Routledge
Taylor & Francis Group

Psychotherapy Research

ISSN: (Print) (Online) Journal homepage: https://www.tandfonline.com/loi/tpsr20

Countertransference in the treatment of patients with personality disorders: A longitudinal study

R. B. Øvstebø, G. Pedersen, T. Wilberg, J. I. Røssberg, H. S. J. Dahl & E. H. Kvarstein

To cite this article: R. B. Øvstebø, G. Pedersen, T. Wilberg, J. I. Røssberg, H. S. J. Dahl & E. H. Kvarstein (14 Nov 2023): Countertransference in the treatment of patients with personality disorders: A longitudinal study, Psychotherapy Research, DOI: 10.1080/10503307.2023.2279645

To link to this article: https://doi.org/10.1080/10503307.2023.2279645

Psychotherapy Research, 2023
https://doi.org/10.1080/10503307.2023.2279645

RESEARCH ARTICLE

Countertransference in the treatment of patients with personality disorders: A longitudinal study

R. B. ØVSTEBØ[1,2], G. PEDERSEN[3,4], T. WILBERG[1,2], J. I. RØSSBERG[1,2], H. S. J. DAHL[5], & E. H. KVARSTEIN[2,3]

[1]*Section for Treatment and Research, Department of Research and Innovation, Division of Mental Health and Addiction, Oslo University Hospital, Oslo, Norway;* [2]*Institute of Clinical Medicine, University of Oslo, Oslo, Norway;* [3]*Section for Personality Psychiatry and Specialized Treatments, Department for National and Regional Functions, Division of Mental Health and Addiction, Oslo University Hospital, Oslo, Norway;* [4]*Institute of Basic Medical Sciences, University of Oslo, Oslo, Norway &* [5]*Department of Psychology, University of Oslo, Oslo, Norway*

(Received 24 November 2022; revised 30 October 2023; accepted 31 October 2023)

ABSTRACT

Objective This study examines how therapist emotional response/countertransference (CT) develops during treatment for patients with personality disorders (PDs) and how pre-treatment patient factors (severity of personality pathology, PD category, level of symptom distress) predict CT responses. Secondly, we explored associations between patient clinical outcome and CT.

Method A longitudinal, observational study including 1956 patients with personality pathology treated at psychotherapy units within specialist mental health services. Therapists' emotional response was repeatedly assessed by the Feeling Word Checklist—Brief Version (FWC-BV) with three subscales—*Inadequate*, *Confident,* and *Idealized*.

Results Levels of *Inadequate* CT were lowest and stable over time while *Confident* and *Idealized* increased over time. Greater severity of personality pathology and borderline PD predicted higher initial *Inadequate*, lower initial *Confident* and decreasing *Inadequate* over time. Antisocial PD predicted decreasing *Confident*. Number of PD criteria had higher impact on therapist CT than level of symptom distress. Clinical improvement was associated with decreasing *Inadequate*.

Conclusion Therapists reported predominantly *Confident* CT when working with PD patients. More severe personality pathology, and borderline PD, specifically, predicted more negative CT initially, but the negative CT decreased over time. Patients who did not improve were associated with increasing *Inadequate*.

Keywords: countertransference; personality disorders; Feeling Word Checklist

Clinical or methodological significance of this article: The study demonstrates noteworthy CT responses among therapists in treatments addressing patients with personality pathology. The findings support former assumptions that therapists should be aware of and reflect over their CT responses, and this awareness might be especially important at the beginning of therapy. The predominance of positive CT responses may indicate that specialized treatment units contribute to greater confidence among therapists in treatment of patients with PD. Further, the findings suggest that focus on personality aspects might be more important than patients' symptom level when treating patients with PD.

Correspondence concerning this article should be addressed to R. B. Øvstebø, Section for Treatment and Research, Department of Research and Innovation, Division of Mental Health and Addiction, Oslo University Hospital, P.O Box 4959 Nydalen, 0424 Oslo, Norway. Email: r.b.ovstebo@medisin.uio.no

Introduction

Being aware of and managing one's emotional response to the patient is a key competence of therapists, related to patients' change in psychotherapy (Biondi, 2022; Hayes et al., 2018). The focus of countertransference (CT) is theorized to be especially important when working with patients with personality disorders (PD), because of the strong emotional reactions these patients tend to stimulate in clinicians (Gabbard & Wilkinson, 2000; McWilliams, 2011). In clinical literature, it is described how a failure to attend to and contain one's CT responses as a therapist may result in boundary violations, disruption in the therapy, and therapist withdrawal (Carsky, 2021).

There is a growing interest around research on therapists' emotional response in many therapy models (Fauth, 2006; Hayes et al., 2015). However, most empirical studies are correlational in nature, investigating therapists' CT at specific time points during therapy. Few have taken a longitudinal perspective, which could provide insight about CT over time during treatment and the relational dynamics in psychotherapy.

Whether a therapist's emotional responses during treatment reveal something about the therapist's own issues or provide important information about the patient's problems is a discussion that has persisted since Freud introduced the concept of CT in 1910 (Freud et al., 1957). Although the following theoretical discussions have emphasized different aspects of CT, and many disagree on where the emphasis should be placed, most clinicians today agree that it is not a question of "either/or" but rather that CT is a jointly created phenomenon involving contributions from both patient and clinician (Gabbard, 2017; Gabbard, 2020). The therapist's feelings may originate from their own unresolved conflicts, but may also give valuable insight into the patient's problems and relational issues.

Similar to the historic theoretical controversies, CT as a scientific construct has been discussed. The debate concerns what can actually be measured empirically. Firstly, CT reactions are considered partly unconscious processes. Secondly, many emotional reactions can be regarded as countertransferential: affective, cognitive, somatic, and/or behavioral (Gelso & Kline, 2022). There are several different definitions of CT (Hayes et al., 2018). In empirical research, therapist's CT has been measured by self-report or observer ratings. In the present study, we assessed the therapists' self-reported affective CT response to the patient. This is considered one part of the total CT construct, i.e., the feelings which therapists become aware, acknowledge, remember, and are willing to report after a session (Dahl et al., 2012).

Using self-report, several studies have found significant associations between therapists' affective CT response during therapy and a variety of patient factors assessed before treatment, e.g., Colli et al. (2014), Nissen-Lie et al. (2022), and Tanzilli et al. (2016). The studies suggest that clinicians can make diagnostic and therapeutic use of their emotional response to patients. Taking an observer perspective, a recent study also found that therapists' own vulnerabilities and unresolved issues contributed to CT (Tishby & Wiseman, 2022). Thus, studies support CT as a co-created phenomenon. To investigate the complex CT phenomenon, it is therefore necessary to adopt various perspectives in research (Rocco et al., 2021).

Of the correlational studies that exist on PD and CT, several have demonstrated that patients who share the same PD tend to elicit similar, specific emotional reactions in therapists (Betan et al., 2005; Colli et al., 2014; Tanzilli et al., 2016). Borderline PD is, e.g., found to be associated with therapists feeling special or idealized (Betan et al., 2005; Breivik et al., 2020), inadequate, overwhelmed, and devaluated (Colli et al., 2014; Tanzilli et al., 2016) and avoidant PD is, e.g., found to be associated with more positive feelings and parental/protective reactions from the therapist (Colli et al., 2014; Tanzilli et al., 2016).

Other studies have examined different PD clusters and CT (Betan et al., 2005; Rossberg et al., 2007; Thylstrup & Hesse, 2008) and some have examined CT in relation to number of fulfilled PD criteria (Dahl et al., 2012; Nissen-Lie et al., 2022). In sum, several studies reveal that more severe personality pathology in patients is associated with more negative CT (Colli et al., 2022; Dahl et al., 2012; Genova & Gazzillo, 2018; Nissen-Lie et al., 2022).

All the studies that have explored whether therapists' CT responses were influenced by therapists' orientation (Betan et al., 2005; Colli et al., 2014; Tanzilli et al., 2016, 2017), showed that the results were independent of clinicians' theoretical standpoints (Stefana et al., 2020). Further, many studies show coherent empirical findings even though the practitioners involved have different kinds of training and experience (Rossberg et al., 2007; Thylstrup & Hesse, 2008).

Comorbidity of symptom disorders (axis I disorders of DSM-IV) in patients with personality disorders is the norm rather than the exception (Lingiardi et al., 2015; Oldham et al., 1995).

However, few studies have examined the influence of symptom distress on therapists' emotional response. Røssberg and co-workers (2010) found that the patients' symptom distress (assessed with the Symptom Checklist-90-Revised, Derogatis, 1994) was positively associated with therapists' inadequate CT responses and negatively related to confident CT responses. Tanzilli and co-workers (2017) found that patients' symptomatology partially mediated the relationship between patients' personality disorder and therapists' emotional responses, but in general, the impact of symptom severity was less sizable than emotional therapist responses aroused by patients' personality style.

