Does the Therapeutic Relationship Predict Outcomes of Psychiatric Treatment in Patients with Psychosis? A Systematic Review

Stefan Priebe, Michelle Richardson, Maire Cooney, Oluwatoyin Adedeji, Rosemarie McCabe

Abstract

Background: Numerous studies have shown that the quality of the therapeutic relationship (TR) between the patient and the clinician is an important predictor of the outcome of different forms of psychotherapy. It is less clear whether the TR also predicts outcomes of psychiatric treatment programmes in patients with psychosis (i.e. outside conventional psychotherapy). Methods: We conducted a systematic review and identified 9 primary studies that prospectively tested the association of the TR with 3 outcomes, i.e. hospitalisation, symptom levels and functioning. Because of the heterogeneity of the methods used, a meta-analysis was not feasible. A vote counting method was used to determine the number of statistically significant effects in the hypothesised direction (i.e. that a more positive TR predicts more favourable outcomes). Results: For each outcome, a \( \chi^2 \) analysis showed that the number of statistically significant findings in the hypothesised direction was greater than expected if the null hypothesis of no association were true. However, studies had methodological shortcomings, and the effect sizes of positive associations were rather small. Conclusion: It may be concluded that there is some, but not overwhelming, evidence that the TR predicts outcomes of complex psychiatric treatment programmes in patients with psychosis, and that methodologically more rigorous research is required. Such research should measure the TR at initial stages of treatment and use validated assessment instruments for both TR and outcomes.

Key Words

Therapeutic relationship · Psychosis · Hospitalisation · Symptoms · Functioning

Introduction

The therapeutic relationship (TR) between a patient and a clinician – also referred to as helping, working or therapeutic alliance [1] – is at the centre of the delivery of psychiatric treatment. In surveys, patients consider it to be the most important component of care [2]. Qualitative research suggests that the TR plays a major role for patients with severe mental illness to engage with services [3, 4]. Although there is no universal consensus on how the TR should be defined and measured, it is widely regarded as an important non-specific factor in determining treatment outcome [5]. However, what is the evidence that the TR predicts outcomes of psychiatric treatments in patients with psychosis?
There has been more research on the TR in psychotherapy and psychosomatics, where for several decades, it has been regarded as a central and important concept. Numerous studies reported an association between a more positive TR and more favourable treatment outcomes of psychotherapy. In a meta-analytic review of 79 studies, a significant association between the TR and a composite outcome of psychotherapeutic treatment was found, with an overall small effect size of $r = 0.22$ [6]. There was no significant variation of findings across studies so that the finding can be seen as applicable to different settings in psychotherapy. Although 18 of these studies included patients with severe mental illness, psychotherapeutic settings are substantially different from those of psychiatric treatment commonly provided for patients with psychotic disorders. Psychiatric treatment can include coercive measures, typically uses a range of psychological, social and pharmacological interventions, and is more open-ended and more variable in terms of the frequency, length and aims of meetings than psychotherapy [7].

Various measures have been used to measure the TR in psychiatric settings. Most of them had originally been developed for psychotherapeutic settings, and some instruments were designed ad hoc for psychiatric settings [8]. Recently, a scale specifically designed to assess the TR in community mental health care has been published [9]. Scales often have separate ratings for the clinician and the patient. Their perspectives on the TR may only be weakly to moderately correlated [10]. More positive ratings of the TR have repeatedly been found to be associated with lower symptom levels [11], but these correlations are based on cross-sectional studies and do not constitute evidence that the TR predicts outcomes of subsequent treatment.

In this paper we report the findings of a systematic review of empirical studies that tested the association between the TR and subsequent outcomes of psychiatric treatments for patients with psychosis. The review was guided by the hypothesis that a more positive TR would be associated with better treatment outcomes.

