HHS Public Access

Author manuscript

Clin Psychol Psychother. Author manuscript; available in PMC 2024 July 12.

Published in final edited form as:

Clin Psychol Psychother. 2022 January ; 29(1): 339–350. doi:10.1002/cpp.2606.

Operationalizing alliance rupture–repair events using control chart methods

Lauren M. Lipner[1], **J. Christopher Muran**[2,1], **Catherine F. Eubanks**[3,1], **Bernard S. Gorman**[2], **Arnold Winston**[4]

[1]Brief Psychotherapy Research Program, Mount Sinai Beth Israel, New York, New York, USA

[2]Derner School of Psychology, Adelphi University, Garden City, New York, USA

[3]Ferkauf Graduate School in Psychology, Yeshiva University, New York, New York, USA

[4]Department of Psychiatry, Icahn School of Medicine at Mount Sinai, New York, New York, USA

Abstract

This study aimed to determine how control charts—a form of time-series line graphs—can be implemented in psychotherapy research to indirectly identify probable rupture–repair episodes that are associated with psychotherapy outcome. There is no current standard in psychotherapy research with regard to how to use control charts to identify rupture–repair events. Control charts were generated for each patient ($N = 73$) using patient-rated Working Alliance Inventory (WAI) scores obtained at the end of every session in a 30-session therapy protocol of either brief relational therapy (BRT) or cognitive behavioural therapy (CBT). Empirically derived cut-off points were used to identify rupture and repair based on each dyad's control chart. Coded rupture–repair episodes were correlated with outcome measures to assess for their relationships. The results of these analyses provide preliminary support for the utility of control charts in psychotherapy research for the indirect identification of probable rupture repair events that are associated with psychotherapy outcome.

Keywords

alliance; outcome & methods; process; rupture

1 | INTRODUCTION

The therapeutic alliance has been consistently found to be one of the most robust predictors of treatment retention and therapeutic outcome, even when controlling for various potential confounding variables such as patient characteristics and treatment processes (Flückiger et al., 2018, 2020; Horvath et al., 2011; Horvath & Symonds, 1991). In his seminal transtheoretical conceptualization, Bordin (1979) emphasized purposeful collaboration on

Correspondence: Lauren M. Lipner, Brief Psychotherapy Research Program, Mount Sinai Beth Israel, 10 Nathan D Perlman Place, New York, NY 10003, USA., lmlipner@gmail.com.

CONFLICT OF INTEREST

The authors report no conflict of interest.

tasks and goals, and the affective bond between patient and therapist, as essential. Based on the results of their meta-analysis, Flückigeret al. (2018; Horvath & Symonds, 1991) concluded that the relation between the working alliance and therapeutic outcome does not appear to be a function of the modality of therapy employed, length of treatment, or other psychotherapy process factors but, instead, a predictive factor in its own right. As such, the construct of the therapeutic alliance remains central in the psychotherapy literature in contributing to a greater understanding of why some patients prematurely terminate treatment, and how best to maximize potential therapeutic outcome.

Throughout the course of treatment, ruptures in the therapeutic alliance occur commonly, and may contribute to the development of an overall weakened alliance (Eubanks et al., 2018). A rupture, in the context of Bordin's (1979) definition of the alliance, can be defined as a strain in the alliance as manifested by a lack of collaboration on goals and tasks and/or a deterioration in the emotional bond. Although the term "rupture" implies a major break in the therapeutic relationship, it is also used to describe the subtle tensions that often occur within the dyad that may go unnoticed by the therapist. When ruptures are identified and addressed by the therapist, the opportunity for successful resolution arises, allowing for getting the treatment back on track, or a means for a corrective emotional experience for the patient, which could have implications for overall change (Eubanks-Carter et al., 2010; Safran & Muran, 2000).

However, the literature suggests that therapists are often unsuccessful in identifying when a rupture has occurred in a session and that patients can sometimes be unwilling or unable to be straightforward with the therapist regarding any disagreements or discomfort (Rennie, 1994; Safran et al., 2002). Regan and Hill (1992) instructed patients and therapists to report on thoughts or feelings that went unsaid in therapy sessions and found only a 17% match between therapist guesses and patients' report. In a later study, when patients were asked to discuss their own treatment, they disclosed negative feelings about their therapists. However, when these patients were uncomfortable addressing these issues with their therapists directly, they concealed them such that they went unnoticed by the therapist, and thus unaddressed, often leading to premature dropout (Rhodes et al., 1994). Another study found that even experienced therapists were only able to guess when a patient had withheld negative feelings 45% of the time (Hill et al., 1993). Hill et al. (1996) found that therapists were often unaware that patients were unsatisfied until after treatment termination. This lack of clarity on the occurrence of rupture on the part of therapists results in researchers using various methods and techniques to infer the presence of rupture in a session.

Approaches to rupture identification can be categorized into two methods: within-session methods and between-session methods, which derive from postsession measures of alliance (Eubanks-Carter et al., 2010). Within-session methods refer to those that assess moment-to-moment process to assess for the occurrence of rupture, such as the Rupture Resolution Rating Scale (3RS; Eubanks et al., 2015) and the Collaborative Interactions Scale (CIS; Colli & Lingiardi, 2009). Between-session methods, or indirect self-report methods, glean the occurrence of rupture from postsession alliance data. Between-session methods, based on postsession measures of alliance, are widely used in psychotherapy research, in part

due to their efficiency and their ability to be applied to archival data. These methods are classified as indirect methods because these methods glean the occurrence of rupture based on some other measure, in comparison to approaches which directly ask therapists or patients whether a rupture has occurred over the course of a session, such as those employed by Muran et al. (2005). Due to their widespread application, this study will focus on indirect between-session methods. Previous studies have implemented various statistical between-session methods to identify the occurrence of ruptures in treatment using between-session self-report measures of alliance. Stiles et al. (2004) coded a rupture when the patient's reported score on the Agnew Relationship Measure (ARM; Agnew-Davies et al., 1998) was lower than the predicted value calculated from the client's intercept and slope by two standard deviations. Both Strauss et al. (2006) and McLaughlin et al. (2014) statistically determined a cut-off point based on measures of central tendency and a clinical significance formula respectively, using the California Psychotherapy Alliance Scales (CALPAS), such that when a score on this measure was lower than the determined cut-off point, a rupture was coded. A rupture repair was coded with a subsequent increase in the CALPAS rating. In both of these studies, the CALPAS was not administered at the end of every session in either study, but rather at predetermined time points. In Strauss et al. (2006), alliance was measured up to eight times in a maximum of 52 sessions treatment protocol (i.e., sessions 2, 5, 10, 20, 30, 40, 50, 52). In the McLaughlin et al. study, therapeutic alliance was measured in all even-numbered sessions (i.e., 2, 4, 6, 8, 10). In yet another method, both Stevens et al. (2007) and Larsson et al. (2016) coded a rupture following a drop in Working Alliance Inventory (WAI) score by one or more points from the previous session's WAI score. While many of these studies yielded meaningful conclusions regarding psychotherapy process, none evaluated the method of rupture identification itself to determine their ability to pinpoint clinically meaningful process. See Lipner (2020) for a comprehensive review of rupture identification strategies previously implemented in psychotherapy literature.

