University of Wollongong

Research Online

Faculty of Health and Behavioural Sciences -
Papers (Archive)

Faculty of Science, Medicine and Health

1-1-2010

Calculating clinically significant change: Applications of the Clinical Global Impressions (CGI) Scale to evaluate client outcomes in private practice

Peter Kelly
University of Wollongong, pkelly@uow.edu.au

Follow this and additional works at: https://ro.uow.edu.au/hbspapers

🌀 Part of the Arts and Humanities Commons, Life Sciences Commons, Medicine and Health Sciences Commons, and the Social and Behavioral Sciences Commons

Recommended Citation

Calculating clinically significant change: Applications of the Clinical Global Impressions (CGI) Scale to evaluate client outcomes in private practice

Abstract

The Clinical Global Impressions (CGI) scale is a therapist-rated measure of client outcome that has been widely used within the research literature. The current study aimed to develop reliable and clinically significant change indices for the CGI, and to demonstrate its application in private psychological practice. Following the guidelines developed by Clement, a file review was conducted of the authors' first six years working in private practice. A reliable change on the CGI required the participants score to change by 2-points. Depending on the method used to calculate the clinical change indices, between 23% and 50% of the total participants demonstrated reliable and clinical significant improvement. The CGI proved to be a useful tool to retrospectively evaluate clinical outcomes in private practice.

Keywords

applications, change, clinical, client, cgi, practice, private, global, significant, evaluate, scale, outcomes, calculating, clinically, impressions

Disciplines

Arts and Humanities | Life Sciences | Medicine and Health Sciences | Social and Behavioral Sciences

Publication Details

This is a copy of the author's final peer-reviewed manuscript as accepted for publication. The article was published in Clinical Psychologist. This article may not exactly replicate the final version published in the journal. It is not the copy of record.

Please cite:

Kelly, P.J. (2010). Calculating Clinically Significant Change: Applications of the Clinical Global Impressions (CGI) Scale to Evaluate Client Outcomes in Private Practice. *Clinical Psychologist*, 14, 107-111.

Calculating Clinically Significant Change: Applications of the Clinical Global

Impressions (CGI) Scale to Evaluate Client Outcomes in Private Practice.

Abstract

The Clinical Global Impressions (CGI) scale is a therapist rated measure of client outcome that has been widely used within the research literature. The current study aimed to develop reliable and clinically significant change indices for the CGI, and to demonstrate its application in private psychological practice. Following the guidelines developed by Clement (1994), a file review was conducted of the authors first 6-years working in private practice. A reliable change on the CGI required the participants score to change by 2-points. Depending on the method used to calculate the clinical change indices, between 23% and 50% of the total participants demonstrated reliable and clinical significant improvement. The CGI proved to be a useful tool to retrospectively evaluate clinical outcomes in private practice.

Calculating Clinically Significant Change: Applications of the Clinical Global Impressions (CGI) Scale to Evaluate Client Outcomes in Private Practice.

It is important that private practitioners evaluate their clinical work. The feedback obtained through evaluation provides the psychologist with the opportunity to refine their clinical skills and improve treatment for clients attending their practice (Clement, 1994; Nezu, 1996). In one of the first published studies of its kind, Clement (1994) outlined a file audit process he used to evaluate his 26 years of work within a private psychological setting. As a component of this review he made retrospective ratings of client improvement on a 5-point Likert scale developed for the study. This scale ranged from 1 (much worse than at intake) to 5 (much improved). Clement was then able to use these findings to identify the types of clinical problems he was most effective in treating, and establish areas to improve his overall clinical practice. The study provides a simple and straightforward method for practitioners to evaluate their clinical work, with the retrospective nature of the study being particularly suited to clinicians who have not routinely collected outcome data. Although it was recommended that clinicians should replicate this design within their own private practice (Clement, 1994), only one other study has used Clement's rating scale (see Lin, 1998). Subsequently, the potential for comparison is greatly reduced.

An alternate measure that private practitioners could consider using is the Clinical Global Impressions scale (CGI; Guy, 1976). Like Clement's scale, it provides a clinician rated measure of overall global improvement. However, the CGI has the advantage of being an extremely widely used within the research literature, it has been shown to demonstrate good inter-rater reliability (Lyons Reardon, Cukrowicz, Reeves, & Joiner, 2002) and can be used to calculate reliable and clinically significant change.

