Minimum Clinically Important Difference in the Positive and Negative Syndrome Scale With Data From the Clinical Antipsychotic Trials of Intervention Effectiveness (CATIE)

Eric D. A. Hermes, MD; Daniel Sokoloff, BA; T. Scott Stroup, MD; and Robert A. Rosenheck, MD

ABSTRACT

Context: Establishing the minimum clinically important difference in the Positive and Negative Syndrome Scale (PANSS) is important to the interpretation of the research and clinical work conducted with this scale.

Method: This study employed both anchor-based and distributive methods to estimate the minimum clinically important difference for the PANSS by using data from the Clinical Antipsychotic Trials of Intervention Effectiveness (CATIE) schizophrenia trial, a large, multicenter trial for patients with schizophrenia. By using an equipercentile method, data from 1,442 individuals linked PANSS scores with both clinician and patient ratings on the Clinical Global Impressions scale (CGI). Data were also used to investigate the magnitude of the standard error of measurement (SEM), offering another estimate of the minimum clinically important difference.

Results: Cross-sectional, clinician-rated CGI-Severity of illness scores of 1 through 7 linked to PANSS scores of 32.4, 42.2, 57.5, 74.5, 93.0, 110.9, and 131.0, respectively. The minimum clinically important difference for PANSS scores using this scale equaled a 15.3-point (34.0%) change from baseline. A 1.96 SEM on the PANSS corresponded to a 16.5-point (36.2%) change from baseline. The minimum clinically important difference for a subsample with above-median baseline PANSS scores was 38% higher than a sample with lower baseline scores. With the patient-rated CGI as the anchor, PANSS scores were higher for CGI scores of 1 through 4, and the minimum clinically important difference was lower, 11.2 points (24.6%).

Conclusion: Minimum clinically important difference estimates from a longer-term effectiveness trial were consistent with previous efforts from shorter-term efficacy trials. Minimum clinically important difference estimates can help clinicians and researchers design future studies and interpret treatment change in future research and clinical work.

J Clin Psychiatry 2012;73(4):526–532

Submitted: *May 20, 2011; accepted September 21, 2011 (doi:10.4088/JCP.11m07162).*

Corresponding author: *Eric D. A. Hermes, MD, Department of Psychiatry, Yale School of Medicine, 300 George St, Ste 901, New Haven, CT 06511 (eric.hermes@yale.edu).*

The Positive and Negative Syndrome Scale (PANSS) is the most widely used standardized instrument for assessing symptom severity in schizophrenia.[1] It has been used as an outcome measure in a multitude of treatment efficacy studies and is increasingly used in clinical practice.[2–4] One drawback of the PANSS and other instruments based on summary rating scores is the lack of a gold standard with which to interpret results.[5] Clinicians must rely on experience with individual patients and populations to interpret PANSS scores and the clinical significance of various degrees of change.

The concept of the minimum clinically important difference has emerged as a way of giving clinical relevance to changes in standardized instrument scores when there is no gold standard of meaningful change. The minimum clinically important difference has been defined by Jaescheke and colleagues as "the smallest difference in a score in the domain of interest which patients [or providers] perceive as beneficial and which would mandate, in the absence of troublesome side effects and excessive cost, a meaningful change in the patient's management."[6(p408)] This concept is important both in clinical practice and clinical trials, especially noninferiority trials where a treatment must demonstrate the absence of a clinically important difference in minimum clinically important difference from a comparison treatment. The minimum clinically important difference is additionally important for determining if small statistically significant differences in measurement scores in studies with large sample sizes are great enough to be considered clinically meaningful.[7]

A number of techniques have emerged to estimate the minimum clinically important difference, which fall into 2 categories termed *anchor-based* and *distribution-based methods*.[8] Both approaches are used to estimate the change in a standardized instrument score associated with clinically important change. Anchor-based methods use a measure with established or face-value clinical meaning, such as the Clinical Global Impressions-Severity of illness scale (CGI-S),[9] to anchor scores on the measure of interest.[5] Distribution-based methods generally use the statistical characteristics of the sample, such as the standard deviation, to separate "signal" from "noise."[5] These methods describe observed differences or change in scale scores, but do not provide information about what size of change is clinically important and should ideally be linked to a clinical measure of the minimum clinically important difference.[10]

Recent studies using anchor-based methods linking PANSS change to the Clinical Global Impressions-Improvement scale (CGI-I) have estimated that between a 16% and 24% change in PANSS score corresponds to the minimum clinically important difference for minimal clinical improvement.[11–15] These studies have evaluated diverse populations. However, no study has yet used a distributive method to estimate the minimum clinically important difference of the PANSS or used a patient-rated measure of illness severity to anchor PANSS scores.

Clinical Points

- The minimum clinically important difference is the smallest difference in a measure that patients or providers perceive as beneficial.
- The minimum clinically important difference for the Positive and Negative Syndrome Scale (PANSS), the most commonly used measure of schizophrenic symptomatology, is approximately 15 points (34%) of baseline values.
- The minimum clinically important difference in PANSS scores varies according to the baseline PANSS score.

