Prediction of Therapeutic Outcome in a Naturalistic Setting Using Pretreatment Psychological Distress Indicators

Kaline Mütze    Michael Witthöft    Anne-Kathrin Bräscher
Abteilung für Klinische Psychologie, Psychotherapie und Experimentelle Psychopathologie,
Johannes-Gutenberg-Universität Mainz, Mainz, Germany

Keywords
Naturalistic setting · Patient-focused research · Therapeutic success

Abstract
Background: Outcome predictions allow to improve psychotherapy and to increase economic benefit. The efficient translation into practice requires simple prediction methods. The present study evaluates the prediction of treatment outcome based on initial distress level. Methods: Routine data of a university psychotherapy outpatient clinic were used (N = 3,145, Mage = 35.8, 67% female). Low versus high distress patients (classified by overall psychological distress, symptomatology, and previous treatment) were compared on total reduction in psychopathology, (early) response, remission, and premature treatment discontinuation using t-tests and logistic regressions. Response and remission were assessed via relative (percentage improvement) and absolute measures (Reliable Change Index; RCI). Results: Distress level was inversely related to percentage improvement (OR = 0.62) and remission (OR = 0.34). It was positively related to total reduction in psychopathology (d = 0.63), RCI response (OR = 2.37), and treatment discontinuation (OR = 2.15). Early response and treatment discontinuation partially mediated the relationship between distress level and treatment outcome. Conclusions: Treatment success tends to be lower when initial distress is high, but this finding appears contingent on the operationalization of treatment outcome. The presented classification approach is easy to implement in practice and may be useful in order to counter an excessive workload in psychotherapy trainees.

Die Prognose psychotherapeutischer Behandlungsergebnisse im naturalistischen Setting durch Belastungsindikatoren zu Behandlungsbeginn

Schlüsselwörter
Naturalistisches Setting · Patientenorientierte Forschung · Therapieerfolg

Zusammenfassung
Introduction

Research has shown that patient characteristics explain the largest amount of variance in psychotherapeutic treatment outcome [Bohart and Tallman, 2010; Bohart and Wade, 2013]. The severity of impairment is the key characteristic that is most closely related to treatment results [Garfield, 1994; Lambert, 2015], such that intense psychological distress is associated with worse therapeutic outcome [Castonguay and Beutler, 2006; Bohart and Wade, 2013]. The prognosis of treatment success is highly relevant for assessing the likely effectiveness and economic benefit of psychotherapy [Boswell et al., 2015; Lambert, 2015], for setting individual treatment plans and goals [Lutz et al., 2019], as well as for planning balanced treatments of patients with heterogeneous distress levels. The latter can be particularly advantageous in an educational context, to counter an excessive workload in psychotherapy trainees. Studies that generate data in naturalistic settings contribute to practice-based evidence in psychotherapy research and complement the randomized controlled design [for an overview, see Barkham et al., 2010; Holmqvist et al., 2015].

A principal interest of practice-oriented psychotherapy research is to identify factors that predict treatment outcome before therapy starts [Riedel et al., 2011; Lewis et al., 2012; Uckelstam et al., 2019]. Different conceptions of success make it difficult, however, to establish universal standards for the evaluation of psychological treatment. The following methods have proven themselves in research: pre-post comparisons as an absolute measure of change, (early) response to treatment, and remission when values are no longer in the pathological range of a symptom scale [for a review, see Hiller et al., 2009; Hiller and Schindler, 2011]. Depending on how they are operationalized, response and remission rates may vary in one sample. For example, the method of percentage improvement considers baseline scores in the calculation and thus individual cut-offs are set that classify patients as responders. By contrast, the Reliable Change Index (RCI) [Jacobson and Truax, 1991] defines a critical difference that must be exceeded, regardless of a patient’s individual baseline score. Thus, for the RCI, absolute change is decisive, opposed to relative change in the percentage improvement method. In general, the measures (early) response and remission are designed to assess quantitative improvement, discriminating pathological and non-pathological scores. An alternative criterion for treatment evaluation is, for example, treatment discontinuation [e.g., Hiller et al., 2009; Cooper and Conklin, 2015; Fernandez et al., 2015].

