance of treatment gains according to Jacobson and Truax's[19] criteria (exceptions include studies by Startup et al.[25] and Gaudiano and Herbert[26]). Furthermore, definitions of clinically significant improvement have varied widely in most studies that have attempted to examine these effects. Definitions have included symptomatic reductions ranging from 20% to 50%, arbitrarily and idiosyncratically defined cut-off scores, and unstandardized methods for computing reliable change.[27–30] For example, a study by Tarrier et al.[31,32] defined clinically significant improvement as a 50% reduction in symptoms at post-treatment, but as a 20% reduction in symptoms at follow-up (which is more consistent with definitions of "responder" status in pharmacotherapy trials).

The aim of the present study was to examine the clinical significance of treatment gains in trials using CBT to treat psychotic disorders, most commonly schizophrenia, based on standardized methods. Clinical significance in the current study was defined according to the *degree of symptomatic improvement*. This differs somewhat from the traditional definition representing a return to normal functioning, which was not deemed an appropriate standard to apply to the current population. This strategy is consistent with analyses conducted in several individual trials in this area.[25–27,31] However, analyses in the current study were conducted by pooling data from multiple clinical trials and standardized methods were used for defining clinically significant symptomatic reductions across studies.

A literature search was conducted to identify relevant studies for inclusion. Next, two interrelated criteria were used to assess clinical significance. First, the RCI (adapted for group data by Sheldrick et al.[18]) was computed for each standardized, interviewer-rated psychotic symptom measure used in the studies. Second, for measures showing reliable change, clinically significant improvement was defined as improvement of two standard deviations (SDs) in the direction of symptomatic reduction. Using methods for group data developed by Hageman and Arrindell,[33] the proportion of subjects in each study estimated to meet criteria for clinically significant symptomatic improvement was calculated. Finally, results between CBT and control conditions were compared to determine if CBT added to the effectiveness of treatment as usual.

## METHOD

### Selection of Studies

Relevant studies were identified by conducting PsychINFO and MEDLINE searches using combinations of the key words schizophrenia, psychotic disorders, psychosis, treatment, cognitive therapy, cognitive behavior therapy, and behavior therapy. Furthermore, the reference lists of meta-analyses and review papers were examined.[9,11–16,34–39] A total of 48 reports were identified and evaluated for possible inclusion in the analyses.

What is described as CBT for psychosis often differs considerably from study to study.[11] Therefore, for inclusion in the current study, CBT was broadly defined as an intervention that contained both explicit cognitive (e.g., cognitive restructuring) and behavioral (e.g., skills training, behavioral experiments) components that targeted core symptoms of psychotic disorders (i.e., positive and/or negative symptoms). In addition, only data from RCTs were included in the analyses due to the methodological problems historically found in outcome studies in this area.[13] Of the 48 studies identified, 36 studies were excluded for the following reasons: case study designs ($n$ = 12); non-randomization to conditions ($n$ = 4); insufficient data/non-standardized scoring methods ($n$ = 5); use of subthreshold diagnostic groups ($n$ = 2); provision of "nontraditional" CBT (e.g., mindfulness/acceptance-based strategies) ($n$ = 5); or small-scale pilot studies ($n$ = 8). If necessary information was missing from a study that otherwise met inclusion criteria, an attempt was made to contact the first author via email to obtain the missing data. Additional data were obtained from the author of one study.

Of the 48 reports evaluated, 12 studies met the inclusion/exclusion criteria and were included in the analyses. Although the final sample represented a smaller subsample of the total identified, it was deemed important to limit analyses to methodologically stringent trials that had more similarities than differences due to the general heterogeneity of trials in this area. The characteristics of these studies are shown in Table 1. Diagnoses were based on DSM or ICD criteria for schizophrenia and other psychotic disorders. Routine care plus CBT was compared with routine care plus a non-CBT intervention in 42% of trials. Individual treatment was provided in 92% of studies and patients received an average of 19 sessions. Most studies treated exclusively outpatient samples (75%). Follow-up assessments of symptoms were conducted in 58% of trials ($M$ = 7 months). The aggregated data set represented a total of 687 patients treated with CBT. Table 2 provides a summary description of the different types of strategies and techniques employed in the selected trials.

### Measures

Only standardized outcome measures assessing psychotic and related psychiatric symptoms were included in the analyses. Analyses were limited to interviewer-rated symptom instruments, because this is the typical type of assessment tool used for the population under study.[50] Furthermore, only outcome measures and subscales that possessed adequate evidence of reliability from psychometric studies were included, as this information was necessary to compute the RCI. Studies reporting outcomes based on unstandardized composite score calculations or individual-item analyses were excluded because the reliability and validity of these

**Table 1. Characteristics of selected clinical trials**

| *Trial* | *CBT n* | *Format* | *# of Sessions/ Length of Treatment* | *Setting/ Diagnoses* | *Follow-Up Period** | *Comparator* |
|---|---|---|---|---|---|---|
| 1. Barrowclough et al.[40] Haddock et al.[41] | 17 | I + F | I = 29, F = 16 9 months | OP SCZ + SUD | 9 months | RC |
| 2. Bechdolf et al.[42] | 40 | G | 16 2 months | IP to OP PD | 6 months | PE |
| 3. Durham et al.[43] | 22 | I | 20 9 months | OP SCZ/SCZA/DD | 3 months | RC or ST |
| 4. Gumley et al.[44] | 72 | I | *Med* = 10 12 months | OP PD | — | RC |
| 5. Kuipers et al.[27,45] | 27 | I | 19 9 months | OP SCZ/SCZA/DD | 9 months | RC |
| 6. Lewis et al.[46] | 101 | I | *M* = 16 70 days | IP to OP PD | — | RC or ST |
| 7. Pinto et al.[47] | 19 | I | 24 6 months | OP SCZ | — | ST |
| 8. Rector et al.[29] | 24 | I | 20 6 months | OP SCZ/SCZA | 6 months | RC |
| 9. Sensky et al.[30] | 46 | I | *M* = 19 9 months | OP SCZ | 9 months | ST |
| 10. Startup et al.[25] | 47 | I | *M* = 13 12 months | IP to OP SCZ/SCZA/SCZF | — | RC |
| 11. Trower et al.[48] | 15 | I | *Med* = 16 6 months | OP PD | 6 months | RC |
| 12. Turkington et al.[49] | 257 | I | 6 3 months | OP SCZ | — | RC |