Though all of these studies use the term "rupture," they all employ different methods to determine when ruptures have occurred in treatment. Thus, it is likely that each method would yield inconsistent numbers of ruptures for the same therapist–patient dyad (Lipner et al., 2018). In methodologies that take a nomothetic approach, such as when a cut-off point or specific value decrease is determined to operationalize a rupture event across all patients, the information may not be applicable at the individual level as it does not allow for individual variability in reporting of alliance strength. For example, patients who report an unstable fluctuating alliance strength would appear to have many ruptures using a specified cut-off point, while patients who have a stable reporting style will show few, if any at all. A method that takes into account factors such as response style bias and intraindividual variability, rather than relying on statistically or theoretically determined cut-off points which cut across all patients in a uniform manner, would provide a more nuanced and individualized lens into the therapeutic relationship at the dyadic level. While these methods have been applied widely, the validity of these approaches has not been established. This paper aims to establish the validity of the control chart approach.

Eubanks-Carter et al. (2012) provide an overview of various quantitative naturalistic methods for detecting potential change points in psychotherapy, including criterion-based methods, partitioning methods, regression methods, and control charts. Given the previous

application of control charts to rehabilitation psychology (Callahan & Barisa, 2005) and monitoring change in behaviour analysis (Pfadt & Wheeler, 1995), this paper will focus on the application of this statistically generated line graph as a method to reliably identify ruptures in the therapeutic alliance. Control charts were initially developed for use in the manufacturing industry for the detection of unpredictable variation requiring intervention; this method of change detection has been widely used by major companies including Motorola and General Electric (Eubanks-Carter et al., 2012; Pande et al., 2000). Of the existing types of control charts, the Shewhart chart in particular has been recommended for use in clinical settings (Callahan & Barisa, 2005). On these graphs, control limits are calculated based on individual-level statistics to be an equidistant band from the mean score based on that standard deviation (referred to as a sigma level), creating a statistically defined confidence interval representative of standard deviations above and below the mean (Figure 1). Time, or session number, is plotted on the X-axis, while the dependent variable, or alliance score in this instance, is plotted on the Y-axis. Thus, similar to other methods that model intraindividual change using multiple time point assessments (e.g., multilevel modelling), control charts produce person-specific estimates (e.g., mean) of within-person variation in the variable of interest over time. The mean alliance score across all sessions is plotted as a solid line across the graph (Eubanks-Carter et al., 2012). As such, in the application of control charts in rupture identification, a rupture can be coded when a score on a measure of alliance drops below the lower control limit. The idiographic nature of this method of coding for rupture and repair episodes in the alliance allows for personalization for each patient, as control limits are set differently depending on each patient's parameters on the alliance measure. The graph generated by control charts allows the viewer to visualize the course of the therapeutic alliance as a whole as it waxes and wanes over the course of the treatment, while making it readily apparent when a rupture has occurred.

However, as outlined by Eubanks-Carter et al. (2012), there is currently no standard in psychotherapy research regarding how to set the control limits or boundaries of the confidence intervals in control charts in order to identify possible rupture events. Although the two-sigma level is commonly used, there is a potential for neglecting subtle events that could play a role in the overall therapeutic alliance, while smaller sigma levels increase the risk of Type I error. Further, there is no standard for defining a rupture–repair that is associated with outcome using this method. There are two potential methods of coding rupture–repair using control chart methods: once a session WAI score has returned to meet or exceed the mean WAI score for that particular patient, or once a session WAI score has met or exceeded the lower control limit value for that patient.

In psychotherapy research, a statistical method that would aid in tracking movement in the alliance would allow for the identification of potential change points in treatment for further examination. As such, the main aims of this study are to extend the efforts of Eubanks-Carter et al. (2012) in their identification of control charts being a useful tool to identify probable rupture and repair episodes by first confirming their utility using a larger sample and second determining where to set the control limits in order to best identify ruptures that are associated with outcome. Although previous studies in psychotherapy research have implemented various methods of rupture repair identification, these have been applied without the confirmation that the predetermined rupture criteria is identifying

process predictive of psychotherapy outcome. Control charts will be examined here using three levels of control limits in order to determine how best to configure the charts for the identification of potential change points. This method of inquiry into the application of control charts for rupture identification assesses the tool as a scientific, evaluative one rather than a live, clinical tool. This retrospective, post hoc analysis allows for an in-depth analysis of ruptures or weak alliances and, as a result, allows researchers to better understand what therapists are doing in the event of such ruptures, as well as their subsequent repair. As such, this method is not a statistical test of significance but rather a method that flags a potentially significant session worthy of further examination for rupture process, consistent with the suggestions outlined by Eubanks-Carter et al. (2012). This study is exploratory in nature, as the parameters around which control charts are configured have not been tested previously, and therefore we did not set specific hypotheses around which configuration of control charts best identify meaningful rupture events. However, we hypothesize that the three-sigma limits previously identified by Wheeler and Chambers (1992) may be too conservative and put researchers at risk of Type II error, and conversely, one-sigma limits may be too lenient and put researchers at risk of Type I error.

2 | METHOD

2.1 | Patients

Patients in this study were a subsample of those who sought psychotherapy services from the Brief Psychotherapy Research Program: an outpatient mental health clinic in a major metropolitan city. The subsample consisted of 73 outpatient participants between the ages of 21 and 77 (M = 41.9), 63% of whom were female, and 83.6% of whom were Caucasian (see Table 1 for further breakdown of patient demographics). Primary diagnoses, as measured by Structured Clinical Interview for the DSM-IV (SCID; First et al., 1995) I and II for DSM-IV semistructured diagnostic interviews, are reported in Table 1. Seventy-four percent of participants were diagnosed with comorbid Axis I and Axis II diagnoses. The primary inclusion criterion for the purposes of this study was the completion of 25 out of 30 sessions with corresponding postsession questionnaire completion and complete data collection from intake and termination measures (i.e., all outcome measures completed at both time points). Additionally, if a patient participated in psychotherapy more than once, only the patient's initial participation in the trial was included. As a result, the initial subsample of 125 was reduced to a final subsample of 73 unique patient-therapist dyads (32.9% with complete postsession questionnaire data, 46.6% missing one to three sessions, 20.5% missing four or more sessions).

2.2 | Therapists

Therapists (N = 73) participating in the study included psychology externs and interns and psychiatry residents in the psychiatry department of a major metropolitan medical centre. Therapists ranged in age from 21 to 50 (M = 31.3), 73.8% of whom were female and 80% of whom were Caucasian. All therapists attended weekly supervision in a group format throughout their participation in the study.

2.3 | Treatment models and training procedure

2.3.1 | Cognitive behavioural therapy (Turner & Muran, 1992)—This cognitive behavioural therapy (CBT) protocol, based on the Beck et al. (1990) adaptation of cognitive therapy for the treatment of personality disorders, was formulated for use with patients diagnosed with a Cluster C personality disorder or personality disorder "not otherwise specified" in a 30-session, one 45-min session per week format. The treatment protocol begins with a focus on the establishment of a case formulation, including a problem list and clarifying core belief systems. Following this portion of the treatment protocol, two phases followed (1) symptom relief or reduction, largely focused on those symptoms associated with an Axis I diagnosis, and (2) schema change, which emphasized the Axis II condition. The patient was taught skills to better handle the Axis II problem by first addressing the Axis I issues. Therapists were trained in a 16-week didactic seminar on CBT, including the observation of expert videotapes and role-playing exercises. Following case assignment, participating therapists attended weekly 75-min group supervision that focused on case formulation, treatment planning, role-play, in addition to direct feedback on therapists' videotaped sessions. All psychotherapy sessions were videotaped for review in supervision groups.