In addition to demographic variables, there are several patient characteristics that affect therapeutic results, such as personality factors, motivational factors, disease-related, cognitive, and emotional factors [Coleman, 2006; Reimer et al., 2007; Bohart and Wade, 2013]. Therapist characteristics, factors of the therapeutic alliance, and the contextual framework are also related to treatment outcome [Lambert, 2015; Heinonen and Nissen-Lie, 2020]. Treatment success is further affected by patients’ expectations, in that higher expectations at the start of therapy are associated with lower symptom level at termination. This effect is mediated, among other things, by the quality of the therapeutic relationship [Constantino et al., 2011; Vislă et al., 2018].

The present study focuses on disorder-related patient characteristics to assess pretreatment psychological distress. Pretreatment symptom severity has been shown to be positively associated with absolute improvement [Schindler et al., 2013] and negatively associated with remission at treatment termination [Riedel et al., 2011]. The findings suggest that with high baseline distress there is potential for improvement, but the distress is also unlikely to completely disappear during treatment. Furthermore, high baseline distress is associated with greater likelihood of treatment discontinuation [Cinkaya et al., 2011]. Prior studies have shown psychotherapy to be more effective with patients who complete treatment than with those who drop out of treatment prematurely. This is reflected, among other ways, in larger effect sizes for pre to post changes in complete samples compared to intention-to-treat (ITT) samples (including all patients who started treatment) [Hiller et al., 2009]. Early improvement during the first sessions has been demonstrated to be positively related to remission at treatment termination [Tadić et al., 2010; Schindler et al., 2013]. Certain disorder-related patient characteristics are associated with lower therapeutic success: Courtet et al. [2017], for example, found a negative relationship between suicidality and symptom reduction. Suicidal behavior and suicidal ideations are also associated with premature treatment discontinuation [Wnuk et al., 2013]. A personality...
disorder often co-occurs with less therapeutic success and an increased likelihood of treatment discontinuation [Merrill and Strauman, 2004; Swift and Greenberg, 2012]. Previous inpatient treatment is associated with chronic symptom stability, such that persistent symptoms are an indication for inpatient treatment [Koch-Stoecker et al., 2015]. Additionally, Riedel et al. [2011] found a negative association of previous inpatient treatment with response and remission.

The objective of the present study is to determine whether patients with low and high levels of distress in outpatient psychotherapy differ in treatment success, taking into account different outcome measures. Furthermore, the study aims to investigate whether the observed patient characteristics reflect pretreatment distress, which in turn is associated with treatment outcome. Outcome prediction based on pretreatment patient characteristics enables reasonable data-based allocation of treatment resources in naturalistic settings.

We postulate a positive correlation of baseline distress with total reduction in psychopathology and RCI response, since high initial distress is associated with greater absolute change, partly due to regression towards the mean and floor effects [Copay et al., 2007]. In contrast, we expect a negative correlation between baseline distress and response according to percentage improvement and remission. We further expect that early response plays a mediating role in these relationships. In accordance with previous research, a positive correlation between baseline distress and treatment discontinuation (in the following termed dropout) is postulated. Exploratory investigations were conducted to determine whether dropout mediates the relationship between baseline distress and response or remission. In the present study, the following patient characteristics were used to assess baseline distress: overall psychological distress, suicidality, suspected personality disorder, and previous (semi-)inpatient treatment. These characteristics were selected in accordance with research evidence and practical experience.