**Methods**

**Searches and Inclusion Criteria**

A 3-stage systematic search was undertaken to locate primary research papers relevant to the review. Initial search terms contained adjectives or derivatives of ‘therapeutic alliance’, ‘treatment outcome’ and ‘psychosis’ that were combined using a series of boolean ‘AND/OR’ operators and wildcards. These combinations were used to search Medline, Psychinfo, Psych-articles and Cochrane databases between 1990 and 2009. Only English-language journals were considered.

Potentially relevant articles were exported into a reference citation manager where titles and abstracts were screened for relevance. At stage 2, studies were included only if (a) patients were treated in psychiatric settings, (b) at least 50% of the sample were diagnosed as having a psychotic disorder (including schizophrenia, schizoaffective disorder and psychoses), (c) the study used a measure of the TR, and (d) the TR was linked to at least 1 measure of clinical improvement or outcome. Moreover, only prospective studies were considered, i.e. studies that used a longitudinal design measuring the TR prior to the assessment of outcome (not cross-sectionally at the same point of time). Finally, at stage 3, we included only outcomes that were assessed in more than 1 study. Where data were missing, authors were contacted. When a study reported associations of TR ratings at several points of time with the same outcome, we included the associations of only 1 of the ratings and selected the earlier rating, making sure that the interval between TR rating and outcome measurement was at least 6 months.

**Data Coding**

The data listed in table 1 were coded from each primary article where present. In order to minimise bias resulting from statistically dependent findings [12], global composite scores were coded wherever available. Ratings of patients and clinicians were treated separately.

**Quality Criteria**

The following criteria and coding were used to assess for each association the quality of the study reporting it: the response rate of those patients who were eligible and/or approached to participate (<30% or not reported = 0; ≥30% = 1); dropout rates between the assessments of the TR and outcome (≥30 = 0; <30 = 1); the sample size (<30 = 0; ≥30 and <100 = 1; ≥100 = 2); the reliability of the instrument used to measure the TR and outcome (no established scale = 0; established scale with internal consistency of <0.70 or non-reported reliability = 1; established scale with internal reliability of >0.70 = 2); whether the association is adjusted for baseline scores of outcomes (no = 0; yes = 1), and whether the association is adjusted for other potential confounders (no = 0; yes = 1). When the outcome was hospitalisation during the follow-up period, the reliability of assessing hospitalisation was coded as 2, as data were obtained from medical records and assumed to be reliable. Adjusting for baseline scores of outcomes was always coded as 0 for hospitalisation, and adjusting for hospitalisations prior to baseline was considered as adjusting for confounders. Scores were summed across each item to create an overall quality score ranging from 0 to 10, with higher scores indicating better study quality. The studies were then allocated to 1 of 3 groups, i.e. low (0–4), medium (5–7) and high quality (8–10), a distinction used in other reviews [13].

**Interrater Reliability**

All articles were coded by 2 independent researchers. An initial agreement rate of 92% across all judgments was obtained and all disagreements were resolved by discussion.
Analytic Strategy

The heterogeneity of methods prevented us from conducting a meta-analysis. A vote counting method was used to establish the number of statistically significant effect size estimates in the hypothesised direction. \(\chi^2\) tests were used to compare hypothesised versus obtained frequencies of positive significant findings, using the 5% probability criterion of making a type I error.

Results

At stage 1 the search strategy yielded a total of 129 papers. After having scanned abstracts and titles using the specified inclusion criteria, 33 papers were identified as relevant and read in detail. The substantial exclusions at this stage were due to a large number of studies that had assessed the TR, but not studied it as a predictor of subsequent outcome. Finally, 9 [14–22] of the 33 relevant papers were found to meet all inclusion criteria and were included in the review. The search process is summarised in figure 1.

The reported studies were conducted in Canada, Germany, Sweden, the UK and the USA. The percentage of patients with psychosis or schizophrenia, respectively, varied between 55 and 86%.