Few studies have examined CT development in relation to therapy outcome, although there is research on how CT influences the processes in psychotherapy. Tishby and Wiseman (2022) found that negative CT patterns were associated with more ruptures and less resolution with patients. Colli and co-workers (2022) found that CT, in complex ways, mediates the relation between patients' defensive functioning and therapist technique. Nissen Lie and co-workers (2022) reported that working explicitly with the therapeutic relationship (transference work) seemed to reduce negative CT with patients having more personality pathology and low motivation. These studies did however not examine how these associations between CT and process relate to outcome. There is a need for more studies to investigate how CT directly or indirectly influence outcome (Hayes et al., 2018).

The present study is an observational study consisting of a uniquely large patient sample reflecting patients as they naturally present within specialist mental health services—covering a broad range of personality problems with differences in severity and PD category. To our knowledge, no studies have explored therapists' CT responses over time in a large patient population with high level of personality pathology. Because of the strong and more troublesome feelings that tend to be elicited when treating patients with personality disorders (Carsky, 2021; Kernberg, 1989), studying these reactions in treatments for this patient group as they are applied within clinical "real-life" settings is especially relevant. The present study has explored therapists' emotional response with the self-report questionnaire; Feeling Word Checklist - Brief Version (FWC-BV: Breivik et al., 2020). The items of the FWC-BV have shown to be reliably differentiated as three distinct dimensions in a previous study (Breivik et al., 2020): *Confident*, *Inadequate and Idealized*. More specifically, the present study included the following research questions:

1. How do therapists' CT feelings develop during treatment of patients with personality pathology?
2. To what extent do different PD diagnoses, severity of personality pathology and symptom distress predict the development of therapist CT feelings over time?
3. Is there an association between patients' clinical outcome and therapists' CT over time?

Material and Methods

Design

The study has a longitudinal, multi-site, naturalistic design

Setting

Data were retrieved from the quality registry of the Norwegian Network for Personality Disorders (The Network) (Pedersen et al., 2022) - a clinical research collaboration between outpatient PD treatment units on a specialist mental health service level. The Network provides a quality assurance system for treatment evaluation, courses and conferences on PD assessment and treatment. In this study period (2010–2016), the register included totally 5600 patients from 20 treatment units. The units combined psychoeducational, group and individual psychotherapy formats. In the study period, treatment approaches were mainly psychodynamic, often combinations with art, body awareness, and cognitive therapies. Specific PD approaches implemented within some units included mentalization-based therapy, dialectical behavioral therapy and schema-focused therapy.

Participants

Patients: The study sample ($N = 1956$) included all patients who had (1) terminated treatment, (2) were evaluated diagnostically with semi-structured interviews before treatment, and (3) had at least one therapist rating of CT. Mean age was 34 years (SD = 10) and 76% were females ($n = 1478$). Overall, 1343 patients (69%) fulfilled criteria for one or more PDs. The most frequent were avoidant PD (31%, $n = 607$), borderline PD (21%, $n = 406$), and PD NOS (16%, $n = 315$) (Table 1). Number of fulfilled PD criteria of the 94 in SCID-II ranged from 0-50, with a mean of 10.1. The prevalence of symptom disorders was 94%, wherein 68% had mood disorders and 52% anxiety disorders. Mean GAF score was 49.7 (SD 6.1). Mean baseline GSI

Table 1. Pre-treatment patient characteristics.

	N	%	Mean	SD
Age			34	10
Female	1478	76		
No work/study last year	815	46		
Not living in a close relationship	1139	59		
Axis I diagnosis				
Major depression	1044	55		
Dysthymia	144	8		
Panic disorder	315	17		
Agoraphobia	83	4		
Social phobia	426	22		
Generalized anxiety disorder	189	10		
Somatization	40	2		
PTSD	236	12		
Eating disorder	187	10		
Total number symptom disorders			1.7	1.0
No diagnosis	113	6		
Axis II diagnosis				
Paranoid	144	7		
Schizoid	8	0		
Schizotypal	2	0		
Antisocial	20	1		
Borderline	406	21		
Histrionic	1	0		
Narcissistic	10	1		
Avoidant	607	31		
Dependent	74	4		
Obsessive-compulsive	104	5		
PD NOS	315	16		
No PD	613	31		
Number of PD criteria			10.1	6.9
More than one personality disorder	267	14		

was 1.52 (SD = 0.66). Treatment duration averaged 19.8 months (SD = 13.5), 67% completed treatment as planned, 14% dropped out, and 19% ended treatment for other reasons, e.g., advised termination, referred to other treatment, moved out of region.

***Therapists*:** The Network units comprised multidisciplinary teams (psychiatrists, psychologists, psychiatric nurses, social workers). Clinical discussion and supervision of treatment processes and CT responses were important elements. Mean number of therapists per unit was 9 (range 4–18), approximately 75% female, mean age 45 years, mean length of clinical experience 17 years, and 73% were trained group therapists. The registry did not couple therapist and patient data.

Measurements

Assessments by clinicians: The Mini International Neuropsychiatric Interview (MINI) (Sheehan et al., 1994) and the Structured Clinical Interview for DSM-IV Personality Disorders (SCID-II) (First, 1997) were used for systematic assessment following the Diagnostic and Statistical Manual of Mental Disorders (*DSM-IV*, 1994). Diagnostic reliability was not investigated. Clinicians were trained in diagnostic interviews and principles of the Longitudinal, Expert, All-Data (LEAD) procedure (Pedersen et al., 2013; Spitzer, 1983). In this study, PD severity is reflected by total number of SCID-II criteria. Diagnostic assessment was performed at baseline, when patients were referred to treatment, before starting therapy.

The Feeling Word Checklist-BV (Breivik et al., 2020), a brief version of the FWC-58 (Rossberg et al., 2003), comprises 10 feeling words aiming to capture therapist CT in a more applicable and less time-consuming questionnaire (Breivik et al., 2020). Different versions have varying in number of items (from 10 to 58) and dissimilar scale formats (Likert scales or dichotomous yes/no versions). For an overview of FWC questionnaire forms, see Lindqvist et al. (2017). In FWC-BV each item is rated on a 0–4 Likert scale. The heading is "During recent conversations with the patient I have felt … ". FWC-BV was assessed repeatedly during therapy, first every 3 months up to 12 months, then every sixth month until 60 months. A total of 4484 questionnaires were filled in over the study period, mainly by the patient's individual therapist.

The Global Assessment of Functioning (GAF: APA, 1994) provides a single score (range 1-100), incorporating symptom severity and social impairment (split GAF version: Pedersen & Karterud, 2012). GAF reliability was acceptable (generalizability coefficient relative decision: .84, absolute decisions: .82) (Pedersen et al., 2007). GAF < 60 indicates moderate/severe impairment (APA, 1994). The study includes GAF as an observer-rated outcome evaluation performed at baseline and at treatment termination.

Patient-rated outcome measures: The Symptom Check List -90-Revised (Derogatis, 1994) includes 90 items rated on a 0–4 Likert scale. Scale reliability (Norwegian version) was satisfactory (reliability estimates of subscales 0.80–0.90) (Pedersen & Karterud, 2010). This study reports the Global Severity Index (GSI), which is the total mean score. Mean GSI > 1 is usual in clinical samples (Derogatis, 1994; Pedersen & Karterud, 2004). The Circumplex of Interpersonal Problems (CIP) (Pedersen, 2002) is a 48 item Norwegian version of the Inventory of Interpersonal Problems—Circumplex version (Alden et al., 1990) rated on a 0–4 Likert scale. The total mean score of the two versions correlates at a level of 0.99. A mean CIP sum score >1 is usual in clinical samples. The study includes patient-rated outcome measures administered at baseline and at treatment termination.

Ethics

Participation required patients' written consent for transfer of anonymous, clinical data to the Network quality register. Collection procedures were approved by local data protection officers for each contributing unit. Data security procedures for the quality register were approved by the data protection officer at the responsible center for the research. Approval from Medical Research and Ethics Committees is not required for anonymous data.

Statistics

Linear Mixed models (LMM) was used for longitudinal data analyses (Fitzmaurice et al., 2012; Singer et al., 2003). Form a hierarchical viewpoint, the data enabled patient-level analyses and included information on change over time (months since starting treatment). It did not enable analyses of variation between therapists but allowed further investigation of variation across units. The dependent variables were: *Inadequate*, *Confident,* and *Idealized*. Data were analyzed in a stepwise fashion, starting with three separate open models for each of the dependent variable. The final model for further analyses included a random intercept and slope (critical values for chi-square statistic: $p < 0.01$) and an unstructured covariance type. In order to restrict linear inflation of scores, we added all the registered termination CT scores to the corresponding last time point in every longitudinal patient course. To give a simpler interpretation of the intercept estimate, the first CT assessment at 3 months was centered at 0.