**Methods**

**Sample**

The sample comprised patients who were treated consecutively between 2003 and 2017 at a university outpatient clinic in Mainz, Germany (Poliklinische Institutsambulanz der Johannes Gutenberg-Universität Mainz; PoIA). Before starting treatment, all patients gave written consent allowing their anonymized data to be used for research purposes. From a total of 5,253 patients, 795 (15%) were excluded due to incomplete data (see online suppl. File 1; see www.karger.com/doi/10.1159/000510272 for all online suppl. material). Next, those were excluded who did not start treatment or dropped out of therapy before the third session. We chose a minimum of three therapy sessions after the diagnostic phase as an inclusion criterion, since a shorter period does not seem sufficient to achieve measurable results [Hiller et al., 2009]. An additional 752 patients (14%) were excluded because of quality-neutral or therapeutic reasons for dropout (e.g., change of residence, therapist recommended alternative interventions) [Hiller et al., 2011]. The remaining 3,145 patients (60%) form the ITT total sample. The ITT total sample (Table 1) (67% female) ranked in age from 16 to 87 (M = 35.80; SD = 12.78). Primary diagnoses were most often in the area of depressive disorders (32.2%), the lowest proportion being schizophrenia and other psychotic disorders (0.4%).

**Table 1. Sample characteristics**

<table>
<thead>
<tr>
<th>Main diagnosis</th>
<th>Total</th>
<th>Low distress (N = 2,231)</th>
<th>High distress (N = 914)</th>
<th>Test statistic</th>
<th>Effect size (95% CI)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Neurodevelopmental disorders</td>
<td>1.4%</td>
<td>1.6%</td>
<td>1.0%</td>
<td>$\chi^2 = 1.96$</td>
<td>OR = 0.61 (0.29–1.26)</td>
</tr>
<tr>
<td>Schizophrenia spectrum and other psychotic disorders</td>
<td>0.4%</td>
<td>0.4%</td>
<td>0.7%</td>
<td>$\chi^2 = 1.21$</td>
<td>OR = 1.84 (0.64–5.31)</td>
</tr>
<tr>
<td>Bipolar disorders</td>
<td>1.0%</td>
<td>0.9%</td>
<td>1.4%</td>
<td>$\chi^2 = 1.63$</td>
<td>OR = 1.60 (0.79–3.22)</td>
</tr>
<tr>
<td>Depressive disorders</td>
<td>32.2%</td>
<td>30.0%</td>
<td>37.5%</td>
<td>$\chi^2 = 16.65^*$</td>
<td>OR = 1.40 (1.19–1.65)</td>
</tr>
<tr>
<td>Anxiety disorders</td>
<td>14.9%</td>
<td>17.2%</td>
<td>9.4%</td>
<td>$\chi^2 = 33.51^*$</td>
<td>OR = 0.50 (0.39–0.64)</td>
</tr>
<tr>
<td>Obsessive-compulsive disorders</td>
<td>3.3%</td>
<td>3.9%</td>
<td>2.1%</td>
<td>$\chi^2 = 6.93$</td>
<td>OR = 0.53 (0.32–0.88)</td>
</tr>
<tr>
<td>Trauma-and stressor-related disorders</td>
<td>3.4%</td>
<td>3.5%</td>
<td>3.2%</td>
<td>$\chi^2 = 0.21$</td>
<td>OR = 0.90 (0.59–1.40)</td>
</tr>
<tr>
<td>Somatic symptom and related disorders</td>
<td>5.2%</td>
<td>6.1%</td>
<td>3.1%</td>
<td>$\chi^2 = 12.98^*$</td>
<td>OR = 0.49 (0.32–0.74)</td>
</tr>
<tr>
<td>Feeding and eating disorders</td>
<td>9.1%</td>
<td>9.4%</td>
<td>8.2%</td>
<td>$\chi^2 = 1.17$</td>
<td>OR = 0.86 (0.65–1.13)</td>
</tr>
<tr>
<td>Substance-related and addictive disorders</td>
<td>0.7%</td>
<td>0.7%</td>
<td>0.5%</td>
<td>$\chi^2 = 0.29$</td>
<td>OR = 0.76 (0.28–2.09)</td>
</tr>
<tr>
<td>Personality disorders</td>
<td>5.8%</td>
<td>2.2%</td>
<td>14.6%</td>
<td>$\chi^2 = 159.04^*$</td>
<td>OR = 7.43 (5.31–10.39)</td>
</tr>
<tr>
<td>BDI-I, M (SD)</td>
<td>19.71 (10.47)</td>
<td>16.02 (8.36)</td>
<td>28.64 (9.65)</td>
<td>$t_{(1,429)} = -33.64^*$</td>
<td>$d = 1.40$</td>
</tr>
<tr>
<td>BDI-II, M (SD)</td>
<td>21.54 (11.66)</td>
<td>17.80 (9.21)</td>
<td>32.15 (11.42)</td>
<td>$t_{(1,638)} = -5.30^*$</td>
<td>$d = 0.63$</td>
</tr>
</tbody>
</table>