Three outcomes were assessed in 2 or more studies: hospitalisations, symptom levels and measures of functioning. The included studies either (i) measured the TR at baseline and predicted the symptoms/functioning at a later point in time and/or the readmissions or days of hospitalisation between baseline and follow-up assessments [15, 18, 22], or (ii) tested the TR as a predictor of change in symptoms/functioning after having controlled for (a) the same measure at baseline [14, 17, 18, 21] and/or (b) a combination of constructs in a multivariate analysis [17, 18], or (iii) tested the correlation of the TR with computed change scores of symptoms/functioning [16, 19, 20]. Three different types of effect size estimates were reported: correlations (r), standardised betas (\(\beta\)) and F values (F). Table 1 shows all studies and their findings as considered in this review.

In total, 22 associations of the TR with outcomes were reported, i.e. 6 with hospitalisations, 10 with symptoms (5 observer rated, 5 clinician rated) and 6 with a measure of functioning (all clinician rated). Given the relatively small number of studies included, stratification by potential moderating factors was not possible.

Hospitalisations

Three studies assessed how the TR predicted hospitalisations [15, 18, 21] with a total of 6 bivariate associations. Exact outcome measures were readmissions, days spent in hospital or a hospitalisation index reflecting days in full and partial hospitalisation. Across all studies, there were 3 statistically significant associations in the hypothesised direction, i.e. a better TR was associated with fewer hospitalisations. Thus, 50% of the associations obtained statistical significance, which is different from the distribution assumed under the null hypothesis \(\chi^2 (1) = 426.316; p < 0.001\).