*Legend: DD = delusional disorder; F = Family; G = Group; I = individual; IP = inpatient; M = mean; Med = median; OP = outpatient; PD = psychotic disorders; PE = psychoeducation; RC = routine care only; SCZ = schizophrenia; SCZA = schizoaffective disorder; SCZF = schizophreniform disorder; ST = supportive therapy; SUD = substance use disorder;*

**Follow-up period post-treatment; only studies that reported psychotic symptom measures at follow-up are shown.*

procedures are unknown. If a study reported both standardized and non-standardized scoring for outcomes measures, only data from the standardized subscales were analyzed. Changes from pre- to post-treatment and pre-treatment to follow-up (if available) were analyzed. If multiple follow-up periods were reported, the longest period was used in the analyses. The following measures met the above criteria and were analyzed when available: Brief Psychiatric Rating Scale (BPRS),[51,52] Comprehensive Psychopathological Rating Scale (CPRS),[53] Positive and Negative Syndrome Scale (PANSS),[54] Psychotic Symptom Rating Scales (PSYRATS),[55] and Scale for the Assessment of Negative Symptoms/Positive Symptoms (SANS/SAPS).[56,57]

### Statistical Analyses

Analyses were conducted based on the models proposed by Sheldrick et al.[18] and Hageman and Arrindell[33] for evaluating clinically significant change using group data. Results from these analyses are typically reported using descriptive statistics. Symptomatic reductions

**Table 2. Description of cognitive-behavioral interventions used in selected clinical trials**

Barrowclough et al.[40] Integration of motivational interviewing, individual cognitive-behavioral, and family approaches

Bechdolf et al.[42] "Formulation, guided recovery, symptom monitoring, exposure/focusing strategies for managing voices, hypothesis/reality testing, reframing attributions, relational responding, coping strategy enhancement, distraction techniques, role play, anxiety management, depression and self-esteem work, medication compliance/motivational interviewing, schema work, relapse prevention and getting well strategies" (p. 23)

Durham et al.[43] "Initial emphasis on engagement, education and building a therapeutic alliance; functional analysis of key symptoms, leading to a formulation and problem list; development of a normalizing rationale for the patient's psychotic experiences, exploration and enhancement of current coping strategies; acquisition of additional coping strategies for hallucinations and delusions; and focus on accompanying affective symptomatology using relaxation training, personal effectiveness training and problem-solving as appropriate" (p. 305)

Gumley et al.[44] Engagement, psychoeducation, individualized case formulation, symptom monitoring, generation of alternative beliefs, behavioral experiments, coping skills, medication adherence

Kuipers et al.[27] Building alliance, developing coping skills (e.g., activity scheduling, relaxation, skills training), collaborating on a shared model of illness, gentle challenging of delusional beliefs and beliefs about hallucinations, modifying dysfunctional schemas, preventing relapse

Lewis et al.[46] Assessment and engagement, coping skills, psychoeducation, problem formulation, generating alternative hypotheses for hallucinations and delusions, identifying precipitating factors, symptom monitoring, distress reduction techniques

Pinto et al.[47] Engagement, alliance building, psychoeducation, case formulation, social skills training, modeling, rehearsal, positive reinforcement, in vivo exercises, homework, self-care, medication adherence, interpersonal problem-solving, recreation planning, resource management, distraction, relaxation, generation of alternative explanations for delusions and hallucinations, reality testing, symptom monitoring, coping strategies, relapse prevention

Rector et al.[29] Assessment and engagement, normalization of symptoms, psychoeducation, problem formulation, self-monitoring, homework, guided discovery for delusions and hallucinations, activity scheduling, assertiveness training, behavioral experiments, relapse prevention

Sensky et al.[30] Engagement, analysis of the antecedents of psychotic symptoms, normalization of symptoms, shared case formulation, Socratic questioning of delusions, voice diaries, reattribution of the causes of voices, classification of neologisms, thought linkage, activity scheduling

Startup et al.[25] Adapted for inpatients from the Kuipers et al.[27] protocol

Trower et al.[48] Assessment, formulation, Socratic questioning, behavioral experiments, generation of alternative beliefs for hallucinations, behavioral experiments

Turkington et al.[49] "Assessment and engaging, developing explanations, case formulation, symptom management, adherence, working with core beliefs and relapse prevention" and psychoeducational booklets for patients and caregivers covering "treatment, self-care and lifestyle, leisure time and relationships, drug and alcohol advice, symptom management and sources of help" (p. 524)

first must be deemed reliable before any observed improvement can be classified as clinically meaningful. Therefore, the RCI for each study measure (including relevant subscales) was computed (see Appendix 1). In order to compute the RCI, an estimate of the standard error of measurement (*SE*) is required. There are various methods for computing the *SE*. Although traditionally used, test-retest reliability coefficients were available only for some of the outcome measures used in the selected studies. However, other indices of reliability can be used instead if test-retest reliability estimates are unavailable.[58,59] Inter-rater reliability coefficients were available for all measures and were therefore used in the following analyses to maintain consistency and comparability between studies and across measures. As all measures analyzed in the current study were interview-based, inter-rater reliability was judged to be an appropriate standard to use for computing reliable change. If available, the inter-rater reliability coefficients reported in the actual outcome studies were used in the analyses. However, if unavailable, inter-rater reliability coefficients reported in psychometric studies of

the respective measures were used (see Appendix 2). If the RCI exceeds the critical value of 1.96 ($p < 0.05$), the observed improvement represents a statistically significant change attributable to more than known measurement error.[19]

As discussed previously, normative comparisons and equivalency testing for aggregated datasets are not easily applied to certain populations, including patients with severe mental illness, since these data are not typically available and a return to normal functioning would not be required to demonstrate a clinically significant improvement. Understandably, there is no "gold" standard for determining clinically significant improvement in studies of CBT for psychosis, so that various criteria have been applied in previous studies. For example, Morrison et al.[28] analyzed clinical significance by attempting to identify scores that represented "meaningfully low" symptom levels, but their definitions were somewhat arbitrary and varied across measures. In the current study, it was deemed more important to establish a standard that could be applied across studies, since the trials included in the analyses used a variety of outcome measures to assess various symptom constructs. Also, some trials of CBT for psychosis have defined clinical significance as a reduction of at least 20%–50% in the severity of symptoms.[27,30,31,43] However, more sophisticated methods are needed, because the percent reduction criterion possesses known methodological shortcomings and poses difficulties in interpretation.[64] Finally, although it may be preferable to examine objective indices or variables that more naturally correspond to functional improvement,[65] measures of such constructs were not routinely reported in the trials, and results were therefore limited to measures assessing symptomatic improvement.