The following predictors were investigated: (a) Type of PD diagnosis, (b) Total number of SCID-II criteria, and (c) GSI level of symptom distress at the start of treatment. Each PD diagnosis, PD severity, and GSI were investigated as separate predictors added to the models with the three dependent CT variables. The PD categories histrionic, schizotypal and schizoid PD were omitted from the analyses because of low n ($n <$ 10 patients). In a final model, we added all PDs with significantly deviating CT levels. In additional analyses, we analyzed the number of specific SCID-II criteria within each PD category as separate predictors aiming to investigate PD dimensionally.

The equation was $Y_{ij} = \beta_0 + \beta_1 \text{ time}_{ij} + b_{0i} + b_{1i} \text{ time}_{ij} + \varepsilon_{ij}$. Y_{ij} is the dependent variable for all observations, individuals (i), and assessment times (j), β is the fixed effects regression coefficient, b the random effects regression coefficient, and ε indicates residual variation. For each of the predictor (PRED) analyses, the equations were: $Y_{ij} = \beta_0 + \beta_1 \text{ time}_{ij} + \beta_{01} \text{ PRED} + \beta_{11} \text{ PRED time}_{ij} + b_{0i} + b_{1i} \text{ time}_{ij} + \varepsilon_{ij}$

We report LMM estimates for predictor-associated deviation of CT (intercept and slope), log likelihood statistics (AIC) and R Square (Table 2). R Square (Conditional R Square) is a summary statistic that describes the proportion of variance explained by both the fixed and random effects in the model (Nakagawa et al., 2017). The predictors associated with the greatest explained variation (% change from the estimated variation in the initial linear coefficient model), are presented in the text.

We chose GAF as main patient outcome variable. In secondary analyses, we investigated GAF improvement (difference in level from baseline to end of treatment) as a predictor variable added to models with the three CT variables (supplementary file, Table 4). Based on the reliability study of Pedersen et al. (2007) and of the study of the split version of GAF by Pedersen and Karterud (2012), a dichotomous variable was generated to separate patients with an improvement on GAF (symptom and function score separately) at or above 3 points or not. The cutoff of three points improvement is to reduce the possibility of Type-II error due to measurement error. As GAF is a clinician-rated, we also investigated patient-reported GSI and CIP change (difference in levels from baseline to end of treatment).

Missing Assessments

The CT data were unbalanced with different numbers of assessments per patient (range 1–9). Mean number of FWC-BV assessments per patient was 2.3 (SD = 1.5), 42% ($n = 813$) had one CT assessment, 24% ($n = 468$) had two assessments and 34% ($n = 673$) had three or more CT assessments. In LMM, model-based estimations do not require that all subjects have equal number of assessments and the analyses use all available data for each individual trajectory (Norusis, 2008). Missing CT reports in this sample were due to locally occurring, administrative failures of delivery or registration and patients having different treatment duration. To investigate possible systematic bias of missing data, a variable counting the number of assessment points was investigated as a longitudinal predictor in separate models for all dependent variables (Hedeker & Gibbons, 1997). The number of assessments was not associated with significant deviation of baseline or longitudinal change of all three CT dimensions. Second, we compared baseline severity of patient conditions (total PD criteria, GAF and GSI) for patients with one assessment and patients with several assessments. No significant differences between those with only one assessment and those with several assessments were found. In addition, all significant results were reanalyzed in a subsample

Table 2. Pre-treatment patient variables as predictor of CT development.

Dependent variable and parameter	Estimate (SE)	95% Confidence interval	t	df	*p*	AIC	R2
Inadequate							
Intercept	0.51(0.01)	0.482 to 0.540	34.9	1502.68	**<.001**	6158	0.582
Time	−0.001(0.001)	−0.002 to 0.001	−0.7	575.36	.480		
Inadequate × PD Category							
Paranoid	0.15(0.06)	0.035 to 0.256	2.6	1551.87	**.010**	6150	0.582
Time × Paranoid	0.001(0.003)	−0.005 to 0.007	0.2	584.78	.853		
Antisocial	−0.11(0.17)	−0.438 to 0.210	−0.7	2486.87	.489	6160	0.584
Time × Antisocial	0.022(0.014)	−0.005 to 0.048	1.6	3438.60	.112		
Borderline	0.23(0.04)	0.162 to 0.300	6.6	1394.22	**<.001**	6117	0.581
Time × Borderline	−0.005(0.002)	−0.008 to −0.001	−2.8	509.40	**.005**		
Narcissistic	0.10(0.19)	−0.276 to 0.484	0.5	1123.70	.591	6161	0.581
Time × Narcissistic	0.004(0.008)	−0.013 to 0.021	0.5	194.71	.639		
Avoidant	−0.06(0.03)	−0.121 to 0.002	−1.9	1488.16	.056	6157	0.582
Time × Avoidant	0.0004(0.002)	−0.003 to 0.004	0.3	576.49	.789		
Dependent	0.005(0.076)	−0.143 to 0.154	0.1	1469.23	.946	6162	0.582
Time × Dependent	−0.002(0.004)	−0.010 to 0.005	−0.6	709.05	.549		
Obsessive-compulsive	0.08(0.07)	−0.047 to 0.209	1.2	1407.90	.216	6159	0.582
Time × Obsessive-compulsive	0.001(0.003)	−0.005 to 0.007	0.2	455.75	.816		
Final model: Time × all PD predictors						6114	0.581
Inadequate × Comorbidity							
PD criteria	0.014(0.002)	0.010 to 0.019	6.6	1363.33	**<.001**	5789	0.586
Time × PD criteria	−0.0003(0.0001)	−0.0005 to −0.0001	−2.5	491.08	**.013**		
GSI	0.09(0.22)	0.043 to 0.130	3.9	1515.99	**<.001**	6146	0.582
Time × GSI	−0.001(0.001)	−0.003 to 0.001	−0.7	575.87	.484		
Final model: Time × all comorbidity predictors						5789	0.587
Confident							
Intercept	2.67(0.02)	2.623 to 2.710	127.7	1621.78	**<.001**	9092	0.557
Time	0.006(0.001)	0.004 to 0.008	5.8	680.04	**<.001**		
Confident × PD Category							
Paranoid	−0.20(0.08)	−0.358 to −0.042	−2.5	1652.52	**.013**	9090	0.557
Time × Paranoid	0.004(0.004)	−0.004 to 0.013	1.0	689.23	.303		
Antisocial	0.27(0.23)	−0.185 to 0.735	1.2	2597.67	.241	9086	0.559
Time × Antisocial	−0.06(0.020)	−0.092 to −0.018	−2.9	3603.80	**.004**		
Borderline	−0.24(0.05)	−0.341 to −0.142	−4.8	1529.80	**<.001**	9063	0.557
Time × Borderline	0.002(0.002)	−0.003 to 0.007	0.8	619.73	.441		
Narcissistic	−0.87(0.28)	−1.418 to −0.332	−3.2	1227.42	**.002**	9082	0.557
Time × Narcissistic	0.006(0.012)	−0.018 to 0.030	0.5	244.30	.626		
Avoidant	0.07(0.05)	−0.014 to 0.161	1.6	1602.63	.101	9094	0.558
Time × Avoidant	−0.002(0.002)	−0.006 to 0.003	−0.8	677.56	.419		
Dependent	−0.04(0.11)	−0.254 to 0.169	−0.4	1577.56	.696	9096	0.557
Time × Dependent	0.004(0.006)	−0.007 to 0.015	0.7	824.33	.490		
Obsessive-compulsive	−0.06(0.09)	−0.243 to 0.123	−0.6	1529.27	.521	9096	0.557
Time × Obsessive-compulsive	−0.0004(0.004)	−0.009 to 0.008	−0.1	550.51	.936		
Final model: Time × all PD predictors						9052	0.558
Confident × Comorbidity							
PD criteria	−0.02(0.003)	−0.024 to −0.012	−5.9	1494.27	**<.001**	8477	0.559
Time × PD criteria	9.9E-5(0.0002)	−0.002 to 0.0004	0.6	587.12	.529		
GSI	−0.08(0.03)	−0.144 to −0.190	−2.6	1638.75	**.011**	9088	0.557
Time × GSI	0.001(0.002)	−0.002 to 0.004	0.5	688.65	.591		
Final model: Time × all comorbidity predictors						8480	0.559
Idealized							
Intercept	0.98(0.02)	0.942 to 1.025	46.8	1609.34	**<.001**	8985	0.663
Time	0.005(0.001)	0.003 to 0.008	4.9	678.98	**<.001**		
Idealized × PD Category							
Paranoid	0.08(0.08)	−0.084 to 0.235	1.0	1650.45	.354	8987	0.663
Time × Paranoid	0.001(0.005)	−0.008 to 0.010	0.2	718.50	.811		
Antisocial	−0,16 (0.23)	−0.618 to 0.291	−0.7	2506.11	.481	8988	0.663
Time × Antisocial	0.005(0.019)	−0.031 to 0.042	0.3	3257.50	.724		
Borderline	0.09(0.05)	−0.012 to 0.189	1.7	1521.62	.083	8983	0.663
Time × Borderline	0.001(0.003)	−0.004 to 0.006	0.3	600.26	.796		

(*Continued*)

Table 2. Continued.