2.3.2 | Brief relational therapy (Safran & Muran, 2000)—The principles of brief relational therapy (BRT) were based on those outlined in *Negotiating the Therapeutic Alliance: A Relational Treatment Guide* (Safran & Muran, 2000), which served as the treatment manual for this protocol. The following key principles guided BRT therapists: (a) an intensive focus on the here-and-now of the therapeutic relationship, (b) an ongoing collaborative exploration of patient and therapist contributions to the therapeutic relationship, (c) the assumption that ruptures in the therapeutic alliance result from the patient and therapist co-participation in negative interpersonal process, and (d) metacommunication as a cornerstone of the psychotherapy, with particular emphasis on the self-disclosure and affective expression on the part of the therapist (Safran & Muran, 2000). Training occurred in a weekly, 75-min group supervision format. Group supervision sessions consisted of a didactic component emphasizing the definition of rupture-resolution events, as well as an experiential component including awareness exercises "as opportunities for deliberate practice" (Ericsson et al., 1993; Muran et al., 2018.

2.4 | Measures

Patient diagnostic status was determined using the SCID I and II (First et al., 1995) for DSM-IV semistructured diagnostic interviews. Previous studies suggest test–retest reliability was reported for various domains of SCID I, ranging from 0.35 to 0.78 (Zanarini et al., 2000). Test–retest reliability for the various domains of SCID II ranged from 0.24 to 0.76 (First et al., 1995).

The 12-item WIA (Tracey & Kokotovic, 1989) was administered at the end of each session to the patient and the therapist. Items were rated using a 7-point Likert scale. Previous studies have provided support for the psychometric properties of the 12-item WAI, which is widely implemented in psychotherapy research (Flückiger et al., 2018). The overall mean scores of the patient-rated WAI were used for analyses.

Outcome measures, assessed at intake and termination, included the patient- and therapist-rated Inventory of Interpersonal Problems (IIP; Horowitz et al., 1988) and Target Complaints (TC; Battle et al., 1966), as well as patient-rated Symptoms Checklist 90 – Revised (SCL-90; Derogatis, 1983) and therapist-rated Global Assessment Scale (GAS; Endicott et al., 1976). The IIP is a 127-item measure emphasizing interpersonal problems patients experience, and the distress associated with them, rated on a 5-point scale. The overall mean IIP score was used for the purposes of this study. Subscale α values in previous studies ranged from .82 to .93 at time 1, and .82 to .94 at time 2 (Horowitz et al., 1988). Total test–retest reliability was excellent (r = .98). Good convergent reliability was found with the SCL-90 (Horowitz et al., 1988).

On the TC (Battle et al., 1966), patients report three presenting problems, and both patient and therapist independently rate their severity on a Likert scale. For the purposes of this study, patient and therapist ratings of the three presenting problems were averaged for an overall patient-rated target complaint and overall therapist-rated target complaint, respectively. The TC has demonstrated good validity and sensitivity to change (Battle et al., 1966; Deane et al., 1997; Luborsky et al., 1988).

The SCL-90 (Derogatis, 1983) requires patients to rate their general psychiatric symptoms on a 5-point scale indicating their presence and associated amount of distress. The scale has been shown to have high internal consistency and reliability (Nguyen et al., 1983; Rosen, et al., 2000). The overall mean score of the Global Severity Index (GSI) was used as a measure of general psychiatric symptom severity. Derogatis (1983) reported coefficient alphas ranging from .77 to .90 and test–retest reliability scores between .80 and .90 over a 1-week period, demonstrating acceptable psychometric statistics.

The GAS (Endicott et al., 1976), a therapist-rated measure of adaptive functioning, involves a single rating on a continuum from 1, the sickest, to 100, the healthiest. The GAS has previously demonstrated good interrater reliability and validity (Endicott et al., 1976). Therapists were trained to reliable standards (intraclass correlation ≥ .90).

2.5 | Procedure

The sample of patients included in this study is a subsample from a larger psychotherapy research trial (Muran et al., 2005). For the purposes of the psychotherapy trial, patients were first screened through an intensive intake process that began with a phone screen which assessed for suicidality, parasuicidality, substance abuse, and other medical conditions. Patients cleared through the phone screening participated in SCID Axis I and Axis II interviews. The interviews were administered by graduate-level research assistants in clinical psychology who completed a training protocol including a demonstration video, role-playing, and the completion of an interrater reliability test. Research assistants were required to meet interclass correlation of .90 on both Axis I and II sections of the reliability test. Patients were excluded from participation if they met criteria for a cluster B personality disorder diagnosis, or a substance use disorder, due to concerns that these patients would require a more intensive treatment approach. Prior to the start of treatment, patients also completed a battery of questionnaires, including intake measures of symptom severity (SCL-90), presenting problems (TC), and interpersonal problems (IIP).

Patients were then assigned to one of two time-limited treatment conditions: CBT or BRT (see Muran et al., 2005, for assignment protocol). Of the subsample (N = 73) used in this study, 30 patients were randomly assigned to CBT, and 43 patients were randomly assigned to BRT. Patients paid a nominal per-session fee determined using an income-sensitive sliding scale ranging from $20 to $80. Therapists participated in one of the two treatment conditions. A postsession questionnaire was administered to both patient and therapist at the end of each session and included a 12-item version of the WAI (Tracey & Kokotovic, 1989), used to assess the quality of the working alliance. For the purposes of this study, only patient-rated WAI scores were used. At termination, both patients and therapists once again completed the outcome measure battery (patient-rated: IIP, SCL-90 and TC; therapist-rated: IIP, GAS, and TC).

2.6 | Data analysis plan

2.6.1 | Treatment outcome—In order to reduce the number of dependent variables and reduce the potential for type I error, standardized residual gain scores were calculated for all six available patient-rated outcome variables, from the intake to termination interval, consistent with previous efforts to define a core battery (Strupp et al., 1997). Principal components analysis (PCA) with varimax rotation component landings was employed in order to reduce the number of dependent variables and increase power (see Muran et al., 2005 for further explanation of this procedure). A single component was extracted from the PCA explaining 49.6% of the variance in the six indicators, which served as an outcome composite for the primary analysis, such that lower values were associated with better outcome (described below). Component loadings ranged from .60 (patient-rated IIP) to .76 (therapist-rated GAS). To assess the normality of this outcome composite variable, a Shapiro–Wilk test of normality was conducted and indicated the variable had a normal distribution (.993, p = .97).

2.6.2 | Control chart coding—SPSS was used to generate three control charts for each patient–therapist dyad based on the patient-reported WAI scores from each session at sigma levels 2, 1.5, and 1 (in a normal distribution, this would equate to the identification of potential rupture in 2.3%, 6.7%, and 15.9% of observations, respectively) (see Figure 1 for an example)[1]. Control charts were not coded if patients did not complete at least 25 out of 30 session measures or if there was a gap between completed measures by more than two sessions. Beyond this specification, any missing data were not interpolated such that control charts represent consecutive data points rather than consecutive weeks of treatment. Across both treatment conditions, WAI data completion across 30 sessions ranged between 83.33% and 100% (M = 94.19%), with no significant difference between groups.