Reliable and Clinically Significant Change

Traditionally psychotherapy research has focused on statistically significant change as a method to evaluate client outcomes. However, this approach does not provide information regarding how clinically meaningful the results are or provide information on the proportion of clients who demonstrate clinical improvement. Jacobson and Traux (1991) proposed a method to examine clinically significant change. The first step involves ensuring that the change is reliable, and not simply attributable to chance. The second step involves determining the proportion of clients who move from a position that is typical of a dysfunctional population, to a population that is more functional. There are three different ways that this can be calculated: (a) the person's score moving 2 standard deviations away from the dysfunctional population mean, (b) the post-treatment score falling within 2-standard deviations of a functional population mean, or (c) the post treatment scores placing the person closer to the functional population mean than the dysfunctional population mean. The calculation of reliable and clinically significant change provides an opportunity for clinicians to compare their results to other published studies. As CGI ratings are made at both intake and discharge, there is the potential to calculate reliable and clinically significant change using the CGI. The current study examines the first 6-years of a part-time private psychological practice operated by the author within a primary care setting. The aim of the research was to develop reliable and clinically significant change indices for the CGI that could be used to evaluate client outcomes within private practice.

Method

Participants

The private psychological practice is located in the Illawarra, a regional area of New South Wales, Australia (population = 410,000). The author is a registered

Psychologist and used a cognitive-behavioral approach during the research period. The private practice was operated for approximately 1.5 days per week within a primary care facility (i.e. general practice). Two hundred and eleven participants attended the practice between July 2003 and June 2009. Participants signed a confidentiality and privacy statement as a component of the initial intake paperwork. This included a statement indicating that de-identified data might be used for research purposes. One client did not agree to provide consent. Of the remaining 210 participants, 90 (43%) came for assessment or education only and 120 people (57%) commenced treatment. The following analysis focuses on just those individuals who entered treatment.

Measures

At the first appointment all participants were asked to complete an initial Intake Form. It included basic demographic information, referral source, and previous experiences with counseling. Participants also completed the 21-item version of the Depression, Anxiety and Stress Scale (DASS; Lovibond & Lovibond, 1995). The DASS is a widely used measure of psychological distress and it provides subscale scores for the domains of Depression, Anxiety and Stress. The DASS was only used at intake, and as such could not be used to calculate reliable and clinically significant change.

The Clinical Global Impressions scale (CGI; Guy, 1976) is a clinician rated measure of symptom severity and client improvement. The current study focuses on the Severity item of the CGI, as this is the item that is most commonly reported in the research literature. It requires the clinician to rate the Severity of the person's illness (i.e. Consider your clinical experience, how mentally ill is the patient at this time?) and is rated from 1 (Normal, not ill) to 7 (Extremely). Within the current study, CGI severity scores were calculated for the person based on their presentation at intake and

then again for the final session that the person attended. The CGI is considered a valid tool to measure clinical outcomes for routine clinical use (Berk, et al., 2008) and has previously been used to retrospectively rate client improvement based on a file audit (Lyons Reardon, et al., 2002).

Procedure

Following the procedures of Clement (1994), a detailed review was conducted of each client file. Each client was assigned one primary problem area that constituted the main focus of treatment. The CGI intake and discharge scores were retrospectively rated for the first 92 participants that entered the service. CGI scores were routinely rated during treatment for the remaining 30 participants in the study. To examine the validity of the retrospective ratings, DASS scores were not reviewed as part of the file audit and were subsequently compared to CGI intake scores.

Data analysis

To evaluate the size of changes in global functioning, effect sizes were also calculated using the following formula:

$$\underline{\text{Mean (CGI Severity Intake – CGI Severity Discharge)}}$$

$$\text{SD}_{\text{CGI Severity Intake}}$$

Reliable change indices were calculated using the Reliable Change Index (RCI; Christensen & Mendoza, 1986). This is calculated as:

$$RC = (1.96 \times SE_{\text{diff}})$$

Where the Standard Error (SE) of measurement is calculated as:

$$SE_{\text{diff}} = SD_{\text{CGI Intake}} \times \sqrt{2} \times \sqrt{(1 - \alpha)}$$

As inter-rater reliability was not examined in the current research, an alpha co-efficient was used from a study that examined inter-rater reliabilities between doctoral-

level psychology students who also retrospectively rated mental health client files (α = .84) (Lyons Reardon, et al., 2002).