The Clinical Antipsychotic Trials of Intervention Effectiveness (CATIE) is one of the largest and longest schizophrenia trials conducted to date comparing the effectiveness of multiple antipsychotic treatments using broad inclusion criteria in a variety of treatment settings. The current effort seeks to use this sample to estimate the minimum clinically important difference of the PANSS using both anchor-based and distributive techniques in addition to using both patient- and rater-reported CGI-S scales.

METHOD

Sample

Researchers designed CATIE to compare the effectiveness and cost-effectiveness of available second-generation antipsychotics in a large, National Institute of Mental Health–funded, randomized, double-blind trial at 57 US sites, including both academic and community providers. Participants were 18 to 65 years of age with a diagnosis of schizophrenia. Those diagnosed with schizoaffective or cognitive disorder and those who had only 1 schizophrenic episode were excluded. Details of the study design have been presented elsewhere.[16,17] The study sample (n = 1,442) included individuals with data from baseline and follow-up time points.

Measures

The PANSS yields a total average symptom score based on 30 items rated from 1 to 7 (range, 30–210). Higher scores indicate more severe symptoms.[1] This study utilized PANSS results from baseline and 1-, 3-, and 6-month follow-up assessments. Raters were required to undergo both initial and ongoing certification in the PANSS to ensure high interrater reliability.[18]

The CGI-S is a widely used measure of global illness severity scored on a 7-item scale. The clinician is asked to rate the individual based on his or her "total clinical experience with the given population." The following scores can be given: 1 = normal, not at all ill; 2 = borderline mentally ill; 3 = mildly ill; 4 = moderately ill; 5 = markedly ill; 6 = severely ill; and 7 = among the most extremely ill patients.[9] Also included in CATIE was a measure of global illness severity similar to the CGI but scored by the patient on a 7-item scale, the patient-rated CGI (CGI-P). The patient was asked, "On a scale of '1' to '7,' where '1' is not at all ill, and '7' is the worst that your illness has ever been, how would you rate the severity of your schizophrenia symptoms?"

Analysis

This effort used both an anchor-based and distributive technique to estimate the minimum clinically important difference of the PANSS.

Anchor-based minimum clinically important difference estimation. Equipercentile-linking techniques were used to compare scores on the CGI-S/CGI-P and PANSS scales following a method used by Leucht et al[14] and described by Kolen and Brennen.[19] This technique functionally maps scores between the 2 different, but correlated, scales by linking scores on both measures at the same percentile rank.[20] For the purposes of this study, equipercentile linking is preferable because it links the CGI-S and CGI-P scales to the PANSS scale and back in an equivalent manner. A regression would produce different comparisons depending on which scale was treated as the independent variable. Linkings were computed using a process developed by Albano[21] in the R environment for statistical computing.[22]

In a cross-sectional analysis, CGI-S and CGI-P scores were initially mapped to PANSS scores using equipercentile-linking techniques for values at baseline, 1, 3, and 6 months, as well as from all data pooled across all time points. The reductions in CGI-S and CGI-P scores were then linked to corresponding changes in PANSS scores and the percent change from baseline in PANSS scores at 1, 3, and 6 months and across the pooled data. The percent change from baseline in PANSS scores was calculated by first subtracting the 30 baseline points, which correspond to the lowest score of 1 point on each of the 30 PANSS questions, thus establishing a valid 0 score at the bottom of the scale. In a secondary analysis, the population was stratified by the median baseline PANSS score, and linkings were repeated for those with "high" and "low" baseline PANSS scores.

Distribution-based minimum clinically important difference estimation. The distribution-based method estimates the minimum clinically important difference by comparing the observed change in PANSS to the variability in the PANSS, calculated in this study as the standard error of measurement (SEM). The SEM is a measure of the variability in PANSS scores reflecting the test-retest reliability of the scale and is considered to be a characteristic of the measure and not of the sample.[10] The formula for the SEM is $SEM = \delta\sqrt{1 - r}$, where δ is the standard deviation (SD) and r is the reliability as measured by the intraclass correlation coefficient. Previous efforts have indicated that values between 1 and 2.3 SEM approximate the minimum clinically important difference.[5,10,23] In order to calculate the SD and reliability of the PANSS in CATIE, a subset of the population that had stable symptomatology over the first month was chosen by identifying individuals whose CGI-S score did not change from baseline to 1 month, a method similar to that

Table 1. Mean Cross-Sectional and Change Scores of Psychometric Outcomes for the CATIE Schizophrenia Trial

	Baseline			Month 1			Month 3			Month 6		
Measure	n	Mean	SD	n	Mean	SD	n	Mean	SD	n	Mean	SD
Cross-sectional												
PANSS	1,442	75.6	17.6	1,306	71.2	18.1	1,103	69.1	17.8	978	68.1	17.5
CGI-S	1,441	4.0	0.9	1,300	3.8	1.0	1,105	3.7	1.0	971	3.6	1.0
CGI-P	1,439	3.5	1.6	1,301	3.3	1.6	1,110	3.3	1.5	972	3.2	1.4
Change from baseline												
PANSS				1,300	−4.4	12.2	1,096	−6.5	14.4	971	−7.5	16.1
CGI-S				1,291	−0.2	0.8	1,096	−0.3	1.0	963	−0.4	1.0
CGI-P				1,293	−0.2	1.6	1,103	−0.2	1.7	965	−0.3	1.7

Abbreviations: CATIE = Clinical Antipsychotic Trials of Intervention Effectiveness, CGI-P = patient-rated Clinical Global Impressions scale, CGI-S = Clinical Global Impression-Severity of illness scale, PANSS = Positive and Negative Syndrome Scale.

used by Duru and Fantino.[24] The SD of the PANSS scores for this population at baseline was used for the SEM calculation. The intraclass correlation coefficient was calculated by using a 2-way mixed model of PANSS scores at baseline and 1 month.