A potential rupture was coded when a session's reported WAI score fell below the lower control limit value for each individual patient based on sigma levels 1, 1.5, and 2, respectively. Clinically, this translates to the various levels of significance in the decrease in alliance strength. A rupture was coded as having been resolved using two methods:

[1]The present analyses were based on person-centered mean data rather than time detrending in following with recommendations outlined by several authors, suggesting that detrending may suppress actual effects when variables are inherently based on time (Alfi-Yogev, et al., 2020; Falkenström, et al., 2017; Wang & Maxwell, 2015)

(1) when a WAI score surpassed the lower control limit value and (2) when a WAI score exceeded the mean WAI score across all 30 sessions, as both coding methods reflect some level of recovery (see Figure 1). This assesses the necessity of a more or less conservative approach to rupture repair coding that is associated with psychotherapy outcome. Clinically, this translates to a return to the average alliance versus a return to some less than average alliance level, but one that allows the work to continue. Finally, ruptures that took more than three sessions to resolve were considered unresolved. This is a theoretically determined limit based on the notion that beyond three sessions, a month has passed since the initial rupture and that it is unlikely that the rupture would be resolved past this time point. A rupture that was not resolved prior to the end of treatment was coded as an unresolved rupture.

As such, each control chart representative of the fluctuating patient-rated therapeutic alliance over the course of the therapy was coded in two different ways based on rupture resolution operationalization and at 3 sigma levels (i.e., sigma level 1, 1.5, and 2), resulting in six distinct rupture repair frequencies per dyad.

2.6.3 | Main analysis—The following variables were generated from each coding approach (i.e., sigma level and rupture resolution operationalization) of the control charts: number of sessions with unresolved ruptures and number of session with ruptures resolved in 1, 2, or 3 sessions. This resulted in a total of 24 variables. To determine the coding method with the strongest relationship to outcome, the number of ruptures and repairs identified by each of the six coding methods were correlated with the outcome composite of each dyad. All variables were found to be significantly nonnormal using Shapiro–Wilk tests ($p < .001$). As such, Spearman rank sum correlation coefficients were computed in order to account for the nonnormality of the data. Finally, in order to confidently interpret linear effects, we tested quadratic effects in all models found to be significant.

3 | RESULTS

3.1 | Psychotherapy outcome

A repeated-measures multivariate analysis of variance (MANOVA) revealed significant improvement across the six outcome measures, assessed at intake and termination ($F = 36.21$, partial $\eta^2 = 0.77$, $p < .001$), which did not differ by treatment group ($F = 1.57$, $p = .17$).

3.2 | Average alliance

Across all 30 sessions, average alliance across the sample of 73 patients was found to be 5.80 (SD = .76). Average early alliance (sessions 1–5; $M = 5.47$, SD = .85), average mid-alliance (sessions 6–24; $M = 5.81$, SD = .79), and average late alliance (sessions 25–30; $M = 6.05$, SD = .81) seemed to suggest an overall linear increase in alliance scores across treatment. The association between early alliance and the psychotherapy outcome composite (where the lower the outcome score, the more improvement was seen) was found to be $r = -.162$ ($p = .17$). The association between overall average alliance and psychotherapy outcome was $r = -.206$ ($p = .08$).

3.3 | Frequency of coded ruptures

Potential ruptures were identified frequently across dyads, regardless of the coding method applied. Table 2 shows the frequency of coded ruptures across dyads ($n = 73$) for all six coding methods. Average rupture frequency across all dyads ranged from 1.08 per dyad (SD = .81) in the most stringent coding method to 2.88 ruptures per dyad (SD = 1.09) in the most liberal coding method. The operationalization of rupture repair as a return to the lower control limit generated a significantly greater number of ruptures across sigma levels (median = 1.67, min. = 0.00, max. = 3.33) than rupture repair coded as a return to the mean WAI score across 30 sessions (median = 2.33, min. = 0.33, max. = 4.33) (Wilcoxon signed ranks test, $z = -5.88$, $p < .001$), consistent with its less stringent criteria. Rupture frequency was not found to differ by treatment group in five of the six coding methods ($p > .05$), with the exception of sigma level 2, rupture repair operationalized as a return to the lower control limit ($p = .037$). Further, across rupture resolution operationalizations, a Friedman test revealed a significant linear relationship, such that as sigma level increased, coded rupture frequencies decreased (sigma 2 median = 1.0, min. = 0.00, max. = 3.50; sigma 1.5 median = 2.00, min. = 0.00, max. = 4.00; sigma 1 median = 2.50, min. = 0.00, max. = 5.00; $\chi^2 = 72.03$, $p < .01$). Given that significant differences in rupture frequency were identified across all six rupture coding methods, all coding approaches were included in the main analyses.

3.4 | Main analysis

Table 3 summarizes the Spearman correlation results of rupture resolution operationalized as a return to the mean score. With this operationalization, two sigma levels were found to have a significant small to moderate association to psychotherapy outcome, as measured by the outcome composite variable: sigma level 2 ($\rho = -.24$, $p < .05$) and sigma level 1.5 ($\rho = -.26$, $p < .05$). Of note, both were seen only in the ruptures resolved within 1 session category.

Table 3 also summarizes the Spearman correlation results of rupture resolution coded as a return to the established lower control limit of the control chart. With this coding approach, sigma level 2 and sigma level 1.5 both resulted in significant correlations between rupture frequency and psychotherapy outcome. Here, when control charts were generated based on sigma level 2, moderate associations were found in the ruptures resolved within one session, and ruptures resolved within two sessions categories ($\rho = -.32$, $p < .01$, and $\rho = -.31$, $p < .01$, respectively). With control charts based on the 1.5 sigma level, significant small to moderate associations were found in the ruptures resolved within two sessions, and ruptures resolved within three sessions ($\rho = -.27$, $p < .05$, $\rho = -.24$, $p < .05$, respectively). R package “cocor” was employed to test for statistically significant differences between correlation coefficients (Diedenhofen & Musch, 2015). No correlation coefficients were found to be significantly different (all p values > .06) despite magnitude differences as high as .2.

To confirm the linear nature of these relationships, we tested the significance of quadratic relationships between the previously identified significant variables and the outcome composite variable. However, no quadratic terms were found to be significant (ps > .19).

4 | DISCUSSION

This study explores the implementation of control charts in the indirect identification of potential rupture repair events with significant implications for psychotherapy outcome, a method initially proposed using a single-case study by Eubanks-Carter et al. (2012). The results of this exploratory study, the first to our knowledge using a larger sample, provide preliminary evidence for the utility of control charts for both clinicians and researchers, alike. The second aim of this study was to further define the parameters with which to construct control charts so as to best predict outcome. Although there were no statistically significant differences between control limit parameters, multiple differences were small-to-moderate in magnitude, and thus we provide tentative interpretations of patterns of results that warrant further exploration. A more conservative approach (i.e., at least 1.5 to 2 standard deviations below the mean alliance score across treatment [sigma level 2]) to coding rupture may identify potential rupture events that are more strongly associated with psychotherapy outcome than less conservative methods of coding. In comparison, fluctuations in alliance score one standard deviation below the mean (sigma level 1) were not found to be significantly correlated with psychotherapy outcome, which may be explained by the occurrence of within-session rupture–repair episodes not captured by our methodology, or that this criterion may have been overly liberal and identified process that was not associated with outcome, likely reflective of Type I error. In sum, while there was no clear difference between rupture events identified with sigma level 2 and sigma level 1.5 charts, the different magnitude of effects of these thresholds indicates they may out-perform sigma level 1 in the identification of rupture process associated with outcome.

Overall, the results of these analyses support the utility of control charts for the identification of potential rupture repair events, with more conservative definitions of ruptures (sigma level 1.5 and 2). However, given that some evidence was found to support the use of 1.5 sigma level control charts from rupture identification, we suggest that readers consider their research question when applying control charts: For questions aimed at further understanding rupture process which may or may not have implications for outcome (i.e., for training purposes or process coding), sigma level 1.5 control charts may yield additional sessions that may be worth further assessment containing some rupture-like process.