Dependent variable and parameter	Estimate (SE)	95% Confidence interval	t	df	*p*	AIC	R2
Narcissistic	−0.04 (0.28)	−0.588 to 0.516	−0.1	1231.23	.899	8988	0.633
Time × Narcissistic	0.010(0.013)	−0.016 to 0.037	0.8	264.35	.434		
Avoidant	−0.04(0.05)	−0.123 to 0.048	−0.9	1594.92	.369	8982	0.662
Time × Avoidant	−0.003(0.002)	−0.008 to 0.001	0.2	660.89	.169		
Dependent	0.14(0.11)	−0.071 to 0.355	1.3	1568.29	.192	8977	0.662
Time × Dependent	0.010(0.0069)	−0.002 to 0.021	1.7	791.02	.091		
Obsessive-compulsive	0.01(0.09)	−0.179 to 0.190	0.1	1524.38	.953	8986	0.663
Time × Obsessive-compulsive	0.006(0.005)	−0.003 to 0.015	1.3	559.70	.188		
Idealized × Comorbidity							
PD criteria	0.004(0.004)	−0.003 to 0.009	1.1	1495.67	.318	8488	0.664
Time × PD criteria	3.1E-5(0.0002)	−0.0003 to 0.0004	0.2	609.60	.857		
GSI	0.03(0.03)	−0.035 to 0.091	0.9	1622.01	.380	8984	0.663
Time × GSI	0.002(0.002)	−0.002 to 0.005	0.9	679.67	.360		

Note: Linear mixed models. *Inadequate*, *Confident* and *Idealized* with baseline (intercept estimates) and longitudinal deviation (slope estimates) associated with PD category, PD comorbidity (PD criteria = number of SCID-II criteria) and GSI (Global Severity Index) as predictors. Bold = *p*-value is significant at the 0.01 or 0.05 levels. Goodness of fit is indicated by Akaikes Information Criterion (AIC), were smaller is better. R2 = Conditional R2.

including only those with two or more assessments (N = 1141). To investigate the impact of different treatment duration, its associations to CT were investigated as a separate predictor in LMM. Treatment duration was not associated with deviating longitudinal CT levels of all three CT dimensions.

Results

Longitudinal Course of CT

Therapist-reported levels of *Inadequate* did not change significantly over time, whilst *Confident* and *Idealized* increased significantly over time, though modest in magnitude (Table 2 and Figure 1). Initial levels (3 months) were lowest for *Inadequate* and highest for *Confident*. MM estimated grand means (mean values over all time points per patient) were for *Inadequate* 0.50 (SE 0.01), *Confident*: 2.76 (SE 0.01) and *Idealized*: 1.07 (SE 0.01). In order to investigate possible bias caused by local unit differences, we analyzed whether the longitudinal CT course differed across units. Controlling for possible baseline variation, we found a main trend of no significant variation between the different units in change over time for *Confident* and *Inadequate* ($p > 0.05$). Only one of the 20 units had a significant deviant change of the *Idealized* CT dimension ($p = 0.03$). When reanalyzing the results without this unit,

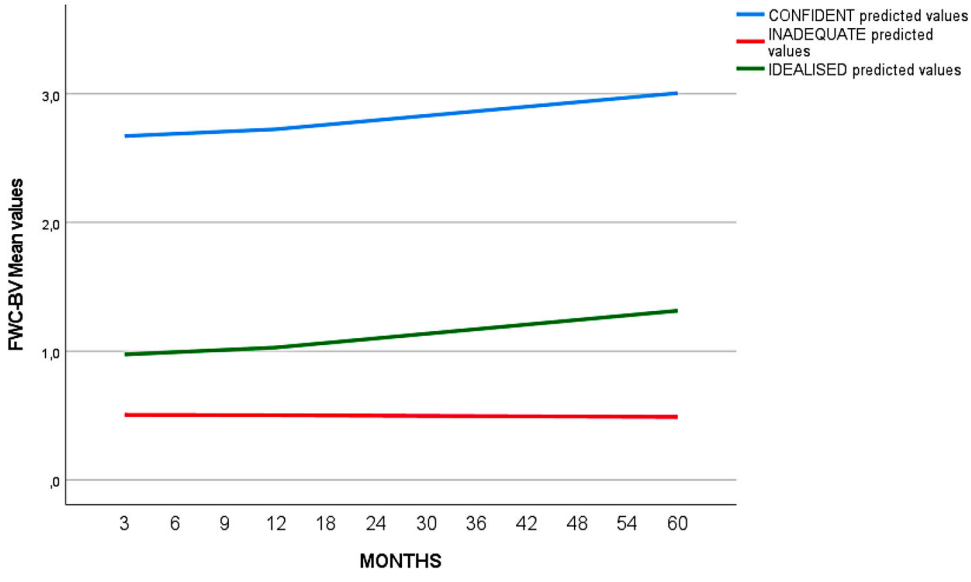

Figure 1. The CT development in therapists in the total sample over time. *Confident* and *Idealized* increased significantly over time ($p < 0.05$), while *Inadequate* did not change significantly over time ($p > 0.05$).

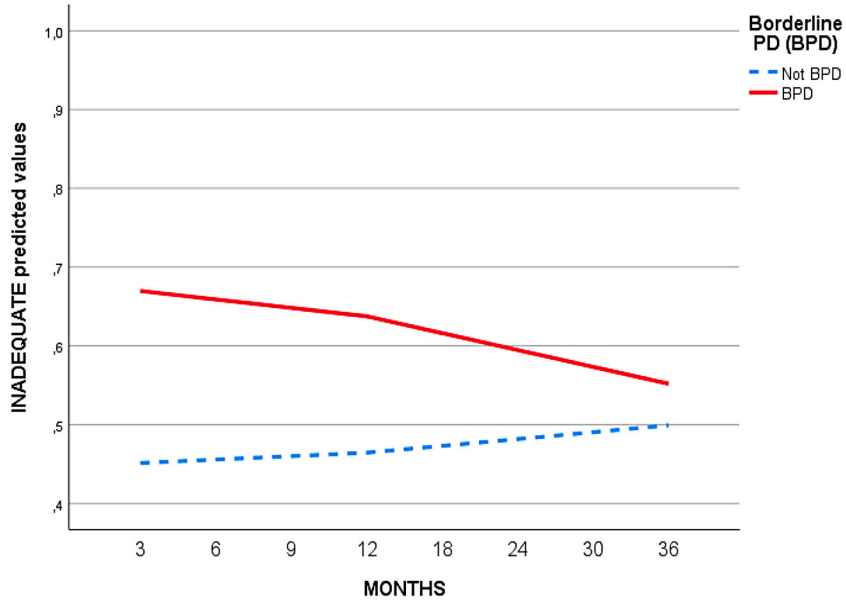

Figure 2. Borderline PD and development of *Inadequate* CT in therapists. LMM trajectories estimated for patients with Borderline PD and those without Borderline PD (dashed line). LMM slope differences were significant ($p < 0.05$). Mean treatment length for BPD was 23 months (SD = 14).

the deviant change disappeared and there was a general trend of no significant difference between units in change over time for *Idealized*.

Predictors

PD Category

The first model included *Inadequate* as the dependent variable. Paranoid PD was associated with significantly higher initial *Inadequate* but not with deviating *Inadequate* over time (Table 2). Borderline PD was associated with significantly higher initial *Inadequate* and decreasing *Inadequate* over time (Table 2, Figure 2). The remaining PDs: avoidant, dependent, and obsessive-compulsive PD, were not associated with significantly deviating initial levels or slope. In a final model adding all PDs with significantly deviating CT levels, only borderline PD remained a significant predictor of

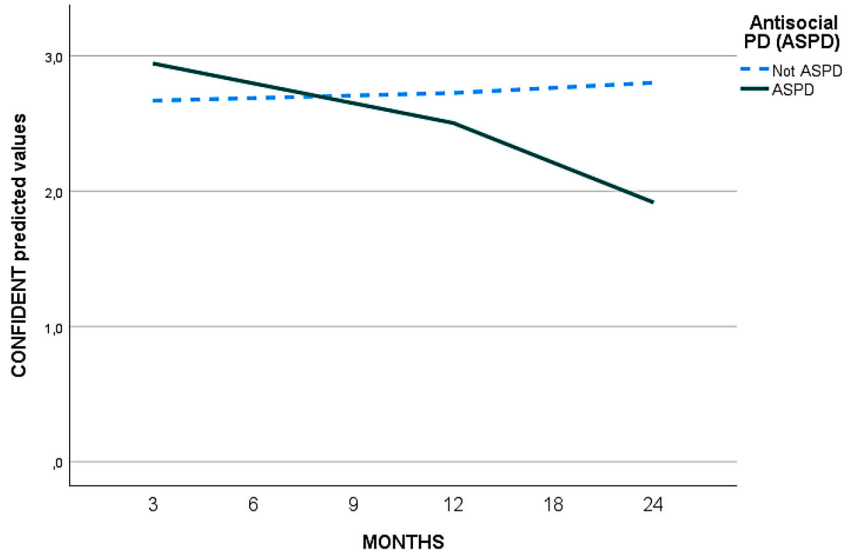

Figure 3. Antisocial PD and development of Confident CT in therapists. LMM trajectories estimated for patients with Antisocial PD and those without Antisocial PD (dashed line). LMM slope differences were significant ($p < 0.05$). Mean treatment duration for Antisocial PD was 14 months (SD = 8).

deviating *Inadequate* levels ($p < 0.05$). The next model included *Confident* as the dependent variable. Paranoid, narcissistic, and borderline PD were associated with less *Confident* response in the initial treatment phase. Neither were associated with deviating rates of change (Table 2). Antisocial PD was associated with significantly decreasing *Confident* levels over time (Table 2, Figure 3). In a final model adding all PDs with significantly deviating CT levels, significant results for borderline, narcissistic, and antisocial PD remained the same ($p < 0.05$). Lastly, *Idealized* was analyzed as dependent variable. None of the PDs predicted deviating initial or change of this CT dimension (Table 2). Of the different CT dimensions, PD category explained most variation in the *Inadequate* dimension, with borderline explaining the most variation (5% explained intercept variation, 3% explained slope variation).