RESULTS

Population Characteristics

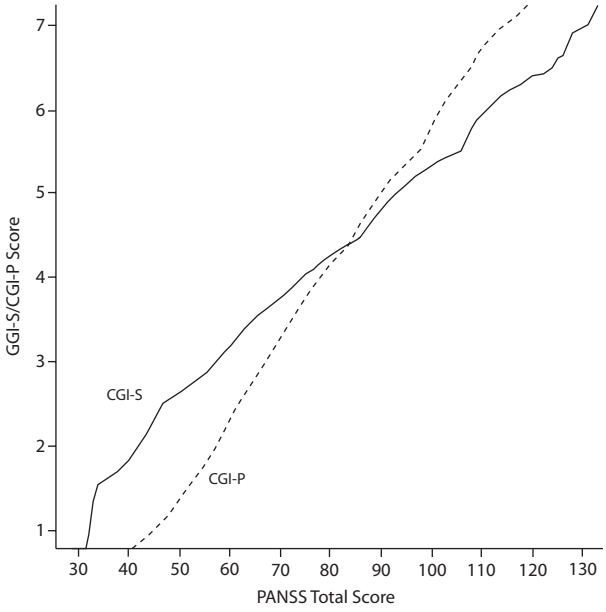
The number of individuals with PANSS, CGI-S, and CGI-P data are reported in Table 1 with mean (SD) scores. The number of individuals declined by 32% between baseline and 6 months due to study dropouts. Scores in all 3 measures decreased from baseline to 6 months by a total of 16.5% on the PANSS, 10.0% on the CGI-S, and 8.6% on the CGI-P.

Anchor-Based Results: Linking CGI and PANSS Total Scores

Linking of cross-sectional scores. Figure 1 depicts the cross-sectional analysis linking CGI-S scores of 1 through 7 to PANSS values of 32.4, 42.2, 57.5, 74.5, 93.0, 110.9, and 131.0, respectively, pooled across the 4 time periods (baseline and months 1, 3, and 6). CGI-P scores of 1 through 7 linked to PANSS values of 45.2, 57.4, 67.4, 78.0, 90.0, 101.8, and 114.4, pooled across the same 4 time periods and reflecting patient judgments of their improvement (Figure 1). The PANSS total scores that linked to CGI-P scores were generally greater than those that linked to CGI-S scores for less severe ratings (1–4). The PANSS scores that linked to CGI-P levels decreased over time for lower CGI-P values (Figure 2), but not when linked to CGI-S levels (data not shown).

Linking of change scores. Figure 3 depicts the change from baseline in CGI-S scores linked to the absolute change from baseline in PANSS scores. For this analysis, we assumed the minimum clinically important difference of the PANSS to be the change linked to a 1-point change in the CGI-S score. When using data pooled from all time points, a 1-point improvement in the CGI-S linked to a 15.3 point decrease (34.0%) in PANSS total score from baseline. The minimum clinically important difference for improvement in the PANSS was slightly lower at 1 month, 14.7 points (32.1%), compared to that calculated at 3 and 6 months, 15.0 points (35.3%) and 16.4 points (35.8), respectively. When the change in PANSS scores was linked to the CGI-P (Figures 2 and 3), the minimum clinically important difference for improvement was somewhat smaller, 11.2 points (24.6%) using pooled data, and showed a similar increase over the course of the study.

Figure 1. Equipercentile Linking of Pooled Cross-Sectional CGI-S and CGI-P With PANSS Total Scores Using All Data[a]