The methods outlined by Jacobson and Traux (1991) were used to calculate the clinical significance cut-off scores. The first approach required the person to move away from the dysfunctional population mean by two standard deviations (criterion a):

$$CS_{cut\text{-}off} = 2 \times SD_{CGI\ Intake}$$

As the distribution of the CGI scores in the current study were overlapping, Jacobson and Traux (1991) recommend that criterion c be used instead of criterion b. As such, the other method used to calculate the clinical cut off scores required the client to move closer to the functional population mean by crossing the midpoint between the dysfunctional and functional population means (criterion c). Only participants who began closer to the dysfunctional population mean were included in this analysis. This is calculated by using the following formula:

$$CS_{cut\text{-}off} = \frac{(mean_{clin} \times SD_{norm}) + (mean_{norm} \times SD_{clin})}{(SD_{norm} + SD_{clin})}$$

Where the $Mean_{clin}$ and SD_{clin} is the mean and standard deviation for the dysfunctional population (the current study), and the $Mean_{norm}$ and SD_{norm} is the mean and standard deviation for the functional population.

Within the current study, the functional population was considered to be participants of a study who had completed a clinical trial that involved cognitive therapy for depressive and anxiety disorders (Foreman, Herbert, Moitra, Yeomans, & Geller, 2007). On average, the participants in this study were 27 years of age and had attended 15-sessions of counseling (Foreman, et al., 2007). Eighty-percent of the participants in the study were female. Whilst this study compared Acceptance and Commitment Therapy with Cognitive Therapy, only the Cognitive Therapy condition

was included in the current analysis. At intake to the study participants were rated between "Mildly Ill" and "Moderately Ill" (M = 3.19, SD = 1.32). At post-treatment participants were rated between "Borderline Mentally Ill" and "Mildly Ill" (M = 2.75, SD = 1.28). The mean discharge CGI score and standard deviation were used to calculate the clinical cut off scores.

Results

Description of Participants attending the practice

The majority of participants were female (66%), and ranged in age between 14 and 74 years (M = 33.90). At intake, 43% of participants reported recently or currently taking medication for their mental health problems, 20% reported previous self-harming behavior, 11% had previously made a suicide attempt, 21% had current suicidal thoughts and 9% had previously been hospitalized for mental health problems.

Average ratings on the CGI scores at intake indicated that participants were 'Mildly' to 'Moderately Ill' (M = 3.77, SD = 1.27). One client was rated as 'Normal, Not Ill At All' (<1%), 19 were 'Borderline Mentally Ill' (16%), 35 'Mildly Ill' (29%), 'Moderately Ill' (25%), 24 'Markedly Ill' (20%), 9 'Severely Ill' (8%), and 2 'Extremely Ill' (2%). Of the 120 people who commenced treatment, 114 completed the DASS at intake (95%). Client ratings of symptom severity on the three DASS subscales indicate that on average participants psychological distress was Moderate to Severe (Depression, M = 19.06, SD = 9.63; Anxiety, M = 15.06, SD = 9.09; Stress, M = 23.02, SD = 9.04). The median number of counseling sessions attended was 5.50; with the mean duration being 7.58 (SD = 5.91) sessions. The majority of people were treated for a mood disorder (42.6%), followed by anxiety disorders (33.6%), substance abuse problems (11.5%), and psychotic disorders (1.6%). Twenty-one percent

completed treatment in 3 or fewer sessions, 57% within 6-sessions, 80% within 9-sessions, 85% within 12-sessions, and 100% within 37-sessions.

Retrospective ratings

In an attempt to examine the validity of the retrospective ratings, independent samples T-tests were used to compare those participants where retrospective CGI ratings were required (n = 92) with participants whose CGI ratings were made following each session (n = 30). There were no statistically significant differences at the .01 level for the CGI Intake score (t = -0.77, df = 118), or CGI Discharge score (t = 0.77, df = 115). Pearsons correlations were also conducted between the DASS subscale scores and the CGI intake score as a further method to examine the validity of the retrospective ratings. These correlations were of a small to moderate range and were statistically significant at the .01 level (Depression, r (111) = 0.36; Anxiety, r (111) = 0.37; Stress, r (111) = 0.24).

Statistical, reliable and clinically significant change

On average, there was a statistically significant improvement on the CGI for all participants accessing the service (t = 12.62, df = 115, p < .01).

For participants to demonstrate Reliable Change there needed to be a 2-point reduction in the persons CGI score from intake to discharge. A 3-point reduction was required for the person to demonstrate clinically significant change using the 2-standard deviation approach (Criterion a). Using Criterion c, participants were required to both demonstrate reliable change (i.e. a 2-point reduction), and to move from a position closer to the dysfunctional population mean (above 3.23) to a position closer to the functional population mean (below 3.23). Only participants who commenced treatment with an intake CGI score of 4 or above were included in the subsequent analysis (n = 65). It would have been impossible for participants with scores below 4

to demonstrate clinical improvement (Lambert & Ogles, 2009; Westbrook & Kirk, 2004).