BDI-I/II, baseline sum scores on the Beck Depression Inventory; CI, confidence interval; M, mean; SD, standard deviation. *p < 0.05.
**Study Context**

The present study investigated data from the outpatient clinic's routine data collection. The PoIA has been DIN EN ISO 9001 certified since 2005.

Overall psychological distress, suicidality, suspected personality disorder, and previous (semi-)inpatient treatment were assessed at intake (see Instruments section). Patients attended cognitive-behavioral therapy, with on average one individual weekly treatment session. In certain cases, third-wave methods were included in the treatment plan (e.g., dialectical behavioral therapy for treatment of patients with borderline personality disorder). Most of the therapists were in advanced training (90.6% trainees; 9.4% licensed therapists). Therapist trainees were supervised every fourth session by approved supervisors. Overall psychological distress was measured every tenth session from the beginning of treatment as well as five sessions before termination. Patients were classified as having low versus high baseline distress based on the characteristics mentioned above (see Operationalization of the Predictors). Treatment success (total reduction in psychopathology, (early) response, remission, and dropout) was compared between the two groups of patients with different distress level.

**Instruments**

Diagnoses were made during the diagnostic phase using the German version of the Structured Clinical Interview (SKID) [Fydrich et al., 1997; Wittchen et al., 1997] for DSM-IV (Diagnostic and Statistical Manual of Mental Disorders) [American Psychiatric Association, 1994]. Overall psychological distress was assessed using the German version of the Brief Symptom Inventory (BSI) [Franke, 2000]. The Global Severity Index (GSI), which is the average of all items, indicates the intensity of overall psychological distress. In the ITT total sample, the GSI showed high internal consistency of all items, indicates the intensity of overall psychological distress was considered fulfilled with a GSI baseline T-score ≥80. Suicidality was operationalized using the corresponding item in the German version of the Beck Depression Inventory (BDI-I) [Hautzinger et al., 1995] and BDI-II (revised version) [Hautzinger et al., 2009]. Over the study period, the BDI-I was replaced by the revised version, BDI-II. At intake, the internal consistency was α = 0.88 for the BDI-I, and α = 0.91 for the BDI-II. The SKID-II questionnaire was used as a screening for personality disorder. At intake, patients were asked about previous (semi-)inpatient psychiatric or psychotherapeutic treatment.

**Operationalization of Predictors**

A classification into low and high distress patients was performed on the basis of the following four characteristics: Overall psychological distress was considered fulfilled with a GSI baseline T-score ≥80. Suicidality was considered fulfilled if the value of the corresponding item in the BDI exceeded 1. For suspected personality disorder, the varying cut-off in the SCID-II-screening had to be reached in at least five sessions. Logistic regression revealed that in cases of ≥5 fulfilled sections, the odds for a corresponding primary diagnosis were not significantly lower than in the reference category = 11 (all) sections fulfilled. Previous (semi-)inpatient treatment was considered fulfilled in cases of psychiatric or psychotherapeutic pretreatment.