In terms of the operationalizing rupture resolution, two definitions were tested, one more stringent than the other. Comparisons between these formulae were inconclusive as no statistically significant differences identified. As such, rupture repair defined as surpassing the lower control limit value (e.g., two standard deviations) and rupture repair defined as returning to average alliance were both found to be sufficient for identifying probable rupture–repair events that were correlated with psychotherapy outcome. Replication with a larger sample size is necessary to determine which approach would be best for operationalizing rupture resolution. Given the pattern of results (e.g., the average magnitude of effects with the lower control limit formula exceeded that of the average alliance formulation), we would hypothesize that the lower control limit operationalization may prove to be more strongly associated with psychotherapy outcome. If this is the case, a repair in the alliance may be related to outcome even if alliance ratings at the repair session did not return to the level of average alliance across treatment.

The idiographic nature of the control chart method of indirect rupture–repair identification allows for a personalized formula for each unique therapist–patient dyad, in comparison to nomothetic approaches, which have historically dominated inquiry in rupture process (Lipner, 2020). A recent single-case study that compared the results of several different rupture repair identification methods found a wide range in the frequency of rupture sessions identified. However, the control chart method was found to be consistent with other idiographic rupture identification methods, each of which identified fewer potential rupture sessions than those methods that employed a nomothetic approach (Lipner et al., 2018). Additionally, control charts provide users with a visual representation of the evolving therapeutic alliance over time. This visual makes control charts an approachable and user-friendly interface for trainees and clinicians who may be interested in identifying ruptures but find the required statistics challenging or time consuming. As such, control charts may be useful tools for researchers and clinicians alike in determining the occurrence of rupture repair process.

In the application of control chart methods for potential rupture repair identification, we emphasize the following considerations. First, control charts are based on postsession self-report ratings of the therapeutic alliance. As a result, any occurrence of rupture process that is subsequently resolved within the same session would go undetected by this method. Further, control charts are generated based on measures of central tendency, namely, mean and standard deviation that summarize the entire time series. As such, the charts are susceptible to shifts based on sudden increases and decreases in alliance scores. Additionally, the well-documented ceiling effect and range restriction in measures of therapeutic alliance may make rupture process more difficult to detect (Tryon et al., 2007). To address these limitations, future studies could consider statistical analyses more recently developed to detect trend turning: methods that function to identify when a shift between two nearby time points occurs within a larger time series in a non-randomly distributed manner (Zuo et al., 2019) as well as methods that include presession measures of alliance to better assess within-session change (Zlotnick et al., 2020). Finally, it is important to acknowledge this study used correlational analyses to assess the relationship between identified rupture repair episodes and psychotherapy outcome, and therefore does not necessarily indicate a causal relationship between rupture or rupture repair and psychotherapy outcome. Further, given that multiple correlations were computed, it is important to acknowledge the potential for Type I error. However, this is mitigated by the intrinsic dependency amongst many of the variables of interest (e.g., multiple methods of coding rupture resolution derived from the same session-level WAI scores).

Previous efforts have compared control chart methods with various other methodological approaches to rupture identification, including exponentially weighted moving average (EWMA) charts and the widely implemented Stiles et al. (2004) method. In this analysis, EWMA charts were not able to identify transient shifts but were better suited to identify general or global trends in alliance over time. As such, they are, perhaps, best suited for tracking gradual changes over time, rather than abrupt or sudden drops in alliance, depending on how the parameters of the chart are set. The control charts explored here function in the inverse: they are sensitive to abrupt or transient shifts but not as sensitive to the small, sustained shifts. Methods such as the one outlined in Stiles et al. (2004),

which was based on shape-of-change parameters, can also detect sudden drops; however, the criteria used to define these sudden drops are somewhat arbitrarily constructed. Each of these methods carry with them pros and cons for their implementation. Here, it would be important to consider the research question at hand: Is the researcher interested in gradual trends or abrupt drops? Is the researcher assessing clinical significance, or to identify sessions which merit further analysis (i.e., with observer-based measures to code rupture process, such as 3RS)? (Refer to Lipner, 2020 for a comprehensive review and comparison of previously implemented methods for rupture repair identification.)

Although this study has provided clarity on the use of control charts for the identification of potential rupture events that are significantly associated with psychotherapy outcome, questions remain. First, it would be helpful, and perhaps necessary, for studies to assess the psychotherapy process in rupture sessions identified by control charts to further validate their utility. One such study, a single case study that applied the 3RS (Eubanks et al., 2015) compared coded rupture process in a control chart-identified rupture session with an “average alliance” session (Lipner et al., 2019). Results indicated that the control chart-identified rupture session had significantly higher rupture marker frequency and an overall more negative impact on the alliance, according to 3RS ratings. Though this study provides additional support for the use of control charts in rupture identification, future studies with larger samples will be more demonstrative of their utility at the aggregate level. Additionally, future case studies focusing on defining the clinical meaning of rupture–repair episodes identified by control charts would serve to illustrate the clinical utility of the control chart tool. Second, the current study does not provide any insight into the minimum number of sessions needed in order to generate a control chart that is associated with outcome given that the present data only assessed control charts post hoc (i.e., posttreatment completion). This question has implications for the use of control charts as a routine outcome monitoring, or live clinical tool for clinicians to track the alliance as it fluctuates over time, and to ensure ruptures do not go unaddressed (Lambert et al., 2004, 2003; Reese et al., 2009; Slade et al., 2008).

It is also important to note that these results may not generalize to all patient populations and treatment modalities. For example, this sample excluded cluster B personality disorders, a patient population with which building and maintaining a strong therapeutic alliance has been shown to be particularly difficult (Boritz et al., 2018; Lingiardi et al., 2005; Muran et al., 1994; Tufekcioglu et al., 2013). In the same vein, though this study included participants in a cognitive behavioural and brief relational protocol, it is possible that other therapeutic approaches, such as dialectical behavioural therapy, may entail unique parameters for identifying rupture process. As such, future studies would be necessary to conclude whether these findings apply to broader patient populations and treatment approaches. Further, this study included only participants who were classified as having completed the psychotherapy protocol (25 out of 30 sessions). As such, these results may not generalize to those cases that prematurely terminate treatment. This is an important consideration, as dropout cases may reflect unresolved ruptures. Future efforts to assess how this method may be applied to identify these cases are imperative for the understanding of negative psychotherapy process. Finally, the results of this study are based on patient ratings of the therapeutic alliance. As such, it cannot be assumed that the same criteria should be applied to therapist ratings,

particularly given the therapist's overall tendency to rate the alliance lower than the patient (Tryon et al., 2007). Therefore, a replication study is required to determine the criteria needed to identify rupture repair process for therapist ratings.

Although future studies would be helpful in further elucidating the implementation of control charts, this study provides initial evidence for the empirical utility of this statistical tool in the identification of complex and meaningful psychotherapy process. The format of these charts makes them approachable for users at all levels to follow the natural ebbs and flows of the therapeutic relationship. The ability of control charts to identify potential rupture process associated with psychotherapy outcome reliably provides both researchers and clinicians alike a method to better understand the complex processes thought to be change points in psychotherapy with significant implications for therapeutic outcome.

FUNDING INFORMATION

This work was supported by the National Institute of Mental Health (Grant MH071768).