Jacobson and Truax's[19] cutoff score "*a*" for determining clinical significance was used in the current study, which is defined as two standard deviations or greater change in the direction of symptomatic improvement following treatment. Hageman and Arrindell[33] adapted Jacobson and Truax's criteria for examining group data to obtain an estimate of the percentage of patients in the sample meeting criteria for both reliable change and clinically significant change, based on separate formulas. However, Hageman and Arrindell's calculation of reliable change could not be used in the current investigation because it requires the calculation of the pre- to post-test correlation coefficients for each study measure, which were unavailable. Therefore, Sheldrick et al.'s[18] "box score" method for calculating reliable change in group data was used instead, which specifies whether or not the group means show reliable change. This method has been used successfully to assess reliable change in treatment studies of conduct disorder[18] and bulimia.[24] If a study measure showed evidence of reliable change based on Sheldrick et al.'s method, Hageman and Arrindell's method for estimating the percentage of subjects meeting criteria for clinically significant change based on Jacobson and Truax's cutoff score $a$ was calculated (see Appendix 3). Estimated group percentages of subjects showing clinically significant change were obtained by computing $z$-scores and comparing them to the standard normal distribution table, $F(z)$.

## RESULTS

### CBT Conditions

***Reliable change.*** Table 3 shows results from the RCI and clinically significant improvement analyses. When post-treatment data alone were considered, routine care plus CBT conditions (referred to simply as "CBT" hereafter) in 33% (4/12) of the studies showed evidence of reliable change on at least one measure. When post-treatment or follow-up data were included as available, reliable change rates rose to 42% (5/12). Of the total 31 measures/subscales analyzed, 29% ($n = 9$) showed reliable change at post-treatment and 32% ($n = 10$) when follow-up data were included. Rates of reliable change in studies at post-treatment or follow-up based on the most common assessment measures used were as follows: PANSS = 17% (1/6), SANS/SAPS = 33% (1/3), BPRS = 67% (2/3), and PSYRATS = 100% (2/2). Rates of reliable change based on the primary assessment of positive versus negative symptoms were 46% (6/13) and 13% (1/8), respectively. Correlation analyses did not reveal significant associations between reliable change and treatment setting (inpatient versus outpatient), treatment format (group versus individual), length of follow-up period, or number of sessions provided (all $p$s = n.s.). Although power was somewhat limited to detect significant associations, coefficients in these analyses were small in magnitude (0.07–0.29).

***Clinically significant improvement.*** Across all measures, the average proportion of patients estimated to achieve clinically significant improvement (i.e., greater than two SDs) was 16% at post-treatment or follow-up. When only those measures showing reliable change were analyzed, the proportion of patients estimated to show clinically significant change was 48%.

**Table 3. Reliable change and percentage of individuals showing clinically significant improvement for CBT and comparison conditions at post-treatment and at follow-up**

| | | *RC + CBT* | | | | *Comparison conditions* | | | |
|---|---|---|---|---|---|---|---|---|---|
| | | *Post* | | *Follow-up* | | *Post* | | *Follow-up* | |
| *Study* | *Measure* | *RCI* | *% CSI* | *RCI* | *% CSI* | *RCI* | *% CSI* | *RCI* | *% CSI* |
| 1. Barrowclough et al.[40] Haddock et al.[41] | PANSS | | | | | | | | |
| | Positive | 0.45 | — | 0.99 | — | 0.44[a] | — | 0.86 | — |
| | Negative | 0.37 | — | 1.42 | — | 1.23[a] | — | 0.89[a] | — |
| | General | M | — | 1.90 | — | M | — | 0.61 | — |
| 2. Bechdolf et al.[42] | PANSS | | | | | | | | |
| | Positive | 0.64 | — | 0.56 | — | 1.04 | — | 1.04 | — |
| | Negative | 0.20 | — | 0.76 | — | 0.90 | — | 0.92 | — |
| | General | 1.15 | — | 1.04 | — | 1.43 | — | 1.21 | — |
| 3. Durham et al.[43,b] | PSYRATS | | | | | | | | |
| | Delusions | 0.58 | — | **2.19*** | **6** | 0.29[a]/0.36 | — | 0.00/1.90 | — |
| | Hallucinations | 1.72 | — | 1.54 | — | 0.44/1.04 | — | 1.07/1.80 | — |
| 4. Gumley et al.[44] | PANSS | | | | | | | | |
| | Positive | 1.38 | — | — | — | 0.52 | — | — | — |
| | Negative | 0.91 | — | — | — | 0.40 | — | — | — |
| | General | 1.25 | — | — | — | 0.44 | — | — | — |
| 5. Kuipers et al.[27,45] | BPRS Total | 1.48 | — | 1.74 | — | 0.41 | — | 0.23 | — |
| 6. Lewis et al.[46.b.c] | PANSS | | | | | | | | |
| | Positive | **2.94*** | **71** | — | — | **2.71*/3.01*** | **65/76** | — | — |
| | PSYRATS | | | | | | | | |
| | Delusions | **9.65*** | **55** | — | — | **8.41*/10.09*** | **52/60** | — | — |
| | Hallucinations | **10.96*** | **92** | — | — | **9.76*/10.87*** | **75/84** | — | — |
| 7. Pinto et al.[47] | BPRS Total[d] | **3.45*** | **99** | — | — | **3.22*** | **60** | — | — |
| | SAPS | **14.87*** | **34** | — | — | **9.37*** | **5** | — | — |
| | SANS | **4.63*** | **18** | — | — | **1.96*** | **10** | — | — |
| 8. Rector et al.[29] | PANSS | | | | | | | | |
| | Positive | 0.70 | — | 1.01 | — | 0.34 | — | 1.01 | — |
| | Negative | 0.61 | — | 1.14 | — | 0.07 | — | 0.05[a] | — |
| | General | 0.98 | — | 1.23 | — | 0.48 | — | 0.89 | — |
| 9. Sensky et al.[30] | CPRS | | | | | | | | |
| | SCZ change | **2.50*** | **10** | **3.05*** | **14** | 1.86 | — | 1.64 | — |
| | Total | **2.62*** | **17** | **3.55*** | **27** | **2.29*** | **19** | 1.65 | — |
| | SANS | 1.53 | — | 1.94 | — | 1.13 | — | 0.65 | — |
| 10. Startup et al.[45] | BPRS Total | **4.06*** | **66** | — | — | **2.15*** | **34** | — | — |
| | SANS | 0.52 | — | — | — | 0.14 | — | — | — |
| 11. Trower et al.[48] | PANSS | | | | | | | | |
| | Positive | 0.94 | — | 1.22 | — | 0.35[a] | — | 0.31[a] | — |
| | Negative | 0.83 | — | 1.20 | — | 0.55[a] | — | 0.11[a] | — |
| | General | 1.27 | — | 1.03 | — | 0.68[a] | — | 1.00[a] | — |
| 12. Turkington et al.[49] | CPRS | | | | | | | | |
| | SCZ change | 0.26 | — | — | — | 0.15 | — | — | — |
| | Total | 0.63 | — | — | — | 0.29 | — | — | — |