In additional analyses investigating the number of PD criteria within each PD category, the results were largely the same as when PD was investigated categorically. That is, borderline criteria remained the most significant predictor of deviating initial *Inadequate* and deviating *Inadequate* change over time in addition to deviating initial *Confident* response ($p < 0.05$). The number of borderline criteria was not associated with significantly deviating initial and longitudinal change of *Idealized,* but specific investigation of patients with severe borderline disorder (>7 borderline criteria) revealed significantly enhanced *Idealized* response in therapists over time ($p < 0.05$).

The Total Number of PD Criteria and Level of Symptom Distress (GSI)

Total number of PD criteria was associated with higher initial *Inadequate* and lower initial *Confident*. Over time, total number of PD criteria was associated with decreasing *Inadequate*. Total number of PD criteria was not associated with deviating *Confident* over time but explained 12% variation. Higher levels of GSI were associated with higher initial *Inadequate* and lower initial *Confident* response, but did not predict change in CT over time. GSI was not a significant predictor of deviating *Idealized* (Table 2). In a final model including both GSI and total PD criteria, only total number of PD criteria remained a significant predictor of therapists *Confident* and *Inadequate* responses ($p < 0.05$).

Patient Outcome

Table 3 demonstrates clinical outcomes (change from start to end of therapy) for the variables GAF, BSI, and CIP. There was an overall improvement reflected by all three outcome variables, self-report and observer-rated ($p < 0.01$) (Table 3). GAF-S (GAF symptom) and GAF-F (GAF function) improvement were associated with lower initial *Inadequate* and GAF-S was associated with significantly decreasing *Inadequate* over time. Oppositely, no improvement/worsening on GAF-S was associated with higher initial *Inadequate* and increasing *Inadequate* over time (Table 4 and Figure 4). Both GAF-S and GAF-F improvement was associated with higher initial *Confident,* but not with deviating change over time. Only GAF-S improvement was associated with deviance of *Idealized* levels, with higher initial *Idealized* but not deviating *Idealized* over time (supplementary file, Table 4). Correspondingly, patient self-report revealed CIP and GSI improvement associated with significantly decreasing *Inadequate* over time. Neither GSI nor CIP improvement was associated with significant deviance of *Confident* or *Idealized* levels. ($p > 0.05$).

Table 3. Patient outcome measures scores.

	Mean	SD
Global severity Index from SCL-90-R		
Pretreatment	1.52	0.66
End of treatment	1.00	0.70
Difference	0.49***	0.66
Global assessment of Functioning Scale[a]		
Pretreatment	49.67	6.09
End of treatment	57.49	9.65
Difference	7.82***	8.98
Inventory of Interpersonal Problems-Circumplex version		
Pretreatment	1.65	0.52
End of treatment	1.30	0.62
Difference	0.35***	0.62

Note: (a) Traditional GAF score based on the most severe of symptom and function. Independent samples T-test. Significant differences (improvement) in patient outcome scores GSI ($N = 1273$), GAF ($N = 1956$), CIP ($N = 1260$) is marked with ** $p < 0.01$, *** $p < 0.001$.

Discussion

The main findings in the present study were firstly, that therapists CT feelings were generally positive and quite stable over time. We registered a modest increase in *Confident* and *Idealized* responses. Levels of *Inadequate* were low and change over time was insignificant. Secondly, increasing severity of personality pathology and borderline PD, specifically, predicted significantly more negative CT in therapists at the beginning of therapy. However, the *Inadequate* response decreased over time. Global symptom

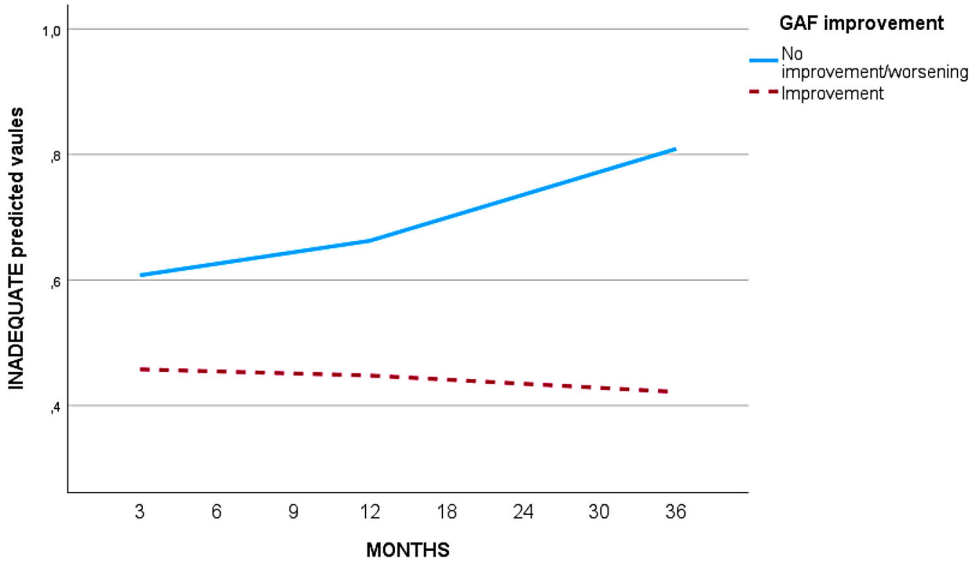

Figure 4. *Inadequate* development associated with patient clinical outcome (GAF-S). LMM trajectories estimated for those with worsening on GAF-S and those with improvement on GAF-S (dashed line). LMM slope differences were significant ($p < 0.05$).

severity among the patients had less impact on therapist CT than patients' severity of personality pathology.

The clinical impact of the significant predictors could however be discussed. That is, we found significant associations between several patient characteristics and CT, but the predictors were associated with small nuances on the CT intensity on the FWC scale. Based on the intensity of CT feelings typically reported in the clinical, theoretical literature on PD treatment, it is also noteworthy how therapists in this study report low levels of negative feelings. However, the score-levels are comparable to those reported in other studies that have examined PD and CT using various versions of FWC as instrument (Breivik et al., 2020; Dam et al., 2021; Rossberg et al., 2010). The levels are also comparable to other studies that have not examined PD specifically (Dahl et al., 2012; Falkenstrøm & Holmqvist, 2022; Holmqvist, 2001; Lindqvist et al., 2017; Ulberg et al., 2013). That is, positive CT responses (e.g., feeling Confident, Calm) are systematically rated higher than negative CT responses (e.g., feeling Inadequate, Disengaged, Overwhelmed), and the intensity of negative CT is generally low. The average "expectable" mean based on the results reported in other FWC studies typically ranges from 0 to 1.5 on dimensions covering negative CT and between 1 and 2.5 on positive CT. An exception is Dahl et al's study (2012) that has noteworthy lower scores on their *Confident* subscales. The heading on the questionnaire in their study is "Countertransference" while the other studies use headings like "Together with the patient during this session I have felt: … " The present study uses the heading: "During recent conversations with the patient I have felt: … " It is reasonable to believe, as also discussed in the study of Dahl and colleagues (2012) that different headings have an impact on the way therapists respond. Immediate reactions after e.g., one therapy session would probably have revealed more intense and possibly varied reactions.