In order to be classified as having low baseline distress, at most one of these characteristic could be fulfilled. This was the case for 2,231 patients in the ITT total sample. In order to be classified as having high baseline distress, at least two characteristics were required to be fulfilled. A total of 914 patients met this criterion. The group classification based on these characteristics was compared with an alternative classification via median split in baseline GSI scores.

**Operationalization of Treatment Success**

Treatment success was measured by total reduction in psychopathology, early response, response, remission, and dropout. Except for the last one, GSI scores at baseline and at the last possible measurement occasion were used (last-observation-carried-forward). Pre to post differences in GSI scores represent the total reduction in psychopathology. Response was operationalized with two competing measures: For percentage improvement, change in distress was relativized to the respective baseline score. To be considered a responder, patients had to have at least 50% improvement within the pathological range of the scale and an additional 25% within the full range. The second criterion was added to counteract the case that for baseline scores close to the cut-off, even minor, clinically insignificant improvements would be considered responders. Percentage improvement was defined in accordance with psychopharmacological research [Hiller and Schindler, 2011]. The cut-off between functional and dysfunctional GSI scores was defined by clinical significance, according to Jacobson and Truax [1991]:

\[ RCI = \frac{X_2 - X_1}{S_{diff}} \]

with

\[ S_{diff} = \sqrt{2(S_x)^2} \]

in which

\[ S_x = S \sqrt{1 - r_n} \]

with \( S = \) standard deviation of the reference population = 0.72 and \( r_n = \) reliability of the measuring instrument = 0.90 [Franke, 2000]. According to the formula, RCI response was achieved for pre-post differences >0.63. Early response was calculated using percentage improvement. There had to be an improvement of at least 25% within the total range at treatment session ten [Schlagert and Hiller, 2015]. As for the operationalization of response, remission was defined using both, percentage improvement and the RCI. In both cases, \( c = \) was used as the threshold and response was required as an additional criterion to prevent clinically insignificant improvements with baseline scores near the cut-off from being considered as remission [Hiller and Schindler, 2011]. Treatment was considered a dropout if it ended for quality-associated reasons (e.g., disregards of agreements). Considering not only improvement but also deterioration, we calculated the proportion of patients whose overall psychological distress increased significantly during treatment (GSI_{pre-post} < -0.63).

---

1. Successful completion of Practical Training I (according to § 2 Psych-ThG-ApV) and at least 280 lessons of theoretical knowledge.
2. Corresponding to high overall psychological distress. The cut-off is based on a sample of outpatient psychiatry patients with an average T-score of 77 [Franke, 2000].
Predicting Therapy Outcome by Pretreatment Psychological Distress

Data Analysis

All analyses were conducted using SPSS (IBM SPSS, version 22.0). Significance level were set at $\alpha = 0.05$. To prevent alpha-error accumulation, Bonferroni correction was applied. Cohen's $d$ was used to calculate effect sizes for mean differences and the odds ratio (OR) of the logistic regression were used to measure the strength of the correlation between categorial variables. To test for differences in total reduction in psychopathology, we ran $t$ tests and AN(C)OVAs. To test for mediation, the variables’ (co)variances, unstandardized coefficients and standard errors from logistic regression analysis were used for the Sobel test [Sobel, 1982; Baron and Kenny, 1986]. By definition of the outcome, it is not possible to achieve (early) response and remission in cases of non-pathological baseline distress. Therefore, patients within the functional range of the GSI at baseline were excluded from corresponding analyses (see online suppl. File 1).

Results

Total Sample

In the ITT total sample, the average overall psychological distress on the GSI at baseline was 1.17 (SD = 0.68), which was significantly reduced to 0.66 at termination (SD = 0.59; $t_{(3,144)} = 46.63; p < 0.01; d = 0.82$). In the ITT sample with pathological baseline scores, 56.7% ($N = 1,361$) responded early to treatment. At treatment termination, 70.2% ($N = 1,742$) had achieved responder status (48.8% or $N = 1,210$ using the RCI), and 46.5% ($N = 1,155$) could be classified as remitted (30.7% or $N = 761$ using the RCI). A total of 10.7% ($N = 337$) dropped out of treatment for quality-related reasons. For 2.4% of patients ($N = 75$), overall psychological distress increased significantly during treatment.