**p < 0 .05 (RCI > 1.96) (bolded text denotes statistically significant results); RC = routine care; RCI = Reliable Change Index; % CSI = estimated percentage of subjects meeting criteria for clinically significant change (Jacobson & Truax[19] cutoff score a); BPRS = Brief Psychiatric Rating Scale; CPRS = Comprehensive Psychopathological Rating Scale; M = missing, data not reported; PANSS = Positive and Negative Syndrome Scale; PSYRATS = Psychotic Symptoms Ratings Scale; SAPS = Scale for the Assessment of Positive Symptoms; SANS = Scale for the Assessment of Negative Symptoms; SCZ = schizophrenia.*

*[a]Non-significant worsening in symptoms.*

*[b]Results calculated separately for RC only and RC + non-CBT conditions and denoted as: (RC alone)/(RC + non-CBT).*

*[c]Due to significant treatment by recruitment site interactions, follow-up data reported in Tarrier et al.[66] were not analyzed.*

*[d]BPRS-Expanded Version (24 items).*

### Comparison Conditions

***Reliable change.*** Table 3 also shows results of clinical significance analyses for the comparison groups in the selected studies. When post-treatment means of comparison conditions were considered, 33% (4/12) of studies showed reliable symptomatic improvement on at least one measure. When post-treatment or follow-up data were analyzed if available, reliable change rates decreased to 25% (3/12) because one study[30] showed reliable change at post-treatment but not at follow-up in the comparison condition. Of the 31 total measures assessed, 26% ($n$ = 8) showed reliable change at post-treatment and 23% ($n$ = 7) when follow-up data were included. The correlation between CBT and comparison condition measures meeting criteria for reliable change was significant ($r$ = 0.78, $p < 0.001$). In other words, if the CBT condition showed reliable change, the comparison condition was likely to as well.

Rates were also calculated based on the type of comparison condition used. For RC only conditions, 22% (2/9) of studies showed reliable change on at least one measure at post-treatment or follow-up. Although sample size was small for evaluating non-CBT conditions, rates were nearly identical to those of CBT conditions. For RC plus non-CBT, 60% (3/5) of studies showed reliable change at post-treatment and 40% (2/5) when follow-up data were included as available.

***Clinically significant change.*** In comparison conditions, clinically significant improvement was estimated to be achieved in an average of 14% of patients at post-treatment or follow-up across all measures. When only those measures showing reliable change were analyzed, the proportion of patients estimated to demonstrate clinically significant change was 52%.

Although sample size for subanalyses was small and findings should be interpreted with caution, results were analyzed based on the type of comparison condition. For RC only conditions, the average percentages of clinically significant improvement were 10% across all measures and 57% for those measures showing reliable change. Although some studies showed symptom worsening over time in RC only conditions, none represented a reliable change according to the RCI. For non-CBT conditions, the average percentages of clinically significant improvement were 20% across all measures and 49% for those measures showing reliable change. Taking into account the relatively small sample sizes, results of clinically significant improvement on measures showing reliable change appeared to be largely indistinguishable between CBT and non-CBT comparison conditions.

### Generalizability of Results

Due to the relatively small sample of studies, the generalizability of the above findings was examined by comparing results to those obtained using a larger sample of studies that included non-randomized trials and pilot studies ($n$ = 23; not depicted in Table 2, list available from the author). Results were similar with 48% of CBT and 26% of comparison conditions showing reliable change on at least one measure at post-treatment or follow-up. Furthermore, CBT conditions in 17% of studies versus 13% of comparison conditions showed absolute improvement of two SDs or greater. Due to the similarity in results and the overall heterogeneity found among studies in this area, only results from the more homogeneous subsample of studies are reported in detail. The 12 selected studies reported are more consistent with those included in traditional meta-analyses in this area.

## DISCUSSION

This study was the first to systematically examine the clinical significance of symptomatic improvement reported in trials of CBT for schizophrenia and related psychotic disorders using methods for group data developed by Sheldrick et al.[18] and Hageman and Arrindell.[33] The study was also unique in that rates of clinically significant symptomatic improvement were contrasted between comparison conditions. In summary, 42% of selected studies showed reliable change on at least one symptom measure in RC plus CBT conditions at post-treatment or at follow-up. This is in contrast to only 25% of comparison conditions (RC only or RC plus non-CBT) in the same studies. However, this discrepancy appeared to be attributable to the inclusion of RC only conditions. The clear advantage to adjunctive CBT in contrast to routine care alone is consistent with the findings of statistically significant differences from traditional meta-analyses in this area.[14] The equivocal results in the current study are similar to those found in the literature when CBT is compared to credible alternative interventions in methodologically stringent trials.[11] Currently, there are insufficient data to indicate that CBT is specifically efficacious for treating psychotic disorders, and supportive interventions appear to be quite effective in their own right. However, the data strongly suggest that, in general, psychosocial interventions

have the potential to be quite beneficial for patients with psychotic disorders when used adjunctively.

For those measures showing reliable change, the estimated proportions of patients demonstrating clinically significant symptomatic improvement (i.e., two SDs or greater change) in the CBT and comparison conditions were virtually identical (48% compared to 52%, respectively). The average percentages of patients demonstrating reliable change *and* clinically significant change between CBT and comparison conditions across all studies were only 16% and 14%, respectively. Jacobson and Truax[19] suggest that changes of two SDs or more in severely ill populations represent a large and clinically noticeable improvement. Results of the current study suggest that such dramatic levels of symptomatic improvement are only likely to occur in a minority of patients treated with adjunctive CBT.