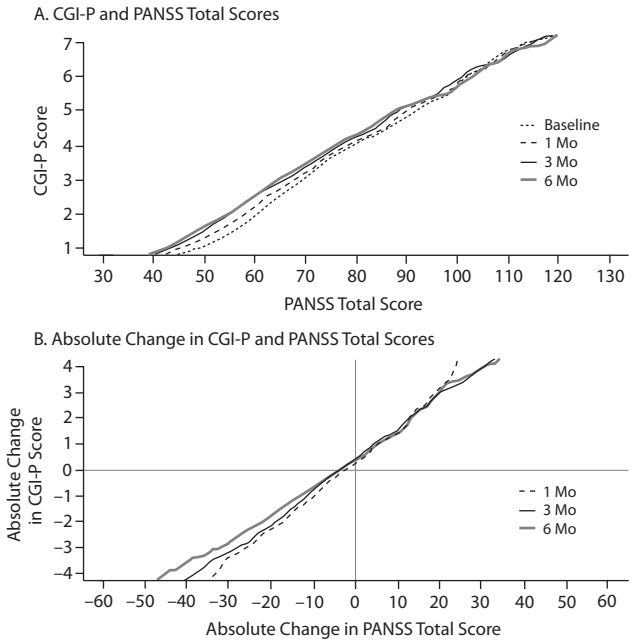


[a]The graph plots the corresponding CGI score for every PANSS score.
Abbreviations: CGI = Clinical Global Impressions, CGI-P = patient-rated Clinical Global Impressions scale, CGI-S = Clinical Global Impressions-Severity of illness scale, PANSS = Positive and Negative Syndrome Scale.

Stratification by baseline PANSS scores. Figure 4 depicts linking CGI-S and CGI-P scores with change in PANSS scores stratified by the median baseline PANSS. In those with lower baseline PANSS scores, CGI-S and CGI-P scores linked to lower PANSS scores compared to those with higher baseline PANSS scores. On the whole, this discrepancy in stratified scores was higher when using the CGI-P than when using the CGI-S. Similarly, the minimum clinically important

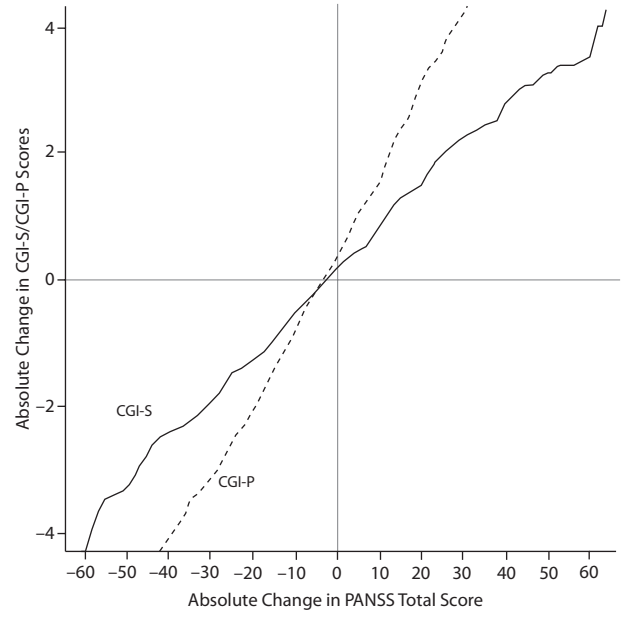
© Copyright 2012 Physicians Postgraduate Press, Inc.

Figure 2. Equipercentile Linking of Cross-Sectional CGI-P and Change in CGI-P Scores With PANSS Total Scores and Change in PANSS Total Scores by Time Period[a]



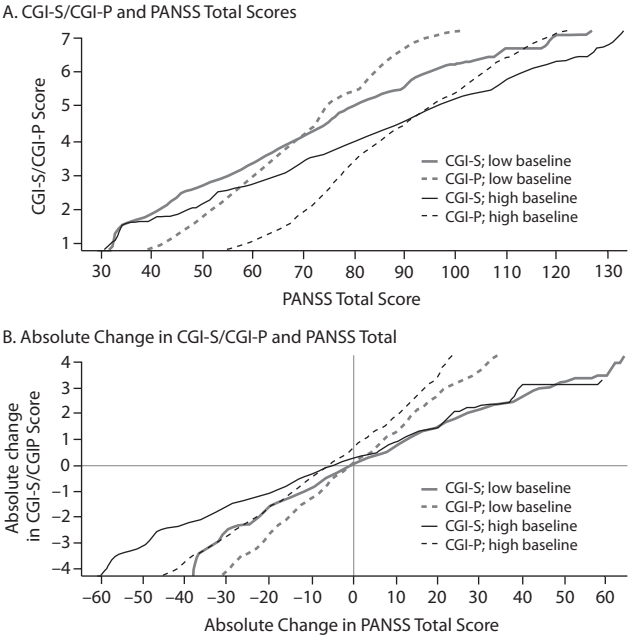
[a]The graph plots the corresponding absolute change in PANSS total score from baseline for every change in CGI score from baseline.
Abbreviations: CGI = Clinical Global Impressions scale, CGI-P = patient-rated Clinical Global Impressions scale, PANSS = Positive and Negative Syndrome Scale.

Figure 3. Equipercentile Linking of Change in Pooled CGI-S and CGI-P Scores With Change in PANSS Total Scores Using Pooled Data From All Time Points[a]



[a]The graph plots the corresponding absolute change in PANSS total score from baseline for every change in CGI score from baseline.
Abbreviations: CGI = Clinical Global Impressions scale, CGI-P = patient-rated Clinical Global Impressions scale, CGI-S = Clinical Global Impressions-Severity of illness scale, PANSS = Positive and Negative Syndrome Scale.