DATA AVAILABILITY STATEMENT

Data may be made available upon direct request to the corresponding author.

REFERENCES

Agnew-Davies R, Stiles WB, Hardy GE, Barkham M, & Shapiro DA (1998). Alliance structure assessed by the Agnew Relationship Measure (ARM). British Journal of Clinical Psychology, 37, 155–172. 10.1111/j.2044-8260.1998.tb01291.x [PubMed: 9631204]

Alfi-Yogev T, Hasson-Ohayon I, Lazarus G, Ziv-Beiman S, & Atzil-Slonim D (2020). When to disclose and to whom? Examining within-and between-client moderators of therapist self-disclosure-outcome associations in psychodynamic psychotherapy. Psychotherapy Research, 1–11. [PubMed: 30092152]

Battle C, Imber S, Hoehn-Saric R, Stone AR, Nash E, & Frank J (1966). Target complaints as criteria of improvement. American Journal of Psychotherapy, 20, 184–192. 10.1176/appi.psychotherapy.1966.20.1.184 [PubMed: 5906081]

Beck AT, Davis DD, & Freeman A (Eds.) (1990). Cognitive therapy of personality disorders. New York: Guilford Press.

Bordin ES (1979). The generalizability of the psychoanalytic concept of the working alliance. Psychotherapy: Theory, Research, and Practice, 16 (3), 252–260. 10.1037/h0085885

Boritz T, Barnhart R, Eubanks CF, & McMain S (2018). Alliance rupture and resolution in dialectical behavioral therapy for borderline personality disorder. Journal of Personality Disorders, 32, 115–128. 10.1521/pedi.2018.32.supp.115 [PubMed: 29388901]

Callahan CD, & Barisa MT (2005). Statistical process control and rehabilitation outcome: The single-subject design. Rehabilitation Psychology, 50(1), 24–33. 10.1037/0090-5550.50.1.24

Colli A, & Lingiardi V (2009). The Collaborative Interactions Scale: A new transcript-based method for the assessment of therapeutic alliance ruptures and resolutions in psychotherapy. Psychotherapy Research, 19 (6), 718–734. 10.1080/10503300903121098 [PubMed: 19742381]

Deane FP, Spicer J, & Todd DM (1997). Validity of a simplified target complaints measure. Assessment, 4(2), 119–130. 10.1177/107319119700400202

Derogatis LR (1983). The SCL-90-R: Administration, scoring, and procedures manual II. Baltimore: Clinical Psychometric Research.

Diedenhofen B, & Musch J (2015). cocor: A comprehensive solution for the statistical comparison of correlations. PLoS One, 10(4), e0121945. [PubMed: 25835001]

Endicott J, Spitzer RL, Fleiss JL, & Cohen J (1976). The Global Assessment Scale: A procedure for measuring overall severity of psychiatric disturbance. Archives of General Psychiatry, 33(6), 766–771. 10.1001/archpsyc.1976.01770060086012 [PubMed: 938196]

Ericsson KA, Krampe RT, & Tesch-Römer C (1993). The role of deliberate practice in the acquisition of expert performance. Psychological Review, 11(3), 363–406.

Eubanks CF, Muran JC, & Safran JD (2015). Rupture Resolution Rating System (3RS): Manual. New York, NY: Mount Sinai Beth Israel.

Eubanks CF, Muran JC, & Safran JD (2018). Alliance rupture repair: A meta-analysis. Psychotherapy, 55(4), 508–519. 10.1037/pst0000185 [PubMed: 30335462]

Eubanks-Carter C, Gorman BS, & Muran JC (2012). Quantitative naturalistic methods for detecting change points in psychotherapy research: An illustration with alliance ruptures. Psychotherapy Research, 22(6), 621–637. 10.1080/10503307.2012.693772 [PubMed: 22708548]

Eubanks-Carter C, Muran JC, & Safran JD (2010). Alliance ruptures and resolution. In Muran JC & Barber JP (Eds.), The therapeutic alliance: An evidence-based guide to practice (pp. 74–96). New York, NY: Guilford.

Falkenström F, Finkel S, Sandell R, Rubel JA, & Holmqvist R (2017). Dynamic models of individual change in psychotherapy process research. Journal of Consulting and Clinical Psychology, 85(6), 537–549. 10.1037/ccp0000203 [PubMed: 28394170]

First MB, Spitzer RL, Gibbon M, & Williams JBW (1995). Structured clinical interview for DSM-IV. New York: New York Psychiatric Institute.

Flückiger C, Del Re AC, Wampold BE, & Horvath AO (2018). The alliance in adult psychotherapy: A meta-analytic synthesis. Psychotherapy, 55(4), 316–340. 10.1037/pst0000172 [PubMed: 29792475]

Flückiger C, Del Re AC, Wlodasch D, Horvath AO, Solomonov N, & Wampold BE (2020). Assessing the alliance-outcome association adjusted for patient characteristics and treatment processes: A meta-analytic summary of direct comparisons. Journal of Counseling Psychology, 67, 706–711. 10.1037/cou0000424 [PubMed: 32212755]

Hill CE, Nutt-Williams E, Heaton KJ, Thompson BJ, & Rhodes RH (1996). Therapist retrospective recall impasses in long-term psychotherapy: A qualitative analysis. Journal of Counseling Psychology, 43(2), 207–217. 10.1037/0022-0167.43.2.207

Hill CE, Thompson BJ, Cogar MC, & Denman DW (1993). Beneath the surface of long-term therapy: Therapist and client report of their own and each other's covert processes. Journal of Counseling Psychology, 40(3), 278–287. 10.1037/0022-0167.40.3.278

Horowitz LM, Rosenberg SE, Baer BA, Ureño G, & Villaseñor VS (1988). Inventory of interpersonal problems: Psychometric properties and clinical applications. Journal of Consulting and Clinical Psychology, 56(6), 885–892. 10.1037/0022-006X.56.6.885 [PubMed: 3204198]

Horvath AO, Del Re AC, Flückiger C, & Symonds D (2011). Alliance in individual psychotherapy. Psychotherapy, 48(1), 9–16. 10.1037/a0022186 [PubMed: 21401269]

Horvath AO, & Symonds BD (1991). Relation between working alliance and outcome in psychotherapy: A meta-analysis. Journal of Counseling Psychology, 38(2), 139–149.

Lambert MJ, Harmon C, Slade K, Whipple JL, & Hawkins EJ (2004). Providing feedback to psychotherapists on their patients' progress: Clinical results and practice suggestions. Journal of Clinical Psychology, 61(2), 165–174.

Lambert MJ, Whipple JL, Hawkins EJ, Vermeersch DA, Nielsen SL, & Smart DW (2003). Is it time for clinicians to routinely track patient outcome? A meta-analysis. Clinical Psychology: Science and Practice, 10(3), 288–301.

Larsson MH, Falkenström F, Andersson G, & Holmqvist R (2016). Alliance ruptures and repairs in psychotherapy in primary care. Psychotherapy Research, 1–14. [PubMed: 25751649]

Lingiardi V, Filippucci L, & Baiocco R (2005). Therapeutic alliance evaluation in personality disorders psychotherapy. Psychotherapy Research, 15(1–2), 45–53. 10.1080/10503300512331327047

Lipner LM (2020). Needle in the haystack: The identification of ruptures in the therapeutic alliance with implications for psychotherapy outcome (Publication No. 27541568) [doctoral dissertation, Adelphi University]. ProQuest Dissertations Publishing.