Group Comparisons

Overall psychological distress was higher in high distress patients, both at baseline ($M = 1.83; SD = 0.60$) and termination ($M = 1.03; SD = 0.72$; $t_{(913)} = 33.38; p < 0.01; d = 1.10$), than in patients with low baseline distress ($M_{pre} = 0.90; SD_{pre} = 0.51; M_{post} = 0.50; SD_{post} = 0.45$; $t_{(2,230)} = 35.66; p < 0.01; d = 0.75$). Total reduction in psychological distress in the high distress group was significantly larger than in the low distress group ($t_{(1,321)} = −15.31; p < 0.01; d = 0.63$). An analysis of covariance showed that this effect was due to different average baseline scores between the groups. Controlling for baseline, the effect was reversed ($F_{(1,3,142)} = 36.36; p < 0.01$; Fig. 1). Patients in the high distress group achieved RCI response significantly more often (OR = 2.37) and dropped out of treatment more frequently (OR = 2.15) than patients in the low distress group. There was no significant group difference in early response, after Bonferroni correction (OR = 0.83). Patients in the high distress group responded significantly less often, using percentage improvement (OR = 0.62). Compared to the low distress group, the odds of remission using percentage improvement were less than half in the high distress group (OR = 0.34). For RCI remission, no significant group differences were found after Bonferroni correction (OR = 0.82; for group comparisons see Table 2 and online suppl. File 2). Using the alternative group classification via median split on the GSI, total reduction in psychological distress within the high distress group was significantly greater than in the low distress group.

Fig. 1. Average total reduction in distress on the Global Severity Index for low and high distress patients. Error bars indicate 95% confidence intervals for mean scores.
(t_{(25,561)} = –31.67; p < 0.01; d = 1.13). Patients in the high distress group significantly more often showed early response (OR = 1.33), RCI response (OR = 6.76), RCI remission (OR = 2.11), and dropout (OR = 1.86). Low distress patients achieved remission significantly more often according to percentage improvement (OR = 0.34). The groups did not differ significantly with regard to response using percentage improvement (OR = 1.03; for group comparisons via median split, see online suppl. File 3).

For patients suffering from depressive disorders, treatment success was somewhat greater, and group differences were smaller compared to the total sample (see online suppl. File 4)\(^3\).

**Mediating Effects**

The results of the Sobel test showed that early response partially mediates both the effect of baseline distress level (low versus high distress group) on response (χ\(^2\) = –15.31; p < 0.01; d = 0.63). Patients in the high distress group significantly more often showed early response compared to the total sample (see online suppl. File 4). As expected, high baseline distress was negatively associated with relative improvement, and remission at treatment termination. It was positively associated with total reduction in psychological distress, RCI response, and dropout. We were also able to confirm that early response and dropout partially mediate the relationship of distress level with response and remission. The finding that less success within the high distress group is partially mediated by dropout leads to the assumption that continuing treatment might have led to better treatment results. Using the alternative group classification via median in GSI, overall treatment success within the high distress group was greater compared to the original classification, in which about one third of the patients were classified as high distress. While the alternative classification highlighted group differences more clearly in absolute measures (total reduction in psychological distress, RCI response, and RCI remission), the original classification via patient characteristics discriminated more precisely in terms of dropout and response using percentage improvement.

The naturalistic setting and patient sample with heterogeneous clinical and demographic characteristics have high external and ecological validity. Since a wide range of mental disorders in adults of all ages are treated at the

\(^3\) The sub-analysis was conducted to improve the generalizability of the results, since depressive patients make up a large proportion of the clientele in university outpatient clinics.