Figure 4. Equipercentile Linking of Cross-Sectional and Change in CGI-S/CGI-P With PANSS Total Scores and Change in PANSS Total Scores Stratified by Median Baseline PANSS Score[a]



[a]The graph plots the corresponding absolute change in PANSS total score from baseline for every change in CGI score from baseline.
Abbreviations: CGI = Clinical Global Impressions scale, CGI-P = patient-rated Clinical Global Impressions scale, CGI-S = Clinical Global Impressions-Severity of illness scale, PANSS = Positive and Negative Syndrome Scale.

difference when using the CGI-S was lower, 11.6 points, in those with lower baseline PANSS scores compared to those with higher baseline PANSS scores, 18.7. A similar difference was observed when using the CGI-P as the anchor, but both values were lower than those when using the CGI-S.

Distribution-Based Results

The subpopulation in which there was no change in the CGI-S between baseline and 1 month consisted of 707 individuals with a mean PANSS score of 76.0 (SD = 17.4). The reliability of the PANSS calculated as the intraclass correlation between PANSS scores at baseline and those at 1 month was 0.77, which computes to 1 SEM = 8.4 (18.4% decrease from baseline) and 1.96 SEM = 16.5 PANSS points (36.2% decrease from baseline).

DISCUSSION

This study estimated the minimum clinically important difference for the PANSS, a widely used measure of symptomatology in schizophrenia. The study used a large sample of participants in CATIE, which employed open-inclusion criteria, allowing for what is likely a generalizable population with chronic schizophrenia. In addition, varying analytic methods were used, including a novel anchor-based approach as well as by a distributive technique. We estimated

Table 2. Comparison of Minimum Clinically Important Difference Estimations for Improvement on the PANSS by Linking PANSS and Clinical Global Impressions Scale Scores Among Published Studies on the Topic

Study	Population	Linking Method	Evaluation of Change	Baseline PANSS Score, Mean (SD)	Minimum Clinically Important Difference[a]	
					Absolute Value	Percent of Baseline[b]
Current study	Single study, multiple sites, enrolled schizophrenics, excluded other psychotic disorders and first episode, inpatient/outpatient; n = 1,442	Equipercentile	Change in CGI-S	75.6 (17.6)	15.3	34
Schennach-Wolff et al[12]	Multicenter follow-up program, enrolled all psychotic disorders, inpatient only; n = 398	Equipercentile	CGI-I	…	5	17
Levine et al[11]	Pooled sample of 4 drug trials, first episode, acute, chronic, schizophrenia, and schizoaffective disorder; n = 2,698	Equipercentile	CGI-I	…	…	22
Leucht et al[14,32]	Pooled sample of 7 drug trials, enrolled all psychotic disorders, included only hospitalized "florid" psychosis; n = 4,091	Equipercentile	CGI-I, change in CGI-S	94 (19)	15	24
Cramer et al[15]	Single sample at multiple VA sites, schizophrenia resistant to at least 2 medications; n = 423	ANOVA	5-Level change scale[c]	91 (15)	…	21

[a]Minimum clinically important difference of the PANSS as calculated by anchor-based methods that link change in PANSS to a measure of global clinical change. For studies using the CGI-I, minimum clinically important difference estimated as the change in PANSS associated with CGI-I score = 3 (minimally improved). Reported estimates are averaged over the time periods analyzed.
[b]The minimum clinically important difference as a percent change from baseline. Cramer et al[15] is the only article specifically accounting for the minimum PANSS total score of 30 points when calculating percent change.
[c]Cramer et al[15] used a 5-level clinician-rated global clinical change scale.
Abbreviations: ANOVA = analysis of variance, CGI-I = Clinical Global Impressions-Improvement scale, CGI-S = Clinical Global Impressions-Severity of illness scale, PANSS = Positive and Negative Syndrome Scale, VA = Veterans Affairs.

a minimum clinically important difference for the PANSS of 15.3 points (34.0%) from baseline by linking PANSS and CGI-S scores. This estimate corresponds to the more conservative estimate of 1.96 SEM, which yielded a similar minimum clinically important difference estimate of 16.5 points. When patient-rated CGI-P values were used as an anchor, lower self-rated scores linked to higher PANSS scores compared to clinician-rated global severity. As a result, the minimum clinically important difference estimated via linking with patient-reported severity was slightly lower than that estimated by clinician report. In addition, the minimum clinically important difference estimate varied according to baseline psychopathology, with a 38% difference in the minimum clinically important difference between those with higher than lower baseline PANSS scores. These results suggest that, for patients in CATIE, in which mean PANSS scores changed, on average, no more than 14.8% during the follow-up period, symptom change may not have reached a clinically meaningful level on average when using a clinician-anchored or distributive estimation of the minimum clinically important difference, although many patients achieved meaningful gains.