Symptom Levels

In 6 studies, a total of 10 associations between the TR and symptom levels as outcomes of subsequent treatment were reported [14, 17, 19, 21, 22]. With respect to the
<table>
<thead>
<tr>
<th>Author(s) and reference No.</th>
<th>Country and setting</th>
<th>Response rates, dropout rates and sample size plus percent with psychosis</th>
<th>TR measure and rater (clinician and/or patient)</th>
<th>Length of time to outcome</th>
<th>Outcome (rater)</th>
<th>Outcome measures</th>
<th>Tested associations and findings (all analyses bivariate unless stated)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Priebe and Gruyters [15]</td>
<td>Germany; community mental health care system</td>
<td>48%; 38%; 72, 69%</td>
<td>Helping Alliance Scale [15]; patient</td>
<td>20 months</td>
<td>Hospitalisation</td>
<td>Hospitalisation Index (days of full hospitalisation multiplied by 3 and days of partial hospitalisation by 2, divided by all days in the observation period)</td>
<td>Correlation (Pearson’s r) between TR scores and Hospitalisation Index; r = –0.20, p &lt; 0.05</td>
</tr>
<tr>
<td>Eklund [19]</td>
<td>Sweden; occupational therapy programme in day care unit</td>
<td>not reported; 20%; 20, 55%</td>
<td>Therapeutic Alliance [29]; clinician and patient</td>
<td>Variable (4–30 months)</td>
<td>Symptoms (patient)</td>
<td>Symptom Check List 90 – Revised Form [30]</td>
<td>Correlations (Spearman’s r) of clinician and patient ratings of TR with symptom change between admission and discharge; n.s. (coefficients not reported)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Symptoms (observer)</td>
<td>Health Sickness Rating Scale [31]</td>
<td>Correlations (Spearman’s r) of clinician and patient ratings of TR with symptom change between admission and discharge; for clinician-rated TR: r = –0.52, p &lt; 0.05; for patient-rated TR: n.s. (coefficients not reported)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Functioning (observer)</td>
<td>Assessment of Occupational Functioning [32]</td>
<td>Correlations (Spearman’s r) of clinician and patient ratings of TR with change in functioning between admission and discharge; n.s. (coefficients not reported)</td>
</tr>
<tr>
<td>Goering et al. [17]</td>
<td>Canada; assertive case management service for homeless patients</td>
<td>12%; 23%; 55, 86%</td>
<td>Working Alliance Inventory [33]; patient</td>
<td>18 months</td>
<td>Symptoms (observer)</td>
<td>Brief Psychiatric Rating Scale [34]</td>
<td>Repeated measures analysis of variance controlling for specific level of functioning at baseline, with symptom levels at baseline, 9 months and 18 months; F = 0.12, n.s.</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Functioning (observer)</td>
<td>Specific Level of Functioning [35]</td>
<td>Repeated measures analysis of variance controlling for symptom levels at baseline, with functioning levels at baseline, 9 months and 18 months; F = 2.60, n.s.</td>
</tr>
<tr>
<td>Klinkenberg et al. [21]</td>
<td>USA; assertive community treatment</td>
<td>not reported; 30%; 74, 66%</td>
<td>15 items patient</td>
<td>15 months</td>
<td>Hospitalisation</td>
<td>Number of hospital days in observation period</td>
<td>Correlation between TR at 2 months and hospitalisation; r_s = 0.14, n.s.</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>12 months</td>
<td>Symptoms (patient)</td>
<td>Global Severity Index of the Brief Symptom Inventory [36]</td>
</tr>
<tr>
<td>Author(s) and reference No.</td>
<td>Country and setting</td>
<td>Response rates, dropout rates and sample size plus percent with psychosis</td>
<td>TR measure and rater (clinician and/or patient)</td>
<td>Length of time to outcome</td>
<td>Outcome (rater)</td>
<td>Outcome measures</td>
<td>Tested associations and findings (all analyses bivariate unless stated)</td>
</tr>
<tr>
<td>---------------------------</td>
<td>---------------------</td>
<td>------------------------------------------------------------------------</td>
<td>-----------------------------------------------</td>
<td>--------------------------</td>
<td>----------------</td>
<td>-----------------</td>
<td>-------------------------------------------------------------</td>
</tr>
<tr>
<td>Svensson and Hansson [16]</td>
<td>Sweden; treatment unit using principles of cognitive therapy</td>
<td>97%; 15%; 28, 68%</td>
<td>Psychotherapy Status Report [37]; clinician; Modified Therapeutic Alliance Scale [29]; patient measured monthly</td>
<td>Variable (M = 62.3 weeks)</td>
<td>Functioning (observer)</td>
<td>Global Assessment of Functioning [38]</td>
<td>Baseline functioning scores were regressed onto functioning scores at discharge to calculate standardised residual change scores, which were subsequently correlated using Spearman’s r with clinician and patient TR ratings; for clinician-rated TR: r = 0.42, p &lt; 0.05; for patient-rated TR: n.s. (correlations not reported)</td>
</tr>
<tr>
<td>Chinman et al. [22]</td>
<td>USA; outreach and intensive case management for homeless patients</td>
<td>not reported; 20%; 2,798, 61%</td>
<td>Therapeutic Alliance [33]; patient</td>
<td>12 months</td>
<td>Symptoms (patient)</td>
<td>Psychotic symptoms in the past month (no standardised scale)</td>
<td>One-way MANCOVA (groups defined TR at baseline and 3 months) including measures of life satisfaction, days homeless, social support, depression, addiction severity index and days worked; both n.s. (F values not reported)</td>
</tr>
<tr>
<td>Calsyn et al. [14]</td>
<td>USA; assertive community treatment with or without substance abuse treatment</td>
<td>not reported; 21%; 98, 68%</td>
<td>Working Alliance Inventory [39]; clinician and patient</td>
<td>18 months</td>
<td>Symptoms (observer)</td>
<td>Brief Psychiatric Rating Scale in revised version</td>
<td>Partial correlations between clinician- and patient-rated TR at 3 months with symptoms at 18 months controlling for baseline symptoms; p&lt;sub&gt;per&lt;/sub&gt; = −0.16, n.s. and −0.21, p &lt; 0.05, for clinician- and patient-rated TR, respectively</td>
</tr>
<tr>
<td>Hopkins and Ramsundar [20]</td>
<td>Canada; treatment for addiction and mental health and housing support services</td>
<td>not reported; not reported; 30, 86%</td>
<td>Short form of Working Alliance Inventory [39]; clinician</td>
<td>1 year</td>
<td>Functioning (observer)</td>
<td>Multnomah Community Ability Scale [41]</td>
<td>Residualised TR scores were formed, using baseline and 11-month scores, and then entered into a multiple regression model predicting functioning, controlling for baseline functioning; β = 0.24, p &lt; 0.05</td>
</tr>
<tr>
<td>Fakhoury et al. [18]</td>
<td>UK; assertive outreach services</td>
<td>not reported; 33%; 133 (newly admitted patients) and 313 (patients &gt;3 months in care), 73%</td>
<td>Helping Alliance Scale [15]; clinician</td>
<td>9 months</td>
<td>Hospitalisation</td>
<td>Rehospitalisation during the study period</td>
<td>In a multiple logistic regression analysis, TR scores were tested for their association with hospitalisation before and after controlling for staff covariates (e.g. working at weekends); for new patients before considering covariates: β = −0.21, n.s.; after considering covariates: β = −0.24, p &lt; 0.05; for established patients: β = 0.04 and −0.03, both n.s.</td>
</tr>
</tbody>
</table>
symptom scales used in the studies, we distinguished between observer-rated [14, 17, 19] and patient-rated measures [19, 21, 22].