### Table 2. Comparison of treatment outcomes between low and high distress patients

<table>
<thead>
<tr>
<th></th>
<th>Total (N = 3,145)</th>
<th>Low distress (N = 2,231)</th>
<th>High distress (N = 914)</th>
<th>Test statistic</th>
<th>Effect size (95% CI)</th>
</tr>
</thead>
</table>
| Total reduction in psychological distress | M = 0.51, SD = 0.62 | M = 0.40, SD = 0.53 | M = 0.80, SD = 0.73 | \( t_{(1,321)} = –15.31^{**} \) | \( d = 0.63 \)
| Early response           | 56.7%            | 58.4%                    | 53.8%                   | \( \chi^2 = 4.83 \) | OR = 0.83 (0.70–0.98) |
| Response                 | 70.2%            | 73.9%                    | 63.7%                   | \( \chi^2 = 28.77^{**} \) | OR = 0.62 (0.52–0.74) |
| RCI response             | 48.8%            | 41.0%                    | 62.3%                   | \( \chi^2 = 103.38^{**} \) | OR = 2.37 (2.00–2.80) |
| Remission                | 46.5%            | 56.0%                    | 30.0%                   | \( \chi^2 = 156.06^{**} \) | OR = 0.34 (0.28–0.40) |
| RCI remission            | 30.7%            | 32.2%                    | 28.0%                   | \( \chi^2 = 4.82 \) | OR = 0.82 (0.68–0.98) |
| Dropout                  | 10.7%            | 8.4%                     | 16.4%                   | \( \chi^2 = 43.69^{**} \) | OR = 2.15 (1.71–2.70) |

Degrees of freedom of the \( \chi^2 \) tests = 1. CI, confidence interval; M, mean; RCI, Reliable Change Index; SD, standard deviation. ** p < 0.01.
Predicting Therapy Outcome by Pretreatment Psychological Distress

Limitations
As a generic instrument, the GSI does not allow for disorder-specific symptom assessment. However, the use of such instruments enable the comparison of patients with heterogeneous diagnoses. A further limitation is that (early) response and remission could only be measured in patients with pathological baseline scores. In addition, the use of the same measurement methods can cause systematic error variance and overestimated correlations. Furthermore, the operationalization of suspected personality disorder requires critical examination. In order to predict outcome as early as possible, this patient characteristic is based on information at intake, before the diagnosis phase. Thus, the percentage of patients suspected of having a personality disorder (33.2%) is larger than those who turn out to have a corresponding diagnosis (5.8%).

It should be further noted that the dichotomous classification into low versus high distress patients is associated with a loss of information. Heterogeneity within the high distress group was considered insofar as treatment outcomes were compared between different combinations of patient characteristics. Since these characteristics have different prognostic validity, and treatment success varies by combination, qualitative classification should be considered as an alternative to quantitative classification. It should also be noted that the present study does not include therapist variables, as for example the level of experience. Since such variables are related to therapeutic outcome, they should be taken into account when analyzing treatment data [Johns et al., 2019; Heinonen and Nissen-Lie, 2020]. Even though longer professional experience is not necessarily associated with greater treatment success [e.g., Goldberg et al., 2016], unbalanced allocation of high distress patients between trainees and licensed therapists may have influenced the results. Furthermore, symptomatic change may be influenced by time or life events, and is thus not causally attributable to treatment outside randomized controlled studies.

Implications for Practice and Research
Predictions of treatment success facilitate indication and allow for tailored treatment planning. Furthermore, a balanced allocation of patients with heterogeneous distress levels can prevent overburden of psychotherapist trainees. For example, therapists at the PoIA start training with two low distress patients and do not treat two high distress cases consecutively. A differentiation between low versus high distress patients based on the selected characteristics proves reasonable and is easy to implement, since all the relevant information are available at intake.

The present results should not be interpreted such that high distress patients are unlikely to experience treatment success. The large effect sizes of pre to post improvement within the high distress group highlight the effectiveness of psychotherapy for these patients. Mechanisms operating between distress level and treatment outcome, such as the therapeutic alliance and motivational factors, must be further investigated [Coleman, 2006; Reimer et al., 2007].