Of the total sample, 55% of participants' demonstrated reliable improvement, (see Table 1). No clients demonstrated reliable deterioration in functioning. Two methods were used to calculate clinically significant change. When using the more stringent approach of a move of 2 standard deviations (criterion a), 23% of the total sample demonstrated reliable change. When comparing to the normative data (criterion c), 50% of the total sample demonstrated reliable and clinically significant change.

Insert Table 1 approximately here

Discussion

Within the current study, participants demonstrated a statistically significant improvement at the .01 level. Whilst this result is encouraging, it provides limited information about the proportion of clients who demonstrate change that is considered to be clinically meaningful. The calculation of reliable and clinically significant change provides an opportunity to identify the proportion of clients who demonstrate clinical improvement or deterioration. As has been found in previous research (Murugesan, 2007), the proportion of participants who demonstrated reliable and clinically significant change varied depending on the criteria used to calculate the change scores. As criterion c uses data from both dysfunctional and functional populations it provides a more "precise determination of which population a subject's score belongs in" (Jacobson & Truax, 1991, p. 13). As such, criterion c is likely to provide the most clinically relevant information.

A key benefit of calculating reliable and clinically significant change is the ability for clinicians to benchmark their results against other published data. For example, using the Beck Depression Inventory, Westbrook and Kirk (2004)

benchmarked the proportion of clients in routine care who demonstrated clinically significant change against both another private practice and a major research trial. This type of analysis provides information on the degree to which results from controlled clinical trials can be translated into routine care. It is also likely to provide individual practitioners with valuable information regarding their own clinical practices. Unfortunately, it was not possible to benchmark the current study in this way as previous researchers have failed to report rates of clinical change for the CGI (e.g. Foreman, et al., 2007). It is not advisable to benchmark change rates between different measures, as rates of change vary between different outcome measures (Murugesan, 2007). With increasing calls for researchers to report rates of reliable and clinically significant change (Lambert & Ogles, 2009), it is hoped that researchers will begin to report these rates. However, the results from the current study might be useful for other private practitioners to generate their own clinical change indices, or to conduct their own benchmarking activities. This will be particularly relevant for psychologists working in primary care facilities, as participants within the current study appear to be fairly representative of clients accessing these types of services (Knowles & McMahon, 1995; Vines, et al., 2004). It is important to note that the results from the current study do not necessarily represent 'best practice' (Hermann & Provost, 2003). As private practitioners use the CGI more regularly and calculate clinical change, the opportunity to establish more meaningful benchmarks will substantially increase.

A major limitation with the current work was the reliance on retrospective ratings of client outcome. Retrospective ratings are highly reliant on the amount of information originally collected and influenced by recall bias. Additionally, the study has only examined the therapists' perspective. Both the study, and possibly the quality of the clinical work conducted, would have been improved by conducting session-by-

session ratings of client progress that incorporated both therapist and client perspectives. For example, brief ratings of improvement could have routinely been made at each session with the CGI and a brief client rated measure, such as the DASS or the Session Rating Scale (Lovibond & Lovibond, 1995; Miller, Duncan, Brown, Sparks, & Claud, 2003). However, for practitioners who have not routinely used outcome measures to evaluate their clinical work, a retrospective file-audit is likely to be a very useful starting place to commence evaluation.

The method outlined by Clement (1994) provided a straightforward approach to conducting an evaluation within a private practice. Calculating reliable and clinically significant change extended this method, providing an opportunity to identify the proportion of people who demonstrated clinically meaningful improvement. The most illuminating aspect of conducting the current study was the process of re-reading the case notes. It provided a rare opportunity to critically reflect on the range of clinical work conducted within the private practice, and highlighted clear areas for improvement (e.g. to regularly use outcome assessment). To echo the sentiment of Clement (1994), clinicians are encouraged to review and publish their clinical work. Whilst this information will be useful to the field in general, it is likely to also be a very insightful experience for the clinician.

References

Berk, M., Ng, F., Dodd, S., Callaly, T., Campbell, S., Bernardo, M., et al. (2008). The validity of the CGI severity and improvement scales as measures of clinical effectiveness suitable for routine clinical use. Journal of Evaluation in Clinical Practice, 14(6), 979-983.