Lipner LM, Liu D, Eubanks CF, & Muran JC (2018, June). Discrepant quantitative methods of alliance rupture identification: A single case study. Paper presented at the 34th annual meeting of the Society for the Exploration of Psychotherapy Integration, New York, New York.

Lipner LM, Muran JC, Eubanks C (2019, June). The validation of control chart-identified rupture sessions using 3RS. Paper presented at the 35th annual meeting of the Society for the Exploration of Psychotherapy Integration, Lisbon, Portugal.

Luborsky L, Crits-Christoph P, Mintz J, & Auerbach A (1988). Who will benefit from psychotherapy?. New York: Basic Books.

McLaughlin AA, Keller SM, Feeny NC, Youngstrom EA, &Zoellner LA (2014). Patterns of therapeutic alliance: Rupture–repair episodes in prolonged exposure for posttraumatic stress disorder. Journal of Consulting and Clinical Psychology, 82(1), 112–121. 10.1037/a0034696 [PubMed: 24188510]

Muran JC, Safran JD, Eubanks CF, & Gorman BS (2018). The effect of alliance-focused training on a cognitive-behavioral therapy for personality disorders. Journal for Consulting and Clinical Psychology, 86(4), 384–397. 10.1037/ccp0000284 [PubMed: 29648858]

Muran JC, Safran JD, Samstag LW, & Winston A (2005). Evaluating an alliance-focused treatment for personality disorders. Psychotherapy: Theory, Research, Practice, Training, 42(4), 532–545.

Muran JC, Segal ZV, Samstag LW, & Crawford CE (1994). Patient pretreatment interpersonal problems and therapeutic alliance in short-term cognitive therapy. Journal of Consulting and Clinical Psychology, 62 (1), 185–190. 10.1037/0022-006X.62.1.185 [PubMed: 8034822]

Nguyen TD, Atkinson CC, & Stenger BL (1983). Assessment of patient satisfaction: Development and refinement of a service evaluation questionnaire. Evaluation and Program Planning, 6, 229314.

Pande PS, Neuman RP, & Cavanagh RR (2000). The six sigma way. New York: McGraw-Hill.

Pfadt A, & Wheeler DJ (1995). Using statistical process control to make data-based clinical decisions. Journal of Applied Behavioral Analysis, 28, 349–370.

Reese RJ, Norsworthy LA, & Rowlands SR (2009). Does a continuous feedback system improve psychotherapy outcome? Psychotherapy: Theory, Research, Practice, Training, 46(4), 418–431.

Regan AM, & Hill CE (1992). Investigation of what clients and counselors do not say in brief therapy. Journal of Counseling Psychology, 39, 168–174. 10.1037/0022-0167.39.2.168

Rennie DL (1994). Clients' deference in psychotherapy. Journal of Counseling Psychology, 41(4), 427–437. 10.1037/0022-0167.41.4.427

Rhodes RH, Hill CE, Thompson BJ, & Elliott R (1994). Client retrospective recall of resolved and unresolved misunderstanding events. Journal of Counseling Psychology, 41(4), 473–483. 10.1037/0022-0167.41.4.473

Rosen CS, Drescher KD, Moos RD, Finney JW, Murphy RT, & Gusman F (2000). Six and ten item indices of psychological distress based on the Symptom Checklist-90. Assessment, 7, 103111.

Safran JD, & Muran JC (2000). Resolving therapeutic alliance ruptures: Diversity and integration. Journal of Clinical Psychology, 56(2), 233–243. 10.1002/(SICI)1097-4679(200002)56:2<233::AID-JCLP9>3.0.CO;2-3 [PubMed: 10718606]

Safran JD, Muran JC, Samstag LW, & Stevens C (2002). Repairing alliance ruptures. In Norcross JC (Ed.), Psychotherapy relationships that work (pp. 235–254). New York, NY: Oxford University Press.

Slade K, Lambert MJ, Harmon SC, Smart DW, & Bailey R (2008). Improving psychotherapy outcome: The use of immediate electronic feedback and revised clinical support tools. Clinical Psychology & Psychotherapy, 15(5), 287–303. 10.1002/cpp.594 [PubMed: 19115449]

Stevens CL, Muran JC, Safran JD, Gorman BS, & Arnold W (2007). Levels and patterns of the therapeutic alliance in brief psychotherapy. American Journal of Psychotherapy, 61(2), 109. [PubMed: 17760317]

Stiles WB, Glick MJ, Osatuke K, Hardy GE, Shapiro DA, Agnew-Davies R, Rees A, & Barkham M (2004). Patterns of alliance development and the rupture–repair hypothesis: Are productive relationships U-shaped or V-shaped? Journal of Counseling Psychology, 51(1), 81–92. 10.1037/0022-0167.51.1.81

Strauss JL, Hayes AM, Johson SL, Newman CF, Brown GK, Barber JP, Laurenceau JP, & Beck AT (2006). Early alliance, alliance ruptures, and symptom change in a nonrandomized trial

of cognitive therapy for avoidant and obsessive-compulsive personality disorders. Journal of Consulting and Clinical Psychology, 74(2), 337–345. 10.1037/0022-006X.74.2.337 [PubMed: 16649878]

Strupp HH, Horowitz LM, & Lambert MJ (1997). Measuring patient changes in mood, anxiety, and personality disorders: Toward a core battery. American Psychological Association. 10.1037/10232-000

Tracey TJ, & Kokotovic AM (1989). Factor structure of the Working Alliance Inventory. Psychological Assessment: A Journal of Consulting and Clinical Psychology, 1(3), 207–210. 10.1037/1040-3590.1.3.207

Tryon GS, Blackwell SC, & Hammel EF (2007). A meta-analytic examination of client-therapist perspectives of the working alliance. Psychotherapy Research, 17(6), 629–642. 10.1080/10503300701320611

Tufekcioglu S, Muran JC, Safran JD, & Winston A (2013). Personality disorder and early therapeutic alliance in two time-limited psychotherapies. Psychotherapy Research, 23(6), 646–657. 10.1080/10503307.2013.843803 [PubMed: 24156505]

Turner AE, & Muran JC (1992). Cognitive-behavioral therapy for personality disorders: A treatment manual. San Diego, CA: Social & Behavioral Documents.

Wang LP, & Maxwell SE (2015). On disaggregating between-person and within-person effects with longitudinal data using multilevel models. Psychological Methods, 20, 63–83. 10.1037/met0000030 [PubMed: 25822206]

Wheeler DJ, & Chambers DS (1992). Understanding statistical process control. SPC Press.

Zanarini MC, Skodol AE, Bender DS, Dolan RT, Sanislow C, Schaefer E, Morey LS, Grilo CM, Shea MT, McGlashan TH, & Gunderson JG (2000). The collaborative longitudinal personality disorders study: Reliability of axis I and II diagnoses. Journal of Personality Disorders, 14(4), 291–299. [PubMed: 11213787]

Zlotnick E, Strauss AY, Ellis P, Abargil M, Tishby O, & Huppert JD (2020). Reevaluating ruptures and repairs in alliance: Between- and within-session processes in cognitive-behavioral therapy and short-term psychodynamic psychotherapy. Journal of Consulting and Clinical Psychology, 88(9), 859–869. 10.1037/ccp0000598 [PubMed: 32672994]

Zuo B, Li J, Sun C, & Zhou X (2019). A new statistical method for detecting trend turning. Theoretical and Applied Climatology, 138, 201–213. 10.1007/s00704-019-02817-9

Key Practitioner Message

- The therapeutic alliance has been consistently found to be one of the most robust predictors of psychotherapy outcome.
- Ruptures, or weakenings, in the alliance can lead to treatment failure if unrepaired.
- Creating methods to identify these critical moments provide the opportunity to better understand their process and impact.