Other possible reasons for the relatively high *Confident* feelings might be that the participating units in this study represent teams of collaborating therapists and are specialized in the treatment of personality disorders. Recent studies highlight negative provider attitudes toward people with personality disorders, particularly those with borderline PD (Bodner et al., 2015; Sheehan et al., 2016). However, several recent studies also illustrate that actual contact between clinicians and patients with PD reduces stigma and increases positive attitudes (Dam et al., 2021; Van den Bosch et al., 2018). Further, the therapists in the present study attend regular supervision meetings and clinical discussions where CT is likely to be a central focus. As indicated by several other studies on CT and therapist experience (Brody & Farber, 1996; Ulberg et al., 2013), and supervision (Ulberg et al., 2013); those with greater clinical experience seem more comfortable with their emotional reactions to patients. However, the risk of defensiveness and/or social desirability bias is always inherent when using self-reports. That is,

they may be more prone to show specific reactions than others. They might "hide" more of their negative reactions or they may simply not be aware of their reactions. This might also influence the results of the present study.

It is also worth noting that although *Confident* was the feeling response ranked strongest by the therapists, it does not imply that this is a constant, prevailing CT response in all patient encounters. Similarly, even though improvement in outcome was the main trend, it does not mean that everyone improves. We found that the subgroup that did not show improvement (approximately 30%), was associated with increasing negative CT. Treatment dropouts is also a relevant group that was not studied in the present study. Although dropout rates were generally quite low (14%) compared to other reviews (Barnicot et al., 2011), and treatment duration was not found to be related to deviating longitudinal CT, there is still a good reason to believe that this group elicits more challenging reactions. Therefore, several aspects may have not been captured in this initial "overview" investigation of the data.

In line with several other empirical studies (Betan et al., 2005; Colli et al., 2014; Tanzilli et al., 2016), we found that borderline PD is associated with more *Inadequate* responses. In fact, borderline pathology was the most significant PD predictor of therapists' *Inadequate* responses, with significantly higher *Inadequate* response at the beginning of therapy, but also, interestingly, the PD associated with the most significant decrease in *Inadequate* response over time. The trend that the negative CT decreases significantly over time could be an encouraging result for the treatment of this group. This might be related to improvement (Kvarstein et al., 2015; Morken et al., 2019). A former study of a corresponding sample from the Norwegian Network of Personality disorders has demonstrated that borderline PD patients who adhered to psychotherapeutic, group-based treatment were associated with high initial symptom levels, but differences from other PDs were compensated by significantly greater rates of improvement over time (Kvarstein & Karterud, 2013). However, specific investigation of interactions between CT responses, treatment factors and outcomes is outside the scope of the present study.

Surprisingly, we did not find that borderline PD predicted significantly more *Idealized* responses in therapists. Our results thus contrast a previous correlational study with a similar patient sample (Breivik et al., 2020). It might happen that the *Idealized* dimension (Important, Exalted, Admired) does not capture the "splitting" mechanism described in clinical literature (Kernberg, 1985), with therapists typically experienced as "all good" and idealized or "all bad" and devaluated by their patients, but rather resembles a more *Confident* response. However, it might also be that such mechanisms are more prominent among patients with more severe borderline conditions. In support of this assumption, is the finding that a high number of borderline PD criteria was a significant predictor of *Idealized* response in therapists.

The largest patient group in the present study was avoidant PD. Some studies show that avoidant PD is associated with severe dysfunction and more modest treatment outcomes (Kvarstein et al., 2021; Wilberg et al., 2009). A few empirical studies have addressed negative responses of helpless (Tanzilli et al., 2016) and disengagement (Genova & Gazillo) in addition to positive CT response. Along the same line, a recent qualitative study (Pettersen, 2021) demonstrated quite heterogonous response to avoidant PD patients, including also negative feelings such as impatience, irritation, and provocation partly elicited by patients' general withdrawal and reluctance to share. From a clinical perspective, it is reason to believe that therapists may experience more negative feelings over time than previous studies have reported, especially toward more poorly functioning avoidant PD patients. We were thus especially curious about the development of *Inadequate* CT over time with avoidant patients. However, we did not find that avoidant PD elicited more *Inadequate* feelings longitudinally. It might be that the brief 10-item questionnaire used in this study does not adequately address the possibly heterogeneous therapist response to avoidant patients.

Of the other significant predictors, we found that antisocial PD was associated with decreasing *Confident* CT over time. This result was quite surprising, given the small number of patients with such traits. In mental health care, these patients are described to evoke strong negative emotions in therapists that often lead to exclusion from treatment programmes (Dam et al., 2021; Sheehan et al., 2016; Van den Bosch et al., 2018). Even though the treatment units focus on personality disorder, no specific programmes existed for antisocial PD at the time, and there is generally a lack of empirical evidence for any effective treatment with this patient pathology (Gibbon et al., 2020). One might speculate if this lack of knowledge contributes to the less positive CT response over time. Another possible explanation is that they may have poor improvement during treatment. Based on treatment duration in this study, this was the PD group with the shortest average treatment length compared to the other PD groups and the total sample, with a mean treatment length of 14.1 (SD = 7.8) months, considerably shorter than the average. However, as mentioned, different

treatment factors and outcome was not the main focus in this study.

Lastly, and one noteworthy finding, was that symptom distress had less impact on therapists' CT feelings than severity of personality pathology. More specifically, the total number of PD criteria had a higher predictive value than the level of global symptom distress for therapist CT reactions. It might support the importance of early targeting and working with relational aspects, instead of focusing solely on patients' level of symptom distress (Grenyer, 2002).

Our additional analyses coupled improvement on clinical outcome measures with decreasing *Inadequate* CT over time, and oppositely, those with no improvement/worsening with increasing *Inadequate*. The scope of this study does not provide detail on concurrent change or possible mediating factors, but nonetheless, highlights a mutual nature of the therapeutic dialogue, process and resulting outcome, in line with recent CT studies (Colli et al., 2022; Nissen Lie et al., 2022; Tishby & Wiseman, 2022).

Strengths and Limitations

The strength of the present study is the uniquely large sample of therapies and number of CT assessments, which is, as far as known, the largest study on CT responses and personality disorders from a longitudinal perspective. Being based on clinical data retrieved from a quality register, the patients included in this study are highly representative of real-life treatments for people with PDs and personality problems.

Missing CT assessments in the present study is a limitation. In addition, differences in treatment duration naturally cause different numbers of CT assessments. However, we have included investigation of possible differences associated with different numbers of assessment and of different treatment length. Analyses did not reveal significant bias.

The scope of this study with a quality register did not provide detailed information on the therapist qualifications nor enable analyses of variation between therapists. More detailed investigation of therapist factors could have given valuable information such as to which extent some therapists were more prone to experience specific CT responses. It would also be interesting to investigate CT in other areas of the public health sector and with trainees instead of experienced therapists, where CT responses may be more intense and varied. The heading of the questionnaire, which asks about responses from "recent conversations", is also a factor that might contribute to a lower range of emotions in general.

This study is based on routine clinical data and as such, its systematic quality can be judged as good. We have described systematic procedures using semi-structured interviews and training for therapists in the Network. However, the reliability of PD classification was not tested. Generally, descriptive DSM diagnoses may be seen as an oversimplification of a more complex reality and reductive about the varieties of personality (McWilliams et al., 2018). Another possible limitation is that the therapists were naturally not blind to the patients' diagnoses, and it is conceivable that therapists' expectations to specific diagnostic characteristics could bias their report on CT. As the study is observational and investigates associations between therapist CT and patient characteristics, its design implies that causality cannot be concluded.

We used self-reports to measure CT. Failure to recognize unconscious feelings and social desirability bias are inherent in this method. A way to overcome this issue in future studies is to use both therapist's self-report and e.g., video-sessions or session transcripts evaluated and assessed by independent raters. In the present large-sampled study, however, such fine evaluation is unrealistic. In addition, the brief 10-item FWC do not include all CT responses described in the clinical literature, for example erotized or angry CT.

Moreover, even though we found associations between improvement in patient outcome and therapist CT, the scope of this study did not allow further investigation of concurrent change patterns, interactions, mediating factors or relative impacts of the different variables.

Conclusion and Implications

This study is one of the first large-scale longitudinal investigations of therapists' CT responses when treating patients with PD. A noteworthy result is that therapists in the specialized PD treatment units reported predominantly *Confident* CT responses. Although more severe personality pathology, and borderline PD, specifically, was associated with more negative CT responses at the beginning of therapy, these negative feelings tended to decrease over time. Patients' clinical improvement related to a decrease in therapists' negative CT. However, we also found that those who did not improve were associated with increasing *Inadequate* CT. The findings point to the importance of supervision in therapies targeting patients' PD problems, helping therapist reflect over and resolve difficult and

counterproductive CT responses. If increasing negative CT feelings are warded off, and not managed, this may contribute to more CT being acted out (Hayes et al., 2018).

There is now increasing empirical support for the importance of being attentive to ruptures in the therapeutic alliance, and the association with awareness of negative CT is therefore particularly relevant. Further, the findings in the present study suggest that focus on personality-related aspects might be more important than patients' symptom level when treating patients with PD. Previous research addresses that discussion of what is happening between therapist and patient may be particularly important for outcome in individuals with personality disorders or low quality of object relations (Dahl et al., 2012; Høglend et al., 2006). The study highlights a need for more detailed research within the field of PD treatment focusing on therapist CT, mutual alliance, and furthering this line, also effects of supervision for therapists and treating PD.