This study used both anchor-based and distributive methods, each with its own conceptual underpinnings, to develop similar estimates of the minimum clinically important difference. This approach provides additional evidence of the validity of previous minimum clinically important difference estimates. Studies evaluating different distributive methods have suggested that the SEM is more concordant with clinically meaningful change than other distributive methods.[25,26] The use of the SEM also mitigates the problems associated with the absence of adequate psychometric investigations into the CGI.[14,27] However, there has been some disagreement as to how many SEMs an individual must change in order for that change to be considered clinically important.[23] Several groups have suggested that 1 SEM corresponds to the minimum clinically important difference on measures of health-related quality of life,[10,28,29] while other studies favor a value of 1.96 SEM.[23,30] It must be remembered that the SEM reflects change that cannot be attributed to measurement error alone,[31] and its application to an estimation of the minimum clinically important difference is theoretical and should be corroborated with clinical evaluation. The current results indicate that using the conservative value of 1.96 SEM may best approximate the minimum clinically important difference. This value corresponds to the value representing a 95% confidence interval proposed by McHorney and Tarlov.[25] In addition, similar studies using the CGI as an anchor to estimate the minimum clinically important difference have used the CGI-I scale,[11,12,14,32] in which clinicians are asked to rate the change in symptoms compared to baseline. Our use of serial CGI-S measurements avoids the recall bias potentially confounding the use of this scale.

Several prior studies using multiple populations have attempted to anchor PANSS scores with a measure of global clinical change. A summary of these studies is reported in Table 2. The studies vary in sample size, in the severity and chronicity of symptoms, and in methods for evaluating clinical change. A majority of these studies report the minimum clinically important difference as the percentage of reduction from baseline in the PANSS,[11,12,14,15,32] while a fraction[12,32] also report the absolute value of the change. Understanding change as a fraction of the whole may be easier than considering the absolute value of a change on an arbitrary scale, which may give validity to these authors' choice to report the minimum clinically important difference as a percentage. However, since the PANSS is a set of 30 questions scored from 1 to 7, the PANSS total score has a minimum value of

© Copyright 2012 Physicians Postgraduate Press, Inc.

30 points, which makes the calculation of percent change in the PANSS score problematic. In calculating the percent change from baseline, 30 points must be subtracted from the baseline value before the percent change can be calculated. In a review of the methods used in the prior efforts at calculating a minimum clinically important difference for the PANSS,[11,12,14,15,32] none specifically mention this issue or that they accounted for the 30 minimum points. Therefore, we feel that comparing the absolute values from these studies is proper.

The current results compare most closely with those reported by Leucht et al,[14,32] who described an absolute value for the minimum clinically important difference of 15 points in a pooled sample of 7 pharmacologic trials in a sample with "florid psychosis," where the mean (SD) baseline PANSS score was 94 (19). A more recent study by Schennach-Wolff et al[12] found a much lower estimate of approximately 5 points. These authors suggest the more generalizable population generated by the open inclusion criteria and probably consisting of individuals with more chronic disease may account for the difference. However, they do not report baseline PANSS values and were unable to quantify chronicity in the population, so comparison with the CATIE sample is difficult. Open inclusion criteria were used in CATIE to develop a representative sample that included both inpatients and outpatients at multiple sites and in multiple treatment settings in addition to providing rigorous training and testing of raters to ensure reliability of PANSS ratings. In doing so, this study may provide a more representative estimate of the minimum clinically important difference for the PANSS in patients with chronic schizophrenia.

Several authors have suggested that the minimum clinically important difference for the PANSS varies depending on the severity and chronicity of symptoms, and 2 efforts along with the current analysis have stratified samples on baseline PANSS scores to evaluate this hypothesis.[12,32] In the current analysis, we found a 38% difference between the minimum clinically important difference of the samples with more and less psychopathology at baseline, while this difference was 43% in Leucht et al[32] and 85% in Schennach-Wolff et al.[12] These results demonstrate that baseline psychopathology greatly affects the minimum clinically important difference of the PANSS and that both clinicians and researchers must account for this in their use of the minimum clinically important difference.

To our knowledge, this is the first study to evaluate the minimum clinically important difference from the patient's perspective. The CGI-S is anchored to the clinician's "total clinical experience with the given population,"[9] while the CGI-P is anchored to patients' knowledge of their past illness severity. While this distinction somewhat lessens the value of comparing the minimum clinically important differences derived from the 2 scales, we feel that the data show, in general terms, that patients can tolerate more symptomatology per unit of subjectively assessed severity and that there is less change in symptomatology associated with each subjectively detectable unit change in severity. These results indicate that the current experience of symptoms are less clinically meaningful to patients compared to formally trained raters but that patients experience the change in symptoms as more clinically meaningful. The former may provide evidence for a decrease in insight in those with schizophrenia or minimization of symptoms, as has been previously reported.[33] In addition, Cramer et al[15] pointed out that patients with schizophrenia may have a difficult time perceiving change due to changes in cognition or the presence of positive symptoms. Our results provide contradictory evidence in that patients identified less symptomatic change as clinically significant compared to trained raters.

However, linking PANSS with CGI-P scores displayed more change over time (Figure 3) compared to linking with CGI-S scores (data not shown). Patient-rated CGI levels linked to less severe PANSS scores over time, and the minimum clinically important difference based on CGI-P linking increased over time, trends that have also been seen in work by other authors.[12,14,15,32] These findings may indicate that the reliability of CGI-P ratings is less than that of CGI-S ratings; however, the extent to which this phenomena is driven by regression to the mean is unknown, especially given that linked CGI-S scores in this study did not show much change over time.