Furthermore, differences were observed based on the type of symptoms assessed, as 46% of measures of positive symptoms in contrast to only 13% of measures of negative symptoms showed reliable change in CBT conditions. These results are consistent with suggestions by Tarrier et al.[67] that CBT may be particularly effective for positive symptoms. Alternatively, the results may suggest that more time between assessments is required to demonstrate changes in negative compared with positive symptoms. Rector et al.[29] reported statistically greater improvement in negative symptoms with CBT compared with an RC condition at 6-month follow-up but not at post-treatment. However, the magnitude of these gains did not meet the criteria used for clinically significant improvement in the current study. In general, the sample sizes of trials in the current study were too small to explore these types of subanalyses fully.

Only two of the selected trials[43,46] contained arms involving routine care alone, routine care plus CBT, and routine care plus a non-CBT intervention within the same study. In the Durham et al.[43] study, the CBT condition achieved reliable change at follow-up on one measure, but the RC alone and RC plus supportive therapy conditions did not (although the supportive therapy condition was approaching significance at follow-up). In the Lewis et al. study,[46] the proportion of patients who met criteria for clinically significant change was nearly identical in the CBT and supportive therapy conditions.

The current results are similar to those found in other studies of CBT that have assessed clinically significant change in difficult-to-treat populations. For example, in an analysis of CBT for bulimia nervosa, Lundgren et al.[24] found rates of reliable change ranging from 0% to 93% (with an average of 62%), depending on the measure used. Furthermore, results of a normative comparison analysis revealed that only two measures met this more stringent criterion for clinical significance. Similarly, Abramowitz[23] found that only 50% of studies of CBT for obsessive-compulsive disorder demonstrated reliable improvement, but that most of the patients meeting criteria for such improvement did not return to normal functioning. Sheldrick et al.[18] reported better results in their investigation of clinically significant change from therapies for children with conduct disorder. Reliable change was demonstrated in 96% of the studies evaluated. However, normative comparisons revealed that only 48% of studies reported end-treatment scores that were equivalent to those of non-clinical samples.

Although the effects found in the current study may appear relatively small, it is important to keep a few points in mind. First, CBT for psychosis is typically provided as an adjunctive treatment. Because individuals with severe mental illness are already likely to be receiving pharmacotherapy and other clinical services, modified standards may be required to assess the value of psychotherapy for these patients relative to other patient populations, such as individuals with mood or anxiety disorders.[68,69] In other words, the additive efficacy of CBT for psychosis would be expected to be less than the absolute effects shown in studies in which CBT is the only intervention provided (e.g., therapy for conduct disorder as examined by Sheldrick et al.[18]). Therefore, even more modest effects could still be considered clinically meaningful in the current context. In the current investigation, in particular, more studies showed reliable change with CBT compared with routine care alone. Even though few patients were estimated to show gains of two SDs or greater, the fact that some did is promising and suggests that it is possible for patients to improve considerably with an adjunctive treatment when evaluated even by these stringent standards. Due to the method used to calculate reliable change, most studies showing reliable change likely demonstrated mean improvement of at least one SD, which still signifies substantial gains.

The current study is not without limitations. The sample of studies was relatively small, particularly for any subanalyses; however, it is important to note that findings based on a larger sample of studies revealed similar results. In addition, analyses of clinical significance cannot determine whether the amount of improvement observed was specifically a result of the intervention or attributable to nonspecific factors, bias, or other confounding variables.[70] Furthermore, some

relevant studies were not included in the analyses. For example, a trial by Tarrier et al.[31] was not analyzed because these researchers reported results using a composite measure of symptomatic improvement that was not standardized.

One reason that more of the studies reported here did not show reliable change may be that mean scores on some study measures were low at baseline, thus leaving little room for large changes. This may be particularly applicable to ratings of negative symptoms or when treating samples of outpatients with residual psychotic symptoms. In other words, one would be more likely to find reliable change when treating certain groups, such as acutely ill patients. Furthermore, schizophrenia has different treatment phases, which range from stabilization in the acutely ill to amelioration and prevention in more chronically ill populations. The targets and expectations for improvement are thus highly variable over time in complicated mental illnesses such as schizophrenia. It should be noted, however, that treatment setting (inpatient versus outpatient) was not significantly associated with reliable change in the current study.

Furthermore, if instrument reliability is low, the magnitude of improvement required to demonstrate reliable change is quite high. Although the interviewer-rated measures used in the current study possessed adequate psychometric properties, reliability and validity estimates ranged considerably across studies. Reliability estimates used in calculations derived from psychometric studies were relatively lenient and were used unless study-specific reliability estimates suggested otherwise. In addition, results from reliable change analyses varied considerably based on the measure used or the type of symptom assessed. This is a known limitation of clinical significance analyses,[71] although some data suggest that results tend to generalize across different measurement domains within a study.[72] Current findings highlight the need to continue refining outcome measures for individuals with psychotic disorders, so that the change attributable to the intervention can be differentiated from measurement error. This issue is of considerable importance in the current context, as the magnitude of effects attributable to any adjunctive treatment will likely be fairly modest.

It is also important to emphasize that many relevant outcome variables for patients were not examined in the current study, which focused on psychotic symptoms. Some have argued that symptom measures assess only one limited dimension of clinically significant improvement and may be a poor indicator of a return to normal *functioning*.[65,73,74] Functional outcome measures, such as rehospitalization rate, work status, interpersonal functioning, social skills, and quality of life, may improve despite the continued presence of residual psychotic symptoms and may better address the issue of "normality." Interestingly, Bach and Hayes[75] and Gaudiano and Herbert[26] found that mindfulness/acceptance-based therapy for psychotic disorders decreased believability and distress associated with positive symptoms and rehospitalization rates, but not frequency of psychotic symptoms relative to routine care. Many would argue that distress and believability associated with psychotic symptoms are the primary targets of CBT for psychosis. In addition, it is believed that relieving stress and improving overall functioning can have a beneficial effect on the frequency and severity of symptoms.