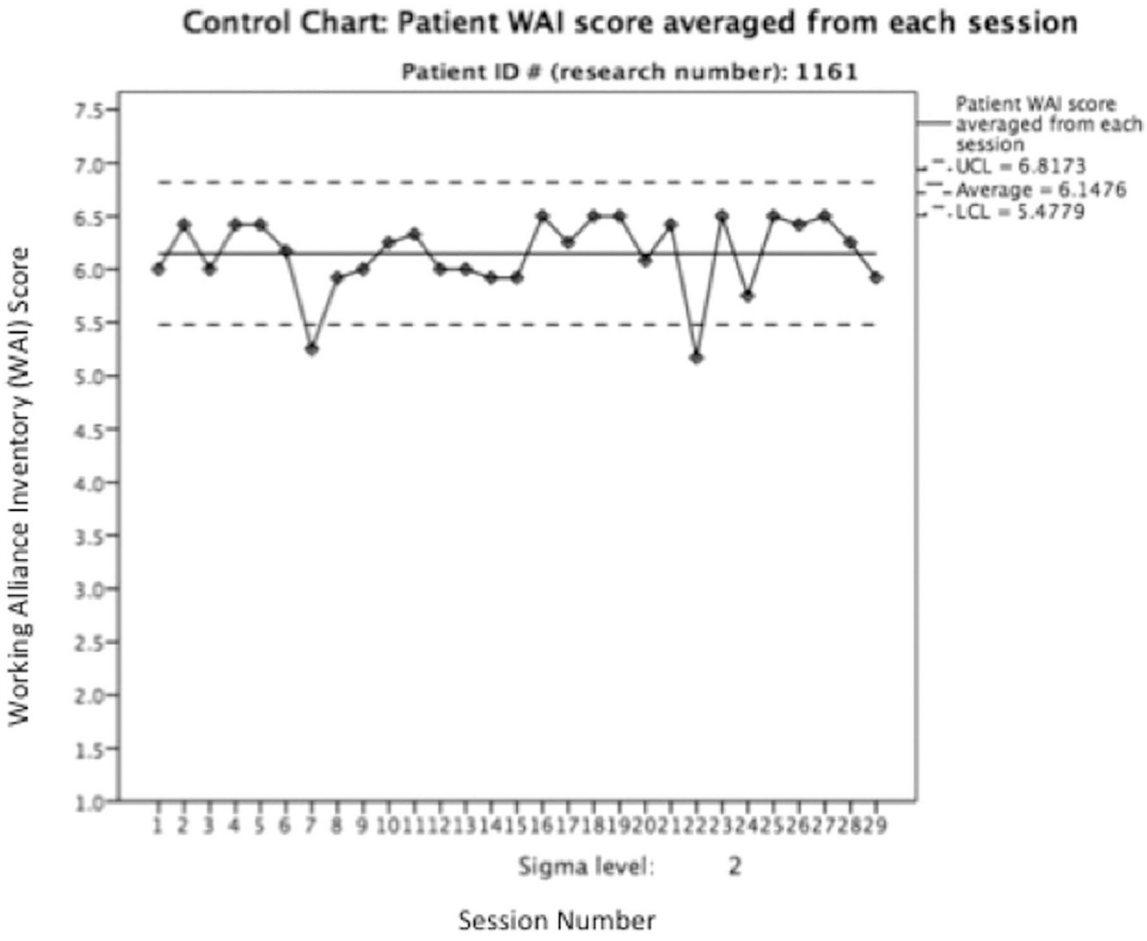

FIGURE 1.

Example of control chart of patient-rated Working Alliance Inventory (WAI) scores at sigma level 2. Solid line represents the average WAI score across 30 sessions, while the dotted lines represent the upper and lower control limits. A rupture was coded when the session average WAI score fell below the lower control limit. Rupture–repair was coded in two methods: when the session average WAI score surpassed the lower control limit or when the session average WAI score surpassed the average WAI score across the 30 sessions

TABLE 1

Patient demographic and diagnostic characteristics (n = 73)

	M	**SD**
Age		
	41.89	13.86
	n	**%**
Gender		
Female	46	63
Male	27	37
Race		
Asian or Pacific Islander	3	4.1
Black, not of Hispanic origin	2	2.7
Hispanic	5	6.8
White, not of Hispanic origin	61	83.6
Primary diagnosis		
Axis I		
Adjustment disorders	5	4.6
Anxiety disorders	22	20.4
Body dysmorphic disorder	1	0.9
Mood disorders	55	51.0
No diagnosis or deferred	8	7.4
Unknown	8	7.4
Axis II		
Avoidant personality disorder	14	13
Dependent personality disorder	1	0.9
Depressive personality disorder	8	7.4
Negativistic personality disorder	1	0.9
Obsessive compulsive personality disorder	10	9.3
Paranoid personality disorder	2	1.9
Personality disorder NOS	21	19.4
No diagnosis or deferred	41	38
Unknown	10	9.3

Note: Patient diagnosis as assessed by Structured Clinical Interview for the DSM-IV (SCID: First et al., 1995) I and II for DSM-IV semistructured diagnostic interviews.

Abbreviations: *M*, mean; SD, standard deviation.

TABLE 2

Frequencies of coded ruptures across dyads

	Resolution to mean	Resolution to LCL
Sigma 2		
Total ruptures	79	110
Unresolved ruptures	19	3
Ruptures resolved in 1 session	41	93
Ruptures resolved in 2 sessions or less	49	102
Ruptures resolved in 3 sessions or less	60	107
Sigma 1.5		
Total ruptures	110	153
Unresolved ruptures	27	5
Ruptures resolved in 1 session	53	115
Ruptures resolved in 2 sessions or less	69	144
Ruptures resolved in 3 sessions or less	83	148
Sigma 1		
Total ruptures	159	210
Unresolved ruptures	37	15
Ruptures resolved in 1 session	82	149
Ruptures resolved in 2 sessions or less	107	184
Ruptures resolved in 3 sessions or less	122	195

Abbreviation: LCL, lower control limit.

TABLE 3

Correlation between rupture repair frequency and psychotherapy outcome

	Sigma level					
	2		1.5		1	
Rupture repair categories	**Rho**	***p***	**Rho**	***p***	**Rho**	***p***
Rupture repair coded as return to mean alliance score						
Total ruptures	−.144	.224	−.083	.132	.132	.267
Unresolved	−.139	.240	−.123	.300	−.039	.744
Resolved in 1 session	−.240	.040 *	−.263	.024 *	−.066	.580
Resolved in 2 sessions	−.196	.096	−.225	.056	−.032	.789
Resolved in 3 sessions	−.078	.510	−.081	.498	.063	.594
Rupture repair coded as return to lower control limit						
Total ruptures	−.194	.101	−.172	.147	.018	.881
Unresolved	.144	.224	.166	.160	.072	.547
Resolved in 1 session	−.322	.005 **	−.210	.074	−.057	.629
Resolved in 2 sessions	−.313	.007 **	−.267	.022 *	−.027	.818
Resolved in 3 sessions	−.225	.056	−.239	.042 *	−.043	.717

Note: Spearman correlations were used for the purposes of this analysis.

* $p < .05$.

** $p < .01$.