Cognitive Therapy and Research (2018) 42:219–229
https://doi.org/10.1007/s10608-018-9904-y

ORIGINAL ARTICLE

Unwanted Events and Side Effects in Cognitive Behavior Therapy

Marie-Luise Schermuly-Haupt[1] · **Michael Linden**[1,2] · **A. John Rush**[3]

Published online: 9 March 2018

Abstract
Side effects (SEs) are negative reactions to an appropriately delivered treatment, which must be discriminated from unwanted events (UEs) or consequences of inadequate treatment. One hundred CBT therapists were interviewed for UEs and SEs in one of their current outpatients. Therapists reported 372 UEs in 98 patients and SEs in 43 patients. Most frequent were "negative wellbeing/distress" (27% of patients), "worsening of symptoms" (9%), "strains in family relations" (6%); 21% of patients suffered from severe or very severe and 5% from persistent SEs. SEs are unavoidable and frequent also in well-delivered CBT. They include both symptoms and the impairment of social life. Knowledge about the side effect profile can improve early recognition of SEs, safeguard patients, and enhance therapy outcome.

Keywords Psychotherapy · Unwanted events · Side effects · Adverse treatment reactions · Quality assurance · Cognitive behavior therapy · Deterioration

Introduction

Psychotherapy has both positive and negative effects (Bergin 1963). Global estimates for the rate of relevant side effects (SEs) across different psychotherapies range from 5 to 20% of patients (Barlow 2010; Berk and Parker 2009; Lilienfeld 2007; Linden and Strauß 2013; Parker et al. 2013). In an online survey, 13% of 1504 patients reported "burdens caused by therapy" (Leitner et al. 2013). If psychotherapists are asked about their own treatments, 20–40% remember harmful effects (Buckley et al. 1981; Macaskill 1992). The most frequent nominations were psychological distress (29%) and marital or family stress (13%), although some therapists thought this to be necessary for effective treatment. Other examples of burdens caused by psychotherapy were patients feeling overwhelmed or being afraid of the therapist, undermining of self-efficacy, deterioration of symptoms, emergence of new symptoms, suicidality, treatment failure, occupational problems, stigmatization, strains in personal relationships, or changes in the social network (Hoffmann et al. 2008).

Over the years, therapists and researchers have repeatedly pointed to the importance of side effects in psychotherapy, though little empirical data are available (Jonsson et al. 2014). So far, there is no generally accepted definition or assessment methodology for psychotherapy SEs. There is the problem of differentiation between SEs on one side and consequences of inappropriate or unethical behavior by the therapists on the other. Also, it must be assumed that different treatments, different types of patients, different types of patient therapist dyads, and even therapist training can have their own side effects, so that detailed data are needed to guide clinical practice (Henry et al. 1993). Finally, the recognition of SEs poses special problems as therapists can have difficulties accepting and recognizing negative consequences of their work. Often they do not expect and therefore not recognize them (Hannan et al. 2005; Hatfield et al. 2010; Lambert 2010).

The assessment of side effects, be it in psychotherapy or pharmacotherapy, follows a stepwise process, as shown in Fig. 1 (Linden 2013; Linden and Schermuly-Haupt 2014). It starts with systematically recording unwanted events (UEs), which occur parallel to treatment and which may be related to the ongoing therapy or not (e.g. suicide). In a second step a professional judgment is needed to determine

✉ Michael Linden
michael.linden@charite.de

[1] Research Group Psychosomatic Rehabilitation at the Charité University Medicine Berlin, CBF Hs II, E01, Hindenburgdamm 30, 12200 Berlin, Germany