As the current study demonstrates, the issue of reliable change remains an important standard for describing the effects of any psychiatric intervention. Less than half of the selected studies showed reliable change on *at least* one symptom measure. Results further suggest that the routine care that is typically delivered to patients with schizophrenia and psychotic disorders accounts for a substantial proportion of this effect. Other studies that have examined clinically significant improvement in group datasets have only examined absolute change in CBT conditions. In the current study, rates of clinically significant improvement were contrasted with those found in conditions involving only treatment as usual, which allowed for a more refined interpretation of effects.

Furthermore, Hageman and Arrindell[33] note that discrepancies could exist between results found at the individual versus group level of analysis using their $\text{Proportion}_{\text{CLINICALLY SIGNIFICANT}}$ formula. The study by Startup et al.[25] was used to examine this potential limitation because it was the only study selected that used Jacobson and Truax's[19] standard criteria to assess clinical significance. These researchers reported that 60% of the CBT group compared with 40% of the treatment as usual group demonstrated reliable change and clinically significant improvement on the Global Assessment of Functioning scale.[76] Unfortunately, they did not report these analyses for measures of psychotic symptoms. Results from the current study showed that 66% of patients receiving CBT and 34% of those receiving treatment as usual were estimated to meet criteria for reliable change and clinically significant improvement on the BPRS. Other studies have reported partial criteria for determining clinically significant treatment

effects. Kuipers et al.[27] reported that 29% of patients receiving CBT met criteria for reliable improvement on the BPRS, although these researchers did not compute results based on the accepted methods described in the current study. Results for the BPRS were not significant based on group level analysis, although the RCI approached significance at follow-up (RCI = 1.74).

In general, heterogeneity both within and between trials has been a major weakness of clinical trials of CBT for psychosis.[11,39] However, previous meta-analytic reviews[13–16,36] of clinical trials in this area have demonstrated that important information can still be obtained from such group analyses. Examination of group data does not diminish the importance of clinical significance analyses conducted at the individual level. In fact, a much richer and more complete picture of improvement is possible when individual patient scores are examined. For example, group findings may obscure the substantial improvement demonstrated by some patients.

The concept of clinically significant change, as originally proposed by Jacobson et al.[17] is best understood as one important tool for describing treatment outcome. The current study demonstrates that modifications in conceptualization and application are needed when attempting to apply these methods to more severely ill populations. Although some may interpret the results of the current study as a cause for cautious optimism, greater refinement of assessment and analytic strategies in future clinical trials will be required before the benefit of adjunctive psychotherapy for schizophrenia can be fully understood. Future studies should report results from clinical significance analyses using standardized procedures and should examine a broader range of outcome measures to investigate the potential of CBT for improving functioning in schizophrenia and other psychotic disorders.

## References


1. Craig T, Hwang MY. Comorbidity in schizophrenia: Epidemiology and clinical implications. Psychiatr Ann 2000;30:76–8.
2. Cassano GB, Pini S, Saettoni M, et al. Occurrence and clinical correlates of psychiatric comorbidity in patients with psychotic disorders. J Clin Psychiatry 1998;59:60–8.
3. Margolese HC, Chouinard G, Beauclair L, et al. Therapeutic tolerance and rebound psychosis during quetiapine maintenance monotherapy in patients with schizophrenia and schizoaffective disorder. J Clin Psychopharmacol 2002;22: 347–52.
4. Ritsner M, Ponizovsky A, Endicott J, et al. The impact of side-effects of antipsychotic agents on life satisfaction of schizophrenia patients: A naturalistic study. Eur Neuropsychopharmacol 2002;12:31–8.
5. Lacro JP, Dunn LB, Dolder CR, et al. Prevalence of and risk factors for medication nonadherence in patients with schizophrenia: A comprehensive review of recent literature. J Clin Psychiatry 2002;63:892–909.
6. Mueser KT, Berenbaum H. Psychodynamic treatment of schizophrenia: Is there a future? Psychol Med 1990;20:253–62.
7. McGlashan TH. The Chestnut Lodge follow-up study. II. Long-term outcome of schizophrenia and the affective disorders. Arch Gen Psychiatry 1984;41:586–601.
8. Tarrier N, Haddock G, Barrowclough C, et al. Are all psychological treatments for psychosis equal? The need for CBT in the treatment of psychosis and not for psychodynamic psychotherapy: Comment on Paley and Shapiro (2002). Psych Psychother: Ther Res Pract 2003;75:365–75.
9. Mueser KT, Corrigan PW, Hilton DW, et al. Illness management and recovery: A review of the research. Psychiatr Serv 2002;53:1272–84.
10. Kingdon DG, Turkington D. Cognitive-behavioral therapy of schizophrenia. New York: Guilford; 1994.
11. Gaudiano BA. Cognitive behavior therapies for psychotic disorders: Current empirical status and future directions. Clin Psych Sci Pract 2005;12:33–50.
12. Turkington D, Dudley R, Warman DM, et al. Cognitive-behavioral therapy for schizophrenia: A review. J Psychiatr Pract 2004;10:5–16.
13. Cormac I, Jones C, Campbell C, et al. Cognitive behaviour therapy for schizophrenia (Cochrane Review). The Cochrane Library, Issue 3. Oxford: Update Software; 2003.
14. Gould RA, Mueser KT, Bolton E, et al. Cognitive therapy for psychosis in schizophrenia: An effect size analysis. Schizophr Res 2001;48:335–42.
15. Pilling S, Bebbington P, Kuipers E, et al. Psychological treatments in schizophrenia: I. Meta-analysis of family intervention and cognitive behaviour therapy. Psychol Med 2002;32: 763–82.
16. Rector NA, Beck AT. Cognitive behavioral therapy for schizophrenia: An empirical review. J Nerv Ment Dis 2002;189: 278–87.
17. Jacobson NS, Follette WC, Revenstorf D. Psychotherapy outcome research: Methods for reporting variability and evaluating clinical significance. Behav Ther 1984;15:336–52.
18. Sheldrick RC, Kendall PC, Heimberg RG. The clinical significance of treatments: A comparison of three treatments for conduct disordered children. Clin Psych Sci Pract 2001;8:418–30.
19. Jacobson NS, Truax P. Clinical significance: A statistical approach to defining meaningful change in psychotherapy research. J Consult Clin Psychol 1991;59:12–9.
20. Follette WC, Callaghan GM. The importance of the principle of clinical significance—defining significant to whom and for what purpose: A response to Tingey, Lambert, Burlingame, and Hansen. Psychother Res 1996;6:133–43.
21. Hafkenscheid A. Psychometric measures of individual change: An empirical comparison with the Brief Psychiatric Rating Scale (BPRS). Acta Psychiatr Scand 2000;101:235–42.
22. Speer DC. What is the role of two-wave designs in clinical research? Comment on Hageman and Arrindell. Behav Res Ther 1999;37:1203–10; discussion 129–33.
23. Abramowitz JS. Does cognitive-behavioral therapy cure obsessive-compulsive disorder? A meta-analytic evaluation of clinical significance. Behav Ther 1998;29:339–55.
24. Lundgren JD, Danoff-Burg S, Anderson DA. Cognitive-behav-