Although the RCI is easy to compute, there are some disadvantages of using this measure. The index gets smaller as the reliability of the measuring instrument increases or as the variance in the norm sample decreases. In both cases, even small, clinically insignificant changes are rated as reliable. In addition, such test parameters differ between samples, which makes comparison with other studies difficult [Zahra and Hedge, 2010]. The RCI has also been criticized for its failure to differentiate between patients with heterogeneous baseline values. While in the present study, high distress patients achieved RCI response more frequently but post scores in the functional range less frequently, low distress patients showed the opposite. Here it becomes clear that for different baseline scores, the measure is more or less liberal. Regarding percentage improvement, it should be noted that even for approved measurement instruments, there are no consistent thresholds that rate change as significant, making comparisons between studies equally difficult. However, this method is advantageous in samples with heterogeneous distress levels since baseline scores are taken into account. For future research, reasonable measures of treatment success should be selected in accordance with the sample population and research question. For the RCI, consistent test parameters should be used. For percentage improvement, thresholds should be defined in accordance with previous research findings [see Hiller et al., 2011]. For outcome in terms of absolute change (e.g., pre to post difference or RCI), prediction based on baseline values is suitable. Further information is required to predict treatment success as defined by relative measures or alternative criteria such as dropout. The presented classification based on patient characteristics provides a prognostically valid approach.

In future research, it should be investigated whether the present findings are reflected in disorder-specific symptoms. Not surprisingly, the group of high distress patients was particularly characterized by the comparatively more frequent occurrence of personality disorders, which highlights the impact of specific disorders on treatment outcome. Considering multiple measurement points over the course of treatment, a quadratic trend (deceleration and acceleration) can be added to the linear trend. For longitudinal data, patient-specific growth can...
be evaluated using multilevel analysis (for an overview, see Lutz et al. [1999], Keller [2003]). This method also allows the inclusion of therapist variables. In order to examine long-term treatment success, the present design could be expanded by a follow-up survey. In addition to symptom change and dropout, it would be interesting to assess the effects of mental disorders on quality of living, social networks and on the working environment. Further, pretreatment information could be expanded by chronicity, taking into account previous episodes and the onset of symptoms.

The anonymized data that support the findings of this study are available on request from the corresponding author: k.muette@uni-mainz.de.

Conclusion

Pretreatment psychological distress, based on the presented patient characteristics, provides a valid forecast of treatment outcome. The results indicate that treatment success tends to be lower when initial distress is high, but this finding appears contingent on the operationalization of treatment outcome. For samples with heterogeneous baseline values, relative measures prove advantageous compared to absolute measures of change. The presented classification of low versus high distress patients is easy to implement and proves useful in naturalistic settings, especially in the educational context, for the allocation of resources, treatment planning, and in order to counter an excessive workload in psychotherapy trainees.

Acknowledgment

The authors would like to thank Prof. Dr. Wolfgang Lutz for his support. We also thank the reviewers for the constructive suggestions in the revision of the manuscript.

Statement of Ethics

The present retrospective study analyzed routine data from a university psychotherapy outpatient clinic in Mainz, Germany. All procedures were performed in accordance with the 1964 Helsinki Declaration and its later amendments. Written, informed consent allowing anonymized data to be used for research purposes was obtained from all participants.

Conflict of Interest Statement

The authors report no conflict of interest.

Funding Sources

K. Mütze is supported by a PhD scholarship from the Johannes Gutenberg University of Mainz, Germany.

Author Contributions

K. Mütze, M. Witthöft, and A.-K. Bräscher conceptualized the study. K. Mütze analyzed and interpreted the data. All authors discussed the results. The manuscript was drafted by K. Mütze and revised by M. Witthöft and A.-K. Bräscher.

References


Predicting Therapy Outcome by Pretreatment Psychological Distress