The studies reported a total of 5 associations between the TR and observer ratings of symptoms. Four of these were bivariate associations, of which 2 were in the hypothesised direction and statistically significant [14, 19]. The remaining 2 associations were from the same 2 studies and non-significant. In 1 study [17] a non-significant association was reported in a model that also included baseline symptom scores and weeks in permanent residence as predictors. Five associations were reported between the TR and patient-rated symptoms. Only 1 of them was in the hypothesised direction [20] and obtained marginal statistical significance.

In summary, of the 10 associations between the TR and symptom outcomes, 3 (i.e. 30%) obtained statistical significance or marginal statistical significance and were in the hypothesised direction. This is statistically different to a hypothetical sampling distribution under the null hypothesis \( \chi^2 (1) = 131.579; p < 0.001 \).

**Functioning**

Four studies were located [16, 17, 19, 20] assessing associations between the TR and measures of functioning, reporting a total of 6 associations. All included clinician-rated measures of functioning. Four associations (3 coded as bivariate, 1 as multivariate) related to global assessments of functioning [1, 16, 20], 2 of which were statistically significant [16, 20]. One study used clinician and patient ratings of the TR and reported 2 associations with measures of occupational functioning [19]. None of them obtained statistical significance.

In summary, out of 6 reported associations between the TR and functioning outcomes, 2 (i.e. 33%) were in the expected direction and statistically significant, both using global assessments of functioning. This indicates that the sampling distribution is different to that assumed under the null hypothesis \( \chi^2 (1) = 165.053; p < 0.001 \).

**Quality of Studies**

Nine of the 22 associations were coded as based on low-quality studies (1 with hospitalisation, 5 with symptoms, and 3 with functioning as outcomes). The remaining 13 were from medium-quality studies (5 with hospitalisation, 5 with symptoms, and 3 with functioning as outcomes). No study met the criteria for high quality.

All 3 significant associations for hospitalisation were from medium-quality studies. Out of the 5 significant associations for symptom change, 2 were from low-quality studies and 3 from medium-quality studies. The 2 significant associations for functioning originated from a low-quality and a medium-quality study, respectively.

**Discussion**

We reviewed the prospective studies on the association of the TR with outcomes of psychiatric treatment programmes in patients with psychosis and included 9 papers reporting studies from 5 countries. The findings were mixed within and across the studies. Several studies showed that a more positive TR was associated with fewer readmissions to hospital and more favourable changes in symptom levels and functioning measures. Overall, there were more significant correlations in the hypothesised direction than would have been expected if there was no association. However, most tested associations were not significant, and the existing evidence for the predictive value of the TR for treatment outcomes in this patient group is limited.