ioral therapy for bulimia nervosa: An empirical analysis of clinical significance. Int J Eat Disord 2004;35:262–74.
25. Startup M, Jackson MC, Bendix S. North Wales randomized controlled trial of cognitive behaviour therapy for acute schizophrenia spectrum disorders: Outcomes at 6 and 12 months. Psychol Med 2004;34:413–22.
26. Gaudiano BA, Herbert JD. Acute treatment of inpatients with psychotic symptoms using Acceptance and Commitment Therapy: Pilot results. Behav Res Ther 2005, May 10 [Epub].
27. Kuipers E, Garety P, Fowler D, et al. London-East Anglia randomised controlled trial of cognitive-behavioural therapy for psychosis. I: Effects of the treatment phase. Br J Psychiatry 1997;171:319–27.
28. Morrison AP, Renton JC, Williams S, et al. Delivering cognitive therapy to people with psychosis in a community mental health setting: An effectiveness study. Acta Psychiatr Scand 2004;110:36–44.
29. Rector NA, Seeman MV, Segal ZV. Cognitive therapy for schizophrenia: A preliminary randomized controlled trial. Schizophr Res 2003;63:1–11.
30. Sensky T, Turkington D, Kingdon D, et al. A randomized controlled trial of cognitive-behavioral therapy for persistent symptoms in schizophrenia resistant to medication. Arch Gen Psychiatry 2000;57:165–72.
31. Tarrier N, Yusupoff L, Kinney C, et al. Randomised controlled trial of intensive cognitive behaviour therapy for patients with chronic schizophrenia. BMJ 1998;317:303–7.
32. Tarrier N, Wittkowski A, Kinney C, et al. Durability of the effects of cognitive-behavioural therapy in the treatment of chronic schizophrenia: 12-month follow-up. Br J Psychiatry 1999;174:500–4.
33. Hageman WJ, Arrindell WA. Establishing clinically significant change: Increment of precision and the distinction between individual and group level of analysis. Behav Res Ther 1999;37:1169–93.
34. Bouchard S, Vallières A, Roy M-A, et al. Cognitive restructuring in the treatment of psychotic symptoms in schizophrenia: A critical analysis. Behav Ther 1996;27:257–77.
35. Dickerson FB. Cognitive behavioral psychotherapy for schizophrenia: A review of recent empirical studies. Schizophr Res 2000;43:71–90.
36. Marshal M, Lockwood A. Early intervention for psychosis (Cochrane Review). The Cochrane Library, Issue 2. Chichester, UK: John Wiley & Sons; 2004.
37. Mueser KT, Bond GR, Drake RE. Community-based treatment of schizophrenia and other severe mental disorders: Treatment outcomes. Medscape General Medicine 2001;3: Available at http://www.medscape.com/viewarticle/430529.
38. Rector NA, Beck AT. Cognitive therapy for schizophrenia: From conceptualization to intervention. Can J Psychiatry 2002;47:39–48.
39. Tarrier N, Wykes T. Is there evidence that cognitive behaviour therapy is an effective treatment for schizophrenia? A cautious or cautionary tale? Behav Res Ther 2004;42:1377–401.
40. Barrowclough C, Haddock G, Tarrier N, et al. Randomized controlled trial of motivational interviewing, cognitive behavior therapy, and family intervention for patients with comorbid schizophrenia and substance use disorders. Am J Psychiatry 2001;158:1706–13.
41. Haddock G, Barrowclough C, Tarrier N, et al. Cognitive-behavioural therapy and motivational intervention for schizophrenia and substance misuse. 18-month outcomes of a randomised controlled trial. Br J Psychiatry 2003;183:418–26.
42. Bechdolf A, Knost B, Kuntermann C, et al. A randomized comparison of group cognitive-behavioural therapy and group psychoeducation in patients with schizophrenia. Acta Psychiatr Scand 2004;110:21–8.
43. Durham RC, Guthrie M, Morton RV, et al. Tayside-Fife clinical trial of cognitive-behavioural therapy for medication-resistant psychotic symptoms. Results to 3-month follow-up. Br J Psychiatry 2003;182:303–11.
44. Gumley A, O'Grady M, McNay L, et al. Early intervention for relapse in schizophrenia: Results of a 12-month randomized controlled trial of cognitive behavioural therapy. Psychol Med 2003;33:419–31.
45. Kuipers E, Fowler D, Garety P, et al. London-east Anglia randomised controlled trial of cognitive-behavioural therapy for psychosis. III: Follow-up and economic evaluation at 18 months. Br J Psychiatry 1998;173:61–8.
46. Lewis S, Tarrier N, Haddock G, et al. Randomised controlled trial of cognitive-behavioural therapy in early schizophrenia: Acute-phase outcomes. Br J Psychiatry Suppl 2002;43:S91–7.
47. Pinto A, La Pia S, Mennella R, et al. Cognitive-behavioral therapy and clozapine for clients with treatment-refractory schizophrenia. Psychiatr Serv 1999;50:901–4.
48. Trower P, Birchwood M, Meaden A, et al. Cognitive therapy for command hallucinations: Randomised controlled trial. Br J Psychiatry 2004;184:312–20.
49. Turkington D, Kingdon D, Turner T. Effectiveness of a brief cognitive-behavioural therapy intervention in the treatment of schizophrenia. Br J Psychiatry 2002;180:523–7.
50. Pratt SI, Mueser KT. Schizophrenia. In: Antony MM, Barlow DH, eds. Handbook of assessment and treatment planning for psychological disorders. New York: Guilford; 2002:375–414.
51. Lukoff D, Nuechterlein KH, Ventura J. Manual for expanded Brief Psychiatric Rating Scale (BPRS). Schizophr Bull 1986;12:594–602.
52. Overall G, Gorham D. The Brief Psychiatric Rating Scale. Psychol Rep 1962;10:799–812.
53. Asberg M, Montgomery SA, Perris C, et al. A comprehensive psychopathological rating scale. Acta Psychiatr Scand Suppl 1978:5–27.
54. Kay SR, Fiszbein A, Opler LA. The positive and negative syndrome scale (PANSS) for schizophrenia. Schizophr Bull 1987;13:261–76.
55. Haddock G, McCarron J, Tarrier N, et al. Scales to measure dimensions of hallucinations and delusions: The psychotic symptom rating scales (PSYRATS). Psychol Med 1999;29:879–89.
56. Andreasen NC, Olsen S. Negative v positive schizophrenia. Definition and validation. Arch Gen Psychiatry 1982;39:789–94.
57. Andreasen NC. Negative symptoms in schizophrenia. Definition and reliability. Arch Gen Psychiatry 1982;39:784–8.
58. Jacobson NS, Wilson L, Tupper C. The clinical significance of treatment gains resulting from exposure-based intervention for agoraphobia: A reanalysis of outcome data. Behav Ther 1988;19:539–54.
59. Jacobson NS, Follette WC, Revenstorf D, et al. Variability in outcome and clinical significance of behavioral marital therapy: A reanalysis of outcome data. Prev Treat 2000:Article 15, Available at: http://journals.apa.org/prevention/volume3