Christensen, L., & Mendoza, J. L. (1986). A Method of Assessing Change in a Single Subject - an Alteration of the Rc Index. Behavior Therapy, 17(3), 305-308.

Clement, P. W. (1994). Quantitative-Evaluation of 26 Years of Private-Practice. Professional Psychology-Research and Practice, 25(2), 173-176.

Foreman, E. M., Herbert, J. D., Moitra, E., Yeomans, P. D., & Geller, P. A. (2007). A randomized controlled effectiveness trial of acceptance and commitment therapy and cognitive therapy for anxiety and depression. Behavior Modification, 31(6), 772-799.

Guy, W. (Ed.). (1976). Clinical Global Impressions. Rockville, MI: National Institute of Mental Health.

Hermann, R. C., & Provost, S. (2003). Interpreting measurement data for quality improvement: Standards, means, norms, and benchmarks. Psychiatric Services, 54(5), 655-657.

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

Knowles, A. D., & McMahon, M. (1995). Expectations and preferences regarding confidentiality in the psychologist and client relationship. Australian Psychologist, 30(3), 175-178.

Lambert, M., & Ogles, B. M. (2009). Using clinical significance in psychotherapy outcome research: The need for a common procedure and validity data. Psychotherapy Research, 19(4-5), 493-501.

Lin, J. C. H. (1998). Descriptive characteristics and length of psychotherapy of Chinese American clients seen in private practice. Professional Psychology-Research and Practice, 29(6), 571-573.

Lovibond, S. H., & Lovibond, P. F. (1995). Manual for the Depression Anxiety Stress Scales (DASS). Sydney: Psychology Foundation.

Lyons Reardon, M., Cukrowicz, K. C., Reeves, M. D., & Joiner, T. E. (2002). Duration and regularity of therapy attendance as predictors of treatment outcome in an adult outpatient population. Psychotherapy Research, 12(3), 273-285.

Miller, S. D., Duncan, B. L., Brown, J., Sparks, J. A., & Claud, D. A. (2003). A preliminary study of the reliability, validity, and feasibility of a brief visual analog measure. Journal of Brief Therapy, 2, 91-100.

Murugesan, G., Amey, C., Deane, F., Jeffery, R., Kelly, B. & Stain, H. (2007). Inpatient psychosocial rehabilitation in rural NSW: assessment of clinically significant change for people with severe mental illness. Australian and New Zealand Journal of Psychiatry, 54(5), 655-657.

Nezu, A. M. (1996). What are we doing to our patients and should we care if anyone else knows? Clinical Psychology-Science and Practice, 3(2), 160-163.

Vines, R. F., Richards, J. C., Thomson, D. M., Brechman-Toussaint, M., Kluin, M., & Vesely, L. (2004). Clinical psychology in general practice: a cohort study. Medical Journal of Australia, 181(2), 74-77.

Westbrook, D., & Kirk, J. (2004). The clinical effectiveness of cognitive behaviour

therapy: outcome for a large sample of adults treated in routine practice. .

Behaviour Research and Therapy, 43, 1243-1261.

Table 1

Client Mean CGI scores, Effect Sizes, Reliable and Clinically Significant Change

	n	Total Sessions (SD)	Intake M (SD)	Discharge M (SD)	Diff (SD)	Effect Size (d)	Reliable Change[1]	% Improved 2SDs[2]	% Improved Crossed midpoint[3]
Total sample	65	9.12 (7.14)	4.74 (0.82)	3.14 (1.46)	1.61 (1.23)	0.98	55%	23%	50%
Gender									
Males	20	9.80 (8.43)	4.90 (0.91)	3.32 (1.34)	1.63 (1.26)	0.90	53%	26%	37%
Females	45	8.82 (6.57)	4.67 (0.77)	3.07 (1.51)	1.60 (1.23)	1.04	56%	22%	56%
Primary problem									
Depressive	34	9.03 (7.12)	4.69 (0.93)	2.85 (1.58)	1.85 (1.33)	0.99	68%	29%	62%
Anxiety	19	9.16 (6.47)	4.74 (0.65)	3.47 (1.43)	1.26 (1.10)	0.97	37%	16%	37%

Note. All participants commenced treatment with a CGI intake score 4 or higher. No clients demonstrated a reliable deterioration. [1]This required a 2-point CGI reduction. [2]This required a 3-point CGI reduction. [3]This required participants to demonstrate reliable change, begin closer to the dysfunctional population mean and finish closer to the functional population mean.