Several limitations of this study must be addressed. There was a high rate of treatment discontinuation in CATIE because of its relatively long duration, potentially introducing some attrition bias.[17] In the order of measures for CATIE, the PANSS was administered prior to the CGI-S, and both measures were most often completed by the same rater at each assessment. This procedure may bias CGI-S ratings in an unknown manner if the rater is aware of the PANSS score. Future investigations should replicate findings using independent raters for the PANSS and CGI, preferably with the clinician managing treatment completing the CGI. In addition, this study used a self-assessed version of the CGI-S, which has not been validated, especially in the light of possible problems with patient-report data from individuals with schizophrenia as discussed above. However, the availability of this measure did allow investigation of differences in patient and independent rater–assessed minimum clinically important difference, which has not been examined in previous studies. Finally, the methods used to estimate the minimum clinically important difference in this study evaluate a sample as a whole, and caution must be used when applying the minimum clinically important difference results to individual patients as might be attempted in measurement-guided treatment initiatives.[34]

In conclusion, we estimated the minimum clinically important difference for the PANSS using a more representative sample than previous studies and novel analytic techniques. We estimate a minimum clinically important difference of approximately 15 points or 34% of the baseline value of the 0-based PANSS in the CATIE sample. Our estimates also varied considerably when stratified by baseline

psychopathology and when anchored to a patient-reported illness severity measure. We found that patients with schizophrenia perceive their symptoms as less severe and judge smaller symptom changes as likely to be clinically important compared to formally trained raters. Our results may give clinicians and researchers a greater understanding of a commonly used measure of schizophrenia symptomatology, especially when the instrument is used as an outcome measure. These findings may allow a more informed assessment of the meaning of change in both research reports and clinical practice.

Author affiliations: Department of Psychiatry, Yale School of Medicine (Dr Hermes); Department of Statistics, Yale University (Mr Sokoloff), New Haven, Connecticut; Program for Intervention Effectiveness Research, New York State Psychiatric Institute, New York (Dr Stroup); and VA New England Mental Illness Research, Education and Clinical Center and the Child Study Center Yale Medical School, and VA Connecticut Health Care System, West Haven, Connecticut (Dr Rosenheck).
Potential conflicts of interest: None reported.
Funding/support: This analysis was supported by the New England Mental Illness Research and Education Center and in part by a grant (N01 MH90001) from the National Institute of Mental Health.
Role of sponsor: The funding source had no role in the design, analysis, or interpretation of data or in the preparation of the report or decision to publish.
Acknowledgments: The authors would like to thank Larry Price, PhD, Texas State University, San Marcos, Texas, and Rolf Engel, PhD, Ludwig Maximillian's University, Munich, Germany, for their assistance on equipercentile matching. The authors would also like to thank Elina Stefanovics, PhD, VA New England Mental Illness Research, Education and Clinical Center, and VA Connecticut Health Care System, West Haven, Connecticut, for her help in data management and analysis. Drs Price, Engel, and Stefanovics declare no conflicts of interest in relation to this research.