[2] Institute for Behavior Therapy Berlin, Berlin, Germany

[3] Duke-National University of Singapore, Singapore, Singapore

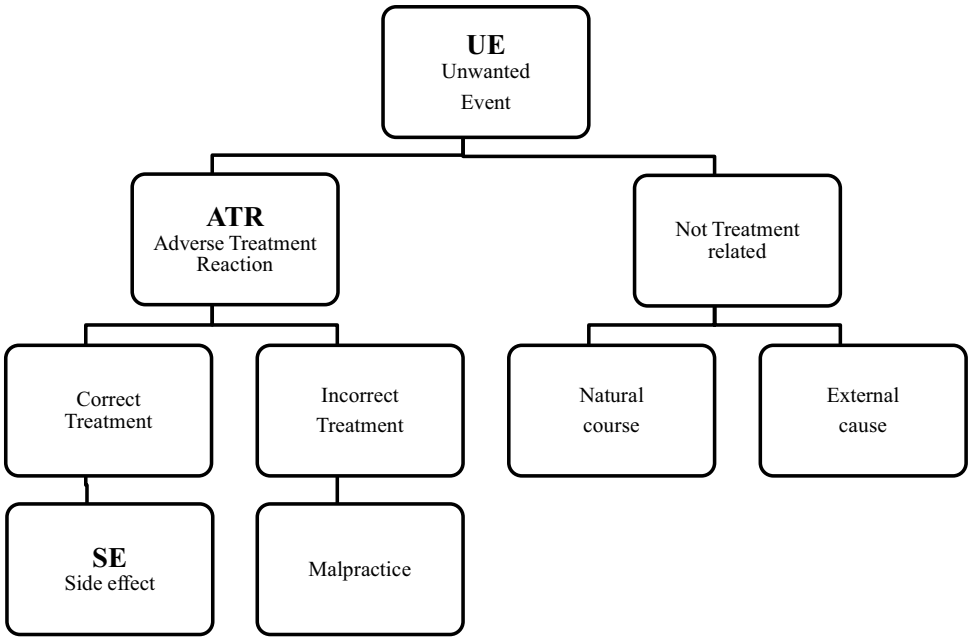

Fig. 1 Differentiation of unwanted treatment events

whether these UEs are adverse treatment reactions (ATRs), i.e. related to treatment (has the suicide been caused by treatment procedures?). Thirdly, it must be determined whether they are caused by treatment according to professional standards (were the treatment procedures which caused suicide following professional standards?). There may be unwanted events during the course of therapy that just happen by coincidence and there may be some unwanted events that are the result of malpractice or even therapeutic unethical behavior. None of those can be considered as side effects. This discrimination helps therapists to accept that their work can have not only beneficial but also negative consequences. The distinction between UEs and ATRs makes clear that SEs are often an unavoidable part of well-delivered treatment. By expecting SEs, we can plan for their management and mitigation.

As the incidence, nature, severity and duration of SE differ across different interventions, conditions and settings, specific data are needed (Lambert and Ogles 2004; Parker et al. 2013; Tarrier et al. 1999). The present study focuses on cognitive behavior therapy (CBT). For this treatment approach it has been reported that some patients with post-traumatic stress disorder (PTSD) or anxiety disorders can exhibit symptom worsening after imaginary exposure (Foa et al. 2002; Mayou et al. 2000; Scheeringa et al. 2011). In PTSD patients with hyper-arousal at baseline, emotional debriefing (meaning single sessions of psychological assistance shortly after the traumatic event where participants are encouraged to talk about detailed aspects of the events and their emotional reactions and thoughts) has resulted in significantly more PTSD symptoms than educational debriefing (single sessions with the focus on information on stress and symptoms) or no debriefing (Rose et al. 2002; Sijbrandij et al. 2006). Treatment of women with agoraphobia resulted in psychological symptoms in their husbands (Hafner 1984; Milton and Hafner 1979). Patients with generalized anxiety disorder treated with behavior therapy that focused on life problems instead of on worrying showed less remission than patients who had attended unspecific group sessions (Fisher and Durham 1990). SEs have also been reported with token economy treatment for substance abuse (Rosen et al. 2010), group psychotherapies (Roback 2000) and psychotherapies for patients with a history of sexual abuse (Brainerd and Reyna 2005; Lambert and Ogles 2004).

Given this background, the present study was designed to assess the nature, frequency, severity and duration of both UEs and SEs in outpatients who were receiving CBT. Such knowledge about the routine SE profile of treatments is needed to inform and safeguard patients. This paper uses a conceptualization and in-depth assessment of SEs that recognizes the distinction between true side effects, malpractice, clinical deterioration and other causes of negative outcomes (Linden 2013). By including a large number of therapists and outpatients, the SE profile of CBT under routine treatment conditions is described (i.e., nature, frequency, severity and duration of SEs). This is what therapists should know about when informing their patients about the upcoming merits and risks of treatment.

Method

Setting and Therapists

A convenience sample of 100 psychotherapists from three outpatient clinics for CBT were asked to participate in an interview on UEs and to report about one current patient who had received at least ten sessions of consecutive CBT, as this duration is estimated to be long enough for therapeutic effects and SEs to manifest (Anderson and Lambert 2001; Harnett et al. 2010). The mean age of therapists was 32.2 (SD = 4.7) years, 78% were female, and 96% had a university degree in psychology. