Treating Treatment-Resistant Patients with Panic Disorder and Agoraphobia Using Psychotherapy: A Randomized Controlled Switching Trial

Methodological Details for Supplemental Online Material

Method

Sample Characteristics

For pharmacology, previous state-of-the-art treatment was defined as the minimum dosage and length as recommended by guidelines for psychopharmacology used internationally and in Germany [1, 2]. When it was unclear if a pharmacological course met guidelines, a psychiatrist who is an expert in the treatment of anxiety reviewed the case (A.S.). The definition of adequate psychological treatment (e.g., exposure in situ, interoceptive exposure, cognitive restructuring, etc) was met by n = 38 (88.4%) patients and the definition of adequate pharmacological treatment was met by n = 14 (32.6%) patients.

Design

This was a randomized, wait-list controlled clinical trial conducted in Germany. Following screening, patients were invited to an initial appointment and completed informed consent procedures. Patients who passed the screening were given a diagnostic appointment. Patients who met all inclusion criteria entered the study and were randomized to either immediate treatment or wait-list delayed treatment. Assessments occurred immediately prior to treatment (baseline), during treatment for process measures (not reported in this paper), immediately following treatment (post treatment), and six months follow-up (FU-6). Treatment consisted of eight sessions conducted twice weekly over four weeks.

Patient enrollment occurred between June 2010 and June 2012, with the final follow-up assessment in December 2012. Patients were blinded to the hypotheses. Patients were recruited from two primary sources. First, patients who took part in our previous study on exposure-based cognitive behavioral therapy [3, 4] and completed the 24-month follow-up [5] were screened for non-response. In these cases, patients were asked if they were still struggling to the point that they need therapy. Only those patients who answered positively were invited for a diagnostic assessment. Second, all patients who presented to the university outpatient center were screened for previous treatment.

In total, 43 patients entered the study. Patients were randomized to either immediate treatment (n = 33) or a four-week wait-list (n = 10). Patients from the WL were offered treatment immediately following the four-week wait period (delayed treatment; n = 8). A total of 51 cases were included in the analyses (See also Flowchart). All patients were treated at the university outpatient clinic of the Technische Universität Dresden.

Randomization

Patients were randomly allocated to immediate treatment or WL with a 3:1 ratio. A statistician not associated with patient care generated the randomization list. To ensure blinding throughout the study, more numbers were drawn than necessary so that treatment condition of final patients remained unpredictable. This procedure resulted in slightly uneven n’s relative to the allocation ratio. Furthermore, a two-step procedure was utilized such that one staff member allocated included patients to a randomization number and a separate staff member decoded this number from a list that matched the randomization code to a condition.

Intervention

A manual of Acceptance and Commitment Therapy (ACT) for anxiety disorders [6] was adapted for this trial [7]. This manual has already been successfully employed in a large randomized clinical trial comparing ACT with CBT [8]. The brief treatment consisted of eight sessions administered twice weekly over four weeks. Sessions were between 90 and 120 minutes. ACT is a behavioral treatment whose aim is the promotion of psychological flexibility that consists of six processes: acceptance, present moment awareness, defusion, self-as-context (observer perspective), values clarification, and committed action. Patients worked towards becoming more present with and accepting of anxiety and other uncomfortable emotions and experiences. This stance was adopted so that they could more willingly engage in important aspects of their life, irrespective of the presence of uncomfortable emotions and thoughts.

Although all six core processes are present to a greater or lesser degree in most ACT sessions [9], we emphasized various components in each session. The order of emphasized processes was selected with the goal of helping the patients get unstuck by developing the skills necessary to interact differently with their thoughts and emotions while cultivating the reason to do so (values). Session 1 began with brief psychoeducation and then used experiential exercises that focused on acceptance, creative hopelessness, valued actions, and present moment awareness. A primary goal of the first
session was to explore how attempts to control or reduce anxiety and other emotions “succeeded” and the costs (i.e., reduction in valued activities and well-being) associated with such efforts. The discussions and exercises oriented the entire therapy towards what they wanted their life to be about (i.e., values) as opposed to what they had been concentrating on (i.e., controlling or reducing anxiety and other negative feelings). Session 2 continued to examine the workability of trying to control thoughts and feelings, practiced present moment awareness, and cultivated cognitive defusion, or the process of experiencing emotions and language (e.g., thoughts, self-talk, etc.) as part of the broader, ongoing stream of present experience as opposed to getting stuck responding to its literal meaning. Session 3 intensified experiences of mindfulness, acceptance, willingness to experience feelings without needless defense and intensified the examination of the qualities and actions of the patient’s values. Session 4 concentrated on helping the patient gain distance and perspective to their thoughts and emotions by discriminating the experience of observing their language and emotion from the psychological place of safety from which they can view these internal experiences for what they are (self-as-context & defusion) versus what they claim to be (cognitive fusion) and continued to practice present moment awareness and values. Beginning in session 5 participants chose important areas of their lives that were negatively impacted by their emotional struggles and began their commitment to willingly and openly engage with their emotions and feelings in the service of their values (i.e., willingness exercises) and continued to integrate concepts of defusion, self-as-context, values, and commitment. Willingness exercises may require contact with uncomfortable stimuli and as such have a formal similarity to exposure. Willingness exercises differ in function, however, in that habituation or extinction is not the goal but rather psychological flexibility for the benefit of their values (see [10] for a detailed discussion). Sessions 6-7 included continued willingness exercises and worked towards helping the patient to recognize their own patterns of behavior and to implement defusion, acceptance, values and commitment when necessary to live life in they way they choose as opposed to letting their anxiety choose. The final session continued this concentration as needed for each patient, and discussed how to continue moving forward, including how they might act in the presence of unwanted future thoughts and emotions.

The advantage of this approach when working with treatment-resistant patients is that the long-standing struggle to eliminate symptoms is reframed as part of the problem. By so doing, psychological flexibility is enhanced and new ways of dealing with difficult life situations become possible. Interestingly, when patients began to focus on and make choices based on their values, they simultaneously reported significant reductions in symptoms. This seeming paradox is predicted by ACT’s theory of language. That is, prior to treatment symptoms are literally interpreted as a barrier to living life as they wish; following treatment they learn that symptoms and the life they value can co-exist.

Therapists

Therapists were graduate students in clinical psychology. None of the therapists had any prior experience treating patients with ACT. Their general therapy experience consisted of treating several patients using traditional cognitive behavioral therapy under supervision (mean = 10.9 patients, SD = 11.3.) Therapists were trained via a three-day intensive training, readings, and self-study. Prior to seeing patients in the study, therapists were required to pass a competency test that consisted of video recording of crucial elements of the manual. Therapists further received weekly supervision from the first two authors. On average, therapists treated 5.9 (range 3–10) patients within this study.

Treatment Integrity

All treatment sessions were videotaped. We wanted to examine treatment integrity (i.e., adherence, therapist competency, and differentiation) for at least 15% of all conducted sessions. We randomly selected 60 sessions from the total of all 328 possible sessions conducted making sure that at least one session and no more than two sessions from each of the 41 participants was included in the analysis. As five patients dropped out at various stages of treatment, the total number of sessions conducted was 305. This meant that 19.7% (60/305) of all conducted sessions were analyzed for treatment adherence and therapist competency. Treatment adherence and therapist competency ratings were made by one of the developers of the manual (G.E.) not involved in study management or clinical supervision. Ratings were made using the Drexel University ACT/ CT Therapist Adherence and Competence Rating Scale (DUACRS); [11], which has previously been used in similar ACT-CBT comparison studies[8].

Assessors

Assessors were blinded to treatment condition. Prior to beginning the study, assessors completed a three-day training, testing, and subsequent certification of the assessment procedures. The training included extensive practice in the proper administration of the instruments (e.g., role-play, and detailed discussion of common rating scenarios) in which all interviews were schooled to the standard
of no variability for the CGI and only 3 points on the SIGH-A across all exam cases. Most raters (80%) were certified as study assessors on the first attempt, the others passed the exam in form of an additional rating after feedback was given on the first. Regular supervision was conducted to maintain consistent strategies across assessors and address questions.

**Assessment**

Patients completed measurements at baseline, post-treatment, and FU-6. Additional assessments were administered following each session (not reported here). Staff blinded to group conditions administered all assessments.

The primary outcome measures assessed symptoms and functioning traditionally used with this disorder, including overall panic and agoraphobia symptomatology (Panic Agoraphobia Scale [PAS]), global clinical impression and functioning (Clinical Global Impression Scale [CGI]), and agoraphobic avoidance (Mobility Inventory [MI]).

**Panic Agoraphobia Scale** (PAS[12]). The PAS is a self-report questionnaire that measures the severity of panic attacks, avoidance, anticipatory anxiety, disability, and worries about health. Scores on the PAS have been demonstrated to have good reliability and sensitivity to change [12]. Response on the PAS was defined as PAS total ≤8.

**Clinical Global Impression Scale – Severity Subscale** (CGI[13]). CGI is a clinician-rated scale that measures the overall severity of a disorder, with scores that range between 1 (no disorder) and 7 (among the most severely ill patients). The scale queries for information across the facets of panic symptoms, anxiety, anticipatory anxiety, avoidance, and overall functional level before making the global rating. Scores on the CGI are sensitive to change in panic treatment [4, 14]. Response on the CGI was defined as scores of “mild”, “borderline”, or “no” disability.

**Mobility Inventory** (MI[15]). The Mobility Inventory is a self-report questionnaire that measures the degree of agoraphobic avoidance across 27 situations. The mean score of the subscale reporting on situations encountered “alone” (range 1–5) are reported. Scores of the MI are highly reliable and sensitive to change [15].

**Composite International Diagnostic Interview** (CIDI). The standardized computer-administered personal CIDI is administered by expert interviewers and systematically assesses all DSM-IV disorders. The diagnoses derived by the CIDI have been demonstrated to be reliable and valid [16-21].

Additional measures targeted three areas: 1) panic-specific processes, 2) general symptomatology, and 3) ACT specific processes. First, panic-specific processes were included that have been found to mediate other forms of CBT for PD &/or AG and are commonly used in the assessment and treatment of PD &/or AG and include content about panic-specific symptomatology in the items. These included: fear related to bodily sensations (Bodily Sensations Questionnaire; BSQ[22]), Agoraphobic Cognitions Questionnaire (ACQ[22]), and the Anxiety Sensitivity Index (ASI[23]). Second, commonly used measures of more general anxiety and depression symptoms were included: Hamilton Anxiety Rating Scale (SIGH-A[24]), depression (Beck Depression Inventory; BDI-II[25]), and anxiety (Beck Anxiety Inventory; BAI[26]). Finally, measures were included that assessed the specific processes assumed to be active in the ACT treatment. These measures do not contain content that specifically references panic symptomatology. Instead, they directly target the treatment elements involved in ACT. These included, difficulty with emotional regulation (Difficulty with Emotion Regulation; DERS[27]), thought suppression (White Bear Suppression Inventory; WBSI[28]), mindfulness (Kentucky Inventory of Mindfulness Skills; KIMS[29]), and defusion (Believability in Anxious Feelings and Thoughts; BAFT[30]). Additionally, a global, diagnostic-unspecific measure of psychological flexibility was also included (Acceptance and Action Questionnaire [AAQ-II[31]]).

**Statistical Analysis**

**Hypothesis 1 (efficacy).** For all primary and secondary outcomes hypothesis 1 was tested using ANCOVA with baseline outcome values as covariates. In each comparison, the ACT treatment group was compared to the WL. It was hypothesized that patients in the ACT group would report improved values relative to WL.

Given that patients from the WL were offered treatment in the active condition following the waiting period, we examined whether this introduced a systematic bias. This was done by testing whether the treatment response on primary outcome variables of these 8 patients differed from the 33 who entered treatment directly after randomization. No significant differences were found for any of the primary outcome variables (results available upon request). Similarly, we wanted to test whether a systematic bias was introduced by including 10 non-responder patients from our previous trial [4] for whom we had more extensive information than the others. Again, no differences were observed between these groups and there was no indication of differential treatment outcome (results available
upon request). Both of these comparisons were done using the same ANCOVA model as described above.

Statistical analyses for hypothesis 1 were based on completers, i.e. patients with nonmissing outcome values at post-treatment. Additionally, we ran ANCOVAs on all 43 patients. To do so we estimated missing post-treatment values in those five patients who dropped out before post-treatment using multiple imputation (five imputations) and then analyzed the five resulting complete datasets pooling the results, as implemented in the R package mice [32]. The following variables were used to impute missing values: prior values of that variable when available, any primary outcome, any secondary outcome, sex, age, number of sessions in previous treatment, number of comorbid variables, process measures at the last session. As results based on these data were comparable to those obtained from the completer analyses, they are presented below but not in the main text.

**Hypothesis 2 (differential response across disorder specific processes, general symptoms, and ACT specific processes)**

To test whether treatment gains would be stronger in ACT-specific processes (i.e., DERS, WBSI, KIMS, & BAFT) than in panic disorder-specific process measures (i.e., BSQ, ACQ, & ASI) or in more general symptom measures (i.e., SIGH-A, BDI-II, & BAI) we set up a multivariate random intercept model of the combined secondary outcomes. In this model each secondary outcome was first standardized to make values among outcomes comparable and all outcomes were then stacked on top of each other to obtain one outcome, the individual outcomes thereby denoting repeated measures. The statistical model included a random intercept for each subject and a compound symmetry covariance pattern for the repeated measures. Fixed effects were the differential category (ACT-specific, panic-disorder specific, and general symptoms), treatment group (ACT vs WL), the interaction between the two, and the baseline values of the secondary outcomes. Because the interaction between differential category and treatment efficacy indicates whether treatment efficacy varies with differential category, this term was used to test hypothesis 2. To test this interaction we used a likelihood ratio test, which compares the fit of the model as described above with the same model but excluding the interaction term.

**Hypothesis 3 (treatment gain and maintenance).** To test hypothesis 3 we used a linear mixed model[33] with a random intercept, assuming equal covariances among the three time points. Linear mixed models assume dropout patterns that follow the “missing at random” terminology (MAR)[34], i.e. dropout depends on observed covariates and observed values of the outcome variable. Contrasts based on least squared means were then computed to test differences between pre- and post-treatment (hypothesis 2), and between post-treatment and six-months follow-up (hypothesis 3). Analyses were conducted using the R packages lme4 [35] and lsmeans [36].

### Results

<table>
<thead>
<tr>
<th>Measure/Time point</th>
<th>ACT (n = 41)</th>
<th>Waitlist (n = 10)</th>
</tr>
</thead>
<tbody>
<tr>
<td>CGI</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Baseline</td>
<td>4.4 (1.0)</td>
<td>3.6 (1.3)</td>
</tr>
<tr>
<td>Post</td>
<td>3.5 (1.0)</td>
<td>4.1 (0.7)</td>
</tr>
<tr>
<td>Follow-Up 6</td>
<td>3.1 (1.4)</td>
<td></td>
</tr>
<tr>
<td>PAS</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Baseline</td>
<td>19.9 (11.3)</td>
<td>13.8 (8.5)</td>
</tr>
<tr>
<td>Post</td>
<td>14.0 (10.8)</td>
<td>16.3 (10.5)</td>
</tr>
<tr>
<td>Follow-Up 6</td>
<td>11.6 (10.7)</td>
<td></td>
</tr>
<tr>
<td>MI</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Baseline</td>
<td>2.3 (1.0)</td>
<td>2.2 (0.8)</td>
</tr>
<tr>
<td>Post</td>
<td>1.9 (0.8)</td>
<td>2.2 (0.7)</td>
</tr>
<tr>
<td>Follow-Up 6</td>
<td>1.7 (0.9)</td>
<td></td>
</tr>
</tbody>
</table>

**Treatment Efficacy (Hypothesis 1) – Primary Outcomes: Results based on Multiple Imputation**

<table>
<thead>
<tr>
<th>Measure</th>
<th>est</th>
<th>se</th>
<th>t</th>
<th>df</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>CGI</td>
<td>-0.2251390</td>
<td>0.1019969</td>
<td>-2.21</td>
<td>40.97</td>
<td>0.03</td>
</tr>
<tr>
<td>PAS</td>
<td>-0.7880071</td>
<td>0.3991534</td>
<td>-1.97</td>
<td>42.96</td>
<td>0.055</td>
</tr>
<tr>
<td>MI</td>
<td>-0.08544262</td>
<td>0.07016456</td>
<td>-1.22</td>
<td>26.28</td>
<td>0.234</td>
</tr>
</tbody>
</table>
We also examined a general, overarching measure of psychological flexibility (AAQ-II). Whereas the effect size compared to the WL was medium ($d = 0.42$), this difference was not significant, which can be considered a result of the relative change insensitivity of this global measure [31, 37]). Measurement of changes in generalized psychological flexibility has proven somewhat difficult to assess due in part to its dynamic and context-dependent nature. We chose the AAQ-II as a variable because it was the most widely utilized measure of this construct at the time the study began. The content neutral nature of the questions, however, is associated with smaller effects than otherwise identical symptom-specific versions[38]. To guard against this, we also included several specific measures of the ACT processes (e.g., mindfulness, defusion, acceptance) and these variables had the highest observed effect sizes. Furthermore, we included a slight variation of the general questionnaire in the weekly sessions that parenthetically added the words “e.g., anxiety, depression, panic, etc.” to items that referenced emotions. Reductions of patients’ scores on this version were three times that of the general version (available upon request).

References


McGrath KB, Forman EM, Herbert JD: Development and validation of the act/ct adherence and competence rating scale. in preparation


35 Bates D, Maechler M, Bolker B: Lme4: Linear mixed-effects models using s4 classes. R package version 0.999999-2, 2013.