REFERENCES

1. Kay SR, Fiszbein A, Opler LA. The positive and negative syndrome scale (PANSS) for schizophrenia. *Schizophr Bull*. 1987;13(2):261–276.
2. Opler LA, Ramirez PM. Use of the Positive and Negative Syndrome Scale (PANSS) in Clinical Practice. *J Prac. Psych. and Behav Hlth*. 1998; 4:157–162.
3. Fleischhacker WW, Kemmler G. The clinical relevance of percentage improvements on the PANSS score. *Neuropsychopharmacology*. 2007; 32(11):2435–2436.
4. Burlingame GM, Dunn TW, Chen S, et al. Selection of outcome assessment instruments for inpatients with severe and persistent mental illness. *Psychiatr Serv*. 2005;56(4):444–451.
5. Crosby RD, Kolotkin RL, Williams GR. Defining clinically meaningful change in health-related quality of life. *J Clin Epidemiol*. 2003;56(5): 395–407.
6. Jaeschke R, Singer J, Guyatt GH. Measurement of health status: ascertaining the minimal clinically important difference. *Control Clin Trials*. 1989;10(4):407–415.
7. Osoba D, Rodrigues G, Myles J, et al. Interpreting the significance of changes in health-related quality-of-life scores. *J Clin Oncol*. 1998;16(1): 139–144.
8. Lydick E, Epstein RS. Interpretation of quality of life changes. *Qual Life Res*. 1993;2(3):221–226.
9. Guy W. *ECDEU Assessment Manual for Psychopharmacology*. Rockville, MD: US Dept. of Health, Education, and Welfare, Public Health Service, Alcohol, Drug Abuse, and Mental Health Administration, National Institute of Mental Health, Psychopharmacology Research Branch, Division of Extramural Research Programs; 1976.
10. Wyrwich KW, Tierney WM, Wolinsky FD. Further evidence supporting an SEM-based criterion for identifying meaningful intra-individual changes in health-related quality of life. *J Clin Epidemiol*. 1999;52(9): 861–873.
11. Levine SZ, Rabinowitz J, Engel R, et al. Extrapolation between measures of symptom severity and change: an examination of the PANSS and CGI. *Schizophr Res*. 2008;98(1–3):318–322.
12. Schennach-Wolff R, Obermeier M, Seemüller F, et al. Does clinical judgment of baseline severity and changes in psychopathology depend on the patient population? results of a CGI and PANSS linking analysis in a naturalistic study. *J Clin Psychopharmacol*. 2010;30(6):726–731.
13. Rabinowitz J, Mehnert A, Eerdekens M. To what extent do the PANSS and CGI-S overlap? *J Clin Psychopharmacol*. 2006;26(3):303–307.
14. Leucht S, Kane JM, Kissling W, et al. What does the PANSS mean? *Schizophr Res*. 2005;79(2–3):231–238.
15. Cramer J, Rosenheck R, Xu W, et al; Department of Veterans Affairs Cooperative Study Group on Clozapine in Refractory Schizophrenia. Detecting improvement in quality of life and symptomatology in schizophrenia. *Schizophr Bull*. 2001;27(2):227–234.
16. Stroup TS, McEvoy JP, Swartz MS, et al. The national institute of mental health clinical antipsychotic trials of intervention effectiveness (CATIE) project: schizophrenia trial design and protocol development. *Schizophr Bull*. 2003;29(1):15–31.
17. Lieberman JA, Stroup TS, McEvoy JP, et al; Clinical Antipsychotic Trials of Intervention Effectiveness (CATIE) Investigators. Effectiveness of antipsychotic drugs in patients with chronic schizophrenia. *N Engl J Med*. 2005;353(12):1209–1223.
18. Swartz MS, Perkins DO, Stroup TS, et al. Assessing clinical and functional outcomes in the Clinical Antipsychotic Trials of Intervention Effectiveness (CATIE) schizophrenia trial. *Schizophr Bull*. 2003;29(1):33–43.
19. Kolen MJ, Brennan RL. *Test Equating. Methods and Practices*. New York, NY: Springer; 1995.
20. Linn RL. Linking results of distinct assessments. *Appl Meas Educ*. 1993; 6(1):83–102.
21. Albano A. Equate: Statistical Methods for Test Score Equating. R package version 1.1-2; 2011. http://cran.r-project.org/web/packages/equate/index.html. Accessed January 10, 2012.
22. R Development Core Team. *R: A language and environment for statistical computing*. Vienna, Austria: R Foundation for Statistical Computing; 2011. http://www.R-project.org. Accessed January 10, 2012.
23. Wyrwich KW. Minimal important difference thresholds and the standard error of measurement: is there a connection? *J Biopharm Stat*. 2004;14(1): 97–110.
24. Duru G, Fantino B. The clinical relevance of changes in the Montgomery-Asberg Depression Rating Scale using the minimum clinically important difference approach. *Curr Med Res Opin*. 2008;24(5):1329–1335.
25. McHorney CA, Tarlov AR. Individual-patient monitoring in clinical practice: are available health status surveys adequate? *Qual Life Res*. 1995;4(4):293–307.
26. Eisen SV, Ranganathan G, Seal P, et al. Measuring clinically meaningful change following mental health treatment. *J Behav Health Serv Res*. 2007; 34(3):272–289.
27. Leucht S, Kane JM, Kissling W, et al. Clinical implications of brief psychiatric rating scale scores. *Br J Psychiatry*. 2005;187(4):366–371.
28. Wolinsky FD, Wan GJ, Tierney WM. Changes in the SF-36 in 12 months in a clinical sample of disadvantaged older adults. *Med Care*. 1998;36(11): 1589–1598.
29. Wyrwich KW, Nienaber NA, Tierney WM, et al. Linking clinical relevance and statistical significance in evaluating intra-individual changes in health-related quality of life. *Med Care*. 1999;37(5):469–478.
30. Jacobson NS, Follette WC, Revenstorf D. Psychotherapy outcome research: methods for reporting variability and evaluating clinical significance. *Behav Ther*. 1984;15(4):336–352.
31. Jacobson NS, Truax P. Clinical significance: a statistical approach to defining meaningful change in psychotherapy research. *J Consult Clin Psychol*. 1991;59(1):12–19.
32. Leucht S, Kane JM, Etschel E, et al. Linking the PANSS, BPRS, and CGI: clinical implications. *Neuropsychopharmacology*. 2006;31(10):2318–2325.
33. Atkinson M, Zibin S, Chuang H. Characterizing quality of life among patients with chronic mental illness: a critical examination of the self-report methodology. *Am J Psychiatry*. 1997;154(1):99–105.
34. Trivedi MH, Daly EJ. Measurement-based care for refractory depression: a clinical decision support model for clinical research and practice. *Drug Alcohol Depend*. 2007;88(suppl 2):S61–S71.

Editor's Note: We encourage authors to submit papers for consideration as a part of our Early Career Psychiatrists section. Please contact Marlene P. Freeman, MD, at mfreeman@psychiatrist.com.

 © COPYRIGHT 2012 PHYSICIANS POSTGRADUATE PRESS, INC.