/pre0030015a.html.
60. Long JD, Brekke JS. Longitudinal factor structure of the Brief Psychiatric Rating Scale in schizophrenia. Psychol Assess 1999;11:498–506.
61. Ventura J, Green MF, Shaner A, et al. Training and quality assurance with the Brief Psychiatric Rating Scale: "The drift busters". Int J Meth Psychiatr Res 1993;3:221–44.
62. Kay SR, Opler LA, Lindenmayer JP. Reliability and validity of the positive and negative syndrome scale for schizophrenics. Psychiatry Res 1988;23:99–110.
63. Schuldberg D, Quinlan DM, Morgenstern H, et al. Positive and negative symptoms in chronic psychiatric outpatients: Reliability, stability, and factor structure. Psychol Assess 1990;2:262–68.
64. Kocsis JH. New issues in the prediction of antidepressant response. Psychopharmacol Bull 1990;26:49–53.
65. Kazdin AE. The meanings and measurement of clinical significance. J Consult Clin Psychol 1999;67:332–9.
66. Tarrier N, Lewis S, Haddock G, et al. Cognitive-behavioural therapy in first-episode and early schizophrenia. 18-month follow-up of a randomised controlled trial. Br J Psychiatry 2004;184:231–9.
67. Tarrier N, Kinney C, McCarthy E, et al. Are some types of psychotic symptoms more responsive to cognitive-behaviour therapy? Behavioural Cog Psychther 2001;29:45–55.
68. Mueser KT. Methodological issues in evaluating empirically supported treatments for severe mental disorders. Paper presented at the 32nd Annual Convention of the Association for the Advancement of Behavior Therapy. Washington, D.C.; 1998.
69. Mueser KT, Torrey WC, Lynde D, et al. Implementing evidence-based practices for people with severe mental illness. Behav Modif 2003;27:387–411.
70. Hsu LM. A comparison of three methods of identifying reliable and clinically significant client changes: Commentary on Hageman and Arrindell. Behav Res Ther 1999;37:1195–202; discussion 1219–33.
71. Chambless DL. Measurement effects on outcome of treatment for agoraphobia. Behav Ther 1989;20:465–66.
72. Lunnen KM, Ogles BM. A multiperspective, multivariable evaluation of reliable change. J Consult Clin Psychol 1998;66:400–10.
73. Chorpita BF. Reflections on clinical significance: What do our best treatments accomplish and how can we best find out? Clin Psych Sci Pract 2001;8:451–54.
74. Gladis MM, Gosch EA, Dishuk NM, et al. Quality of life: Expanding the scope of clinical significance. J Consult Clin Psychol 1999;67:320–31.
75. Bach P, Hayes SC. The use of Acceptance and Commitment Therapy to prevent the rehospitalization of psychotic patients: A randomized controlled trial. J Consult Clin Psych 2002;70:1129–39.
76. American Psychiatric Association. Global Assessment of Functioning scale. In: Diagnostic and statistical manual of mental disorders, 4th edition, text revision. Washington, DC: American Psychiatric Association; 2000:32–4.

## Appendix: Formulas for Clinical Significance Analyses

1. $RCI = (X_2 - X_1) / S_{diff}$, $S_{diff} = \sqrt{2 \times SE^2}$, $SE = SD_{pre} \times \sqrt{1 - r}$, where $X_2$ = post-treatment mean, $X_1$ = pre-treatment mean, $S_{diff}$ = standard error of the difference, $SD_{pre}$ = standard deviation of the Time 1 assessment (in the sample used to determine reliability, if available), and $r$ = reliability coefficient of the measure.[18,19]
2. Inter-rater reliability coefficients used in RCI analyses were BPRS = 0.91;[60] BPRS-expanded version = 0.81;[61] CPRS = 0.92;[30] PANSS general scale = 0.87, positive scale = 0.83, negative scale = 0.85;[62] PANSS general scale = 0.87, positive scale = 0.87, negative scale = 0.73;[42] PANSS general scale = 0.72, positive scale = 0.91, negative scale = 0.87;[44] PSYRATS hallucination scale = 0.95 (item average), delusion scale = 0.97 (item average);[55] SANS = 0.84, SAPS = 0.91;[63] SANS = 0.83; SAPS = 0.88.[40] Note that reliability coefficients reported in outcome studies were used in analyses if available; otherwise, reliability coefficients from other psychometric studies of the corresponding measure were used instead.
3. $\text{Proportion}_{\text{CLINICALLY SIGNIFICANT}} = \text{F}(z_{\text{BEYOND CUTOFF}})$, $z_{\text{BEYOND CUTOFF}} = (TRC - X_2) / (SD_2 \times \sqrt{r_{xx(2)}})$, $r_{xx(2)} = (SD_2^2 - SE^2) / SD_2^2$, where $TCR$ = true cutoff for clinically significant change (cutoff score $a$[10]), $X_2$ = post-treatment mean, $SD_2$ = post-treatment standard deviation, $SE$ = standard error, $r_{xx(2)}$ = reliability of post-treatment scores.[33]