The review has various limitations. Because of the small number of studies and the heterogeneity of methods, no meta-analysis was conducted. Consequently, artefact variance such as sampling and measurement errors could not be accounted for. We used a vote counting method, which does not provide an estimate of the overall effect size. In some cases, 2 or more associations were extracted from the same study which may have led to a bias in the vote counting procedure [12]. However, we considered more than 1 association from the same study because we treated clinician and patient ratings of the TR as distinct and tested separate outcomes. Further limitations of the review reflect the methodological shortcomings of the included studies, none of which met the defined criteria for a high-quality study. Seven out of the 9 included studies had insufficient power to detect a small effect size, with 3 studies having sample sizes of <30. Consequently, the fact that most tested correlations were not statistically significant may be a result of the usually small sample sizes. Only 2 studies had sample sizes of >100. One neither used standardised outcome measures nor reported bivariate associations and had a negative result. The other one showed a significant association of the TR with readmissions for patients who were newly admitted to assertive outreach care, but not for patients who had already been in care of the teams for more than 3 months. Finally, the vote counting method may have been influenced by publication bias.
The findings are consistent with the assumption that the TR is associated with important clinical outcomes of psychiatric treatment programmes in outpatients with psychosis, and the strength of the association represents a small effect size so that it gets identified in some studies and not in others. Such a conclusion would be in line with evidence in psychotherapy which also shows a small effect size [6]. However, whilst the findings are consistent with this assumption, there is too little research and some of the existing studies are of too poor a methodological quality to provide conclusive evidence for it.

Whilst one can only speculate about the reasons for the relative lack of high-quality studies addressing the role of the TR in psychiatric treatments of patients with psychosis, the review underlines the need for methodologically more rigorous research and points to at least 3 requirements for future research in the area.

First, the TR should be assessed by accurate instruments that have been shown to be valid measures of the TR in psychiatric treatments of patients with psychosis. These instruments should possibly distinguish between different aspects of the TR as different components such as the overall collaboration or emotional responses of the clinician and patient and could have different associations with outcomes. Validated instruments should also be used to assess outcomes.

Second, the assessment of the TR and the baseline measure of the outcome criterion should happen early in treatment. If both assessments are not conducted initially and the observation period for the association of the TR with outcomes begins at a later stage of treatment, outcomes may already have improved as a result of a positive TR. Similarly, a longer time period between the measurement of the TR and the follow-up assessment may facilitate the detection of an effect [23]. Thus, the study design can lead to ceiling effects (or floor effects, respectively) and reduce the chance to detect an association of the TR with outcomes in the future, e.g. in terms of further symptom change. Such effects may be the reason why one study on patients in assertive outreach teams [14] identified an association between the TR and readmissions in newly admitted patients, but not in patients who had already been with the team for more than 3 months.

Third, research should consider mediating factors such as treatment adherence in order to understand how the TR may impact on outcomes.

The assumption that a more positive TR is linked to better treatment outcomes in patients with psychosis can have clinical implications. Clinicians can be trained in communication skills and receive supervision to establish better relationships with some or all of their patients [24]. One may also develop and test interventions compared to an appropriate control [25], influencing patient-clinician communication directly to improve both the TR and outcomes. An example of the latter is the DIALOG intervention [26], which is a computer-mediated method of structuring the communication in a patient-centred and forward-looking manner and has been found to be associated with better treatment outcomes in community mental health care [27, 28]. Many clinicians may intuitively agree with the assumption that the quality of the TR is relevant for outcomes of complex psychiatric treatments in patients with psychosis, and that the overall effect of the TR on outcomes is limited within the complex interplay of all specific pharmacological, psychological and social interventions. There is some research evidence for this assumption, but the existing research has serious shortcomings and the evidence is not overwhelming.

**References**


Therapeutic Relationships in Psychiatric Settings

Psychoter Psychosom 2011;80:70–77