Acknowledgements

We wish to thank the patients and staff of the Norwegian Network for Personality Norwegian Network of Personality-Disorder for their contribution to this study.

Disclosure Statement

No potential conflict of interest was reported by the author(s).

Availability of Data and Materials

Due to restrictions imposed by the Regional Medical Ethics Committee regarding patient confidentiality, data are available upon request. Requests for data may be sent to the hospital's Privacy and Data Protection Officer at: personvernous-hf.no.

Authors' Contributions

RB: data analysis and drafting and revising the manuscript; GP: organization of the quality registry, local data collection and revising the current manuscript. TW: revising the manuscript; JE: revising the manuscript; JIR: revising the manuscript; HSD: revising the manuscript; EK: data analysis and drafting and revising the manuscript. All authors read and approved the final manuscript.

References

Alden, L. E., Wiggins, J. S., & Pincus, A. L. (1990). Construction of circumplex scales for the inventory of interpersonal problems. *Journal of Personality Assessment*, *55*(3–4), 521–536. https://doi.org/10.1080/00223891.1990.9674088

American Psychiatric Association (APA). (1994). *Diagnostic and statistical manual of mental disorders* (4th ed.). APA.

Barnicot, K., Katsakou, C., Marougka, S., & Priebe, S. (2011). Treatment completion in psychotherapy for borderline personality disorder—A systematic review and meta-analysis. *Acta Psychiatrica Scandinavica*, *123*(5), 327–338. https://doi.org/10.1111/j.1600-0447.2010.01652.x

Betan, E., Heim, A. K., Zittel Conklin, C., & Westen, D. (2005). Countertransference phenomena and personality pathology in clinical practice: An empirical investigation. *American Journal of Psychiatry*, *162*(5), 890–898. https://doi.org/10.1176/appi.ajp.162.5.890

Biondi, M. (2022). *The clinician in the psychiatric diagnostic process* (pp. 151–163). Springer International Publishing AG. https://doi.org/10.1007/978-3-030-90431-9_10

Bodner, E., Cohen-Fridel, S., Mashiah, M., Segal, M., Grinshpoon, A., Fischel, T., & Iancu, I. (2015). The attitudes of psychiatric hospital staff toward hospitalization and treatment of patients with borderline personality disorder. *BMC Psychiatry*, *15*(1), 2–2. https://doi.org/10.1186/s12888-014-0380-y

Breivik, R., Wilberg, T., Evensen, J., Røssberg, J. I., Dahl, H. S., & Pedersen, G. (2020). Countertransference feelings and personality disorders: A psychometric evaluation of a brief version of the Feeling Word Checklist (FWC-BV). *BMC Psychiatry*, *20*(1), 1–12. https://doi.org/10.1186/s12888-020-02556-6

Brody, E. M., & Farber, B. A. (1996). The effects of therapist experience and patient diagnosis on countertransference. *Psychoterapy: Theory, Research, Practice, Training*, *33*(3), 372–380. https://doi.org/10.1037/0033-3204.33.3.372

Carsky, M. (2021). Managing countertransference in the treatment of personality disorders. *Psychodynamic Psychiatry*, *49*(2), 339–360. https://doi.org/10.1521/pdps.2021.49.2.339

Colli, A., Gagliardini, G., & Gullo, S. (2022). Countertransference responses mediate the relationship between patients' overall defense functioning and therapists' interventions. *Psychotherapy Research*, *32*(1), 32–45. https://doi.org/10.1080/10503307.2021.1884768

Colli, A., Tanzilli, A., Dimaggio, G., & Lingiardi, V. (2014). Patient personality and therapist response: An empirical investigation. *American Journal of Psychiatry*, *171*(1), 102–108. https://doi.org/10.1176/appi.ajp.2013.13020224

Dahl, H. S., Rossberg, J. I., Bogwald, K. P., Gabbard, G. O., & Hoglend, P. A. (2012). Countertransference feelings in one year of individual therapy: An evaluation of the factor structure in the Feeling Word Checklist-58. *Psychotherapy Research*, *22*(1), 12–25. https://doi.org/10.1080/10503307.2011.622312

Dam, A., Rijckmans, M., & Bosch, L. (2021). Explaining the willingness of clinicians to work with patients with antisocial personality disorder using the theory of planned behaviour and emotional reactions. *Clinical Psychology & Psychotherapy*, *29*(2), 676–686.

Derogatis, L. R. (1994). *Symptom checklist-90-R (SCL-90-R): Administration, scoring and procedures manual*. National Computer System.

DSM-IV. (1994). *Diagnostic and statistical manual of mental disorders* (4th ed.). APA.

Falkenstrøm, F., & Holmqvist, R. (2022). Therapist in-session feelings predict change in depressive symptoms in interpersonal and

brief relational psychoterapy. *Psychotherapy Research*, *32*(5), 571–584. https://doi.org/10.1080/10503307.2021.1998700

Fauth, J. (2006). Toward more (and better) countertransference research. *Psychotherapy (Chic)*, *43*(1), 16–31. https://doi.org/10.1037/0033-3204.43.1.16

First, M. B. (1997). *User's guide for the structured clinical interview for DSM-IV axis II personality disorders: SCID-II*. American Psychiatric Press.

Fitzmaurice, G. M., Laird, N. M., & Ware, J. H. (2012). *Applied longitudinal analysis (Vol. 998)*. John Wiley & Sons.

Freud, S., Strachey, J., Freud, A., & Institute of, p.-a. (1957). *The standard edition of the complete psychological works of Sigmund Freud: Vol. 11: Five lectures on psycho-analysis; Leonardo da Vinci, and other works: (1910)* (Vol. 11). Hogarth Press and the Institute of Psycho-analysis.

Gabbard, G. O. (2017). *Long-term psychodynamic psychotherapy: A basic text*. American Psychiatric Pub.

Gabbard, G. O. (2020). The role of countertransference in contemporary psychiatric treatment. *World Psychiatry*, *19*(2), 243. https://doi.org/10.1002/wps.20746

Gabbard, G. O., & Wilkinson, S. M. (2000). *Management of countertransference with borderline patients*. Jason Aronson.

Gelso, C. J., & Kline, K. V. (2022). Some directions for research and theory on countertransference. *Psychotherapy Research*, *32*(1), 46–51. https://doi.org/10.1080/10503307.2021.1968529

Genova, F., & Gazzillo, F. (2018). Personality organization, personality styles, and the emotional reactions of treating clinicians. *Psychodynamic Psychiatry*, *46*(3), 357–392. https://doi.org/10.1521/pdps.2018.46.3.357

Gibbon, S., Gibbon, Khalifa, Cheung, Völlm, & McCarthy. (2020). Psychological interventions for antisocial personality disorder. *Cochrane Library*, *2020*(9). https://doi.org/10.1002/14651858.CD007668.pub3

Grenyer, B. F. A. (2002). *Mastering relationship conflicts: Discoveries in theory, research, and practice*. American Psychological Association.

Hayes, J. A., Gelso, C. J., Goldberg, S., & Kivlighan, D. M. (2018). Countertransference management and effective psychotherapy: Meta-analytic findings. *Psychotherapy (Chic)*, *55*(4), 496–507. https://doi.org/10.1037/pst0000189

Hayes, J. A., Nelson, D. L., & Fauth, J. (2015). Countertransference in successful and unsuccessful cases of psychotherapy. *Psychotherapy (Chic)*, *52*(1), 127–133. https://doi.org/10.1037/a0038827

Hedeker, D., & Gibbons, R. D. (1997). Application of random-effects pattern-mixture models for missing data in longitudinal studies. *Psychological Methods*, *2*(1), 64–78. https://doi.org/10.1037/1082-989X.2.1.64

Høglend, P., Amlo, S., Marble, A., Bøgwald, K., Sørbye, Ø, Sjaastad, M., & Heyerdahl, O. (2006). Analysis of the patient-therapist relationship in dynamic psychotherapy: An experimental study of transference interpretations. *The American Journal of Psychiatry*, *163*(10), 1739–1746. https://doi.org/10.1176/ajp.2006.163.10.1739

Holmqvist, R. (2001). Patterns of consistency and deviation in therapists' countertransference feelings. *The Journal of Psychotherapy Practice and Research*, *10*(2), 104.

Kernberg, O. F. (1985). *Borderline conditions and pathological narcissism*. Rowman & Littlefield.

Kernberg, O. F. (1989). *Psychodynamic psychotherapy of borderline patients*. Basic Books.

Kvarstein, E. H., Antonsen, B. T., Klungsøyr, O., Pedersen, G., & Wilberg, T. (2021). Avoidant personality disorder and social functioning: A longitudinal, observational study investigating predictors of change in a clinical sample. *Personality Disorders*, *12*(6), 594–605. https://doi.org/10.1037/per0000471

Kvarstein, E. H., & Karterud, S. (2013). Large variation of severity and longitudinal change of symptom distress among patients with personality disorders. *Personality and Mental Health*, *7*(4), 265–276. https://doi.org/10.1002/pmh.1226

Kvarstein, E. H., Pedersen, G., Urnes, Ø, Hummelen, B., Wilberg, T., & Karterud, S. (2015). Changing from a traditional psychodynamic treatment programme to mentalization-based treatment for patients with borderline personality disorder—does it make a difference? *Psychology and Psychotherapy: Theory, Research and Practice*, *88*(1), 71–86. https://doi.org/10.1111/papt.12036

Lindqvist, K., Falkenstrom, F., Sandell, R., Holmqvist, R., Ekeblad, A., & Thoren, A. (2017). Multilevel exploratory factor analysis of the Feeling Word Checklist-24. *Assessment*, *24*(7), 907–918. https://doi.org/10.1177/1073191116632336

Lingiardi, V., Tanzilli, A., & Colli, A. (2015). Does the severity of psychopathological symptoms mediate the relationship between patient personality and therapist response? *Psychotherapy (Chic)*, *52*(2), 228–237. https://doi.org/10.1037/a0037919

McWilliams, N. (2011). *Psychoanalytic diagnosis: Understanding personality structure in the clinical process*. Guilford Press.

McWilliams, N., Grenyer, B. F. S., & Shedler, J. (2018). *Personality in PDM-2: Controversial issues* [299-305]. Hillsdale, N.J.

Morken, K. T., Binder, P. E., Arefjord, N., & Karterud, S. (2019). Juggling thoughts and feelings: How do female patients with borderline symptomology and substance use disorder experience change in mentalization-based treatment? *Psychotherapy Research*, *29*(2), 251–266. https://doi.org/10.1080/10503307.2017.1325021

Nakagawa, S., Johnson, P. C., & Schielzeth, H. (2017). The coefficient of determination R 2 and intra-class correlation coefficient from generalized linear mixed-effects models revisited and expanded. *Journal of the Royal Society Interface*, *14*(134), 20170213. https://doi.org/10.1098/rsif.2017.0213

Nissen-Lie, H. A., Dahl, H.-S. J., & Høglend, P. A. (2022). Patient factors predict therapists' emotional countertransference differently depending on whether therapists use transference work in psychodynamic therapy. *Psychotherapy Research*, *32*(1), 3–15. https://doi.org/10.1080/10503307.2020.1762947

Norusis, M. J. (2008). *SPSS 16.0 advanced statistical procedures companion*. Prentice Hall Inc.

Oldham, J. M., Skodol, A. E., Kellman, H. D., Hyler, S. E., Doidge, N., Rosnick, L., & Gallaher, P. E. (1995). Comorbidity of axis I and axis II disorders. *American Journal of Psychiatry*, *152*(4), 571–578. https://doi.org/10.1176/ajp.152.4.571

Pedersen, G. (2002). Revised Norwegian version of inventory of interpersonal problems—Circumplex (IIP-C). *Tidsskrift Norsk Psykologforening*, *39*(1), 25–34.

Pedersen, G., Hagtvet, K. A., & Karterud, S. (2007). Generalizability studies of the global assessment of functioning—split version. *Comprehensive Psychiatry*, *48*(1), 88–94. https://doi.org/10.1016/j.comppsych.2006.03.008

Pedersen, G., & Karterud, S. (2004). Is SCL-90R helpful for the clinician in assessing DSM-IV symptom disorders? *Acta Psychiatrica Scandinavica*, *110*(3), 215–224. https://doi.org/10.1111/j.1600-0447.2004.00321.x

Pedersen, G., & Karterud, S. (2010). Using measures from the SCL-90-R to screen for personality disorders. *Personality and Mental Health*, *4*(2), 121–132. https://doi.org/10.1002/pmh.122

Pedersen, G., & Karterud, S. (2012). The symptom and function dimensions of the Global Assessment of Functioning (GAF) scale. *Comprehensive Psychiatry*, *53*(3), 292–298. https://doi.org/10.1016/j.comppsych.2011.04.007

Pedersen, G., Karterud, S., Hummelen, B., & Wilberg, T. (2013). The impact of extended longitudinal observation on the

assessment of personality disorders. *Personality and Mental Health*, *7*(4), 277–287. https://doi.org/10.1002/pmh.1234

Pedersen, G., Wilberg, T., Hummelen, B., & Kvarstein, E. H. (2022). The Norwegian network for personality disorders—Development, contributions and challenges through 30 years. *Nordic Journal of Psychiatry*. https://doi.org/10.1080/08039488.2022.2147995

Pettersen, M. I. S. (2021). Therapists' experiences with mentalization-based treatment for avoidant personality disorder. *European Journal of Qualitative Research in Psychotherapy*, *11*, 143–159.

Rocco, D., De Bei, F., Negri, A., & Filipponi, L. (2021). The relationship between self-observed and other-observed countertransference and session outcome. *Psychotherapy*, 301–309. https://doi.org/10.1037/pst0000356

Rossberg, J. I., Hoffart, A., & Friis, S. (2003). Psychiatric staff members' emotional reactions toward patients. A psychometric evaluation of an extended version of the Feeling Word Checklist (FWC-58). *Nordic Journal of Psychiatry*, *57*(1), 45–53. https://doi.org/10.1080/08039480310000257

Rossberg, J. I., Karterud, S., Pedersen, G., & Friis, S. (2007). An empirical study of countertransference reactions toward patients with personality disorders. *Comprehensive Psychiatry*, *48*(3), 225–230. https://doi.org/10.1016/j.comppsych.2007.02.002

Rossberg, J. I., Karterud, S., Pedersen, G., & Friis, S. (2010). Psychiatric symptoms and countertransference feelings: An empirical investigation. *Psychiatry Research*, *178*(1), 191–195. https://doi.org/10.1016/j.psychres.2009.09.019

Sheehan, D., Lecrubier, Y., & Janavs, J. (1994). *Mini-International Neuropsychiatric Interview (MINI)*. University of South Florida. Institute for Research in Psychiatry.

Sheehan, L., Nieweglowski, K., & Corrigan, P. (2016). The stigma of personality disorders. *Current Psychiatry Reports*, *18*(1), 1–7. https://doi.org/10.1007/s11920-015-0654-1

Singer, J. D., Willett, J. B., & Willett, J. B. (2003). *Applied longitudinal data analysis: Modeling change and event occurrence*. Oxford university press.

Spitzer, R. L. (1983). Psychiatric diagnosis: Are clinicians still necessary? *Comprehensive Psychiatry*, *24*(5), 399–411. https://doi.org/10.1016/0010-440X(83)90032-9

Stefana, A., Bulgari, V., Youngstrom, E. A., Dakanalis, A., Bordin, C., & Hopwood, C. J. (2020). Patient personality and psychotherapist reactions in individual psychotherapy setting: A systematic review. *Clinical Psychology & Psychotherapy*, *27*(5), 697–713. https://doi.org/10.1002/cpp.2455

Tanzilli, A., Colli, A., Del Corno, F., & Lingiardi, V. (2016). Factor structure, reliability, and validity of the therapist response questionnaire. *Personality Disorders: Theory, Research, and Treatment*, *7*(2), 147–158. https://doi.org/10.1037/per0000146

Tanzilli, A., Muzi, L., Ronningstam, E., & Lingiardi, V. (2017). Countertransference when working with narcissistic personality disorder: An empirical investigation. *Psychotherapy*, *54*(2), 184–194. https://doi.org/10.1037/pst0000111

Thylstrup, B., & Hesse, M. (2008). Substance abusers' personality disorders and staff members' emotional reactions. *BMC Psychiatry*, *8*(1), 1–6. https://doi.org/10.1186/1471-244X-8-21

Tishby, O., & Wiseman, H. (2022). Countertransference types and their relation to rupture and repair in the alliance. *Psychotherapy Research*, *32*(1), 16–31. https://doi.org/10.1080/10503307.2020.1862934

Ulberg, R., Falkenberg, A., Naerdal, T., Johannessen, H., Olsen, J., Eide, T., Hersoug, A., Dahl, H.-S., & Ulberg, R. (2013). Countertransference feelings when treating teenagers. A psychometric evaluation of the Feeling Word Checklist-24. *American Journal of Psychotherapy*, *67*(4), 347–358. https://doi.org/10.1176/appi.psychotherapy.2013.67.4.347

Van den Bosch, L. M. C., Rijckmans, M. J. N., Decoene, S., & Chapman, A. L. (2018). Treatment of antisocial personality disorder: Development of a practice focused framework. *International Journal of law and Psychiatry*, *58*, 72–78. https://doi.org/10.1016/j.ijlp.2018.03.002

Wilberg, T., Karterud, S., Pedersen, G., & Urnes, Ø. (2009). The impact of avoidant personality disorder on psychosocial impairment is substantial. *Nordic Journal of Psychiatry*, *63*(5), 390–396. https://doi.org/10.1080/08039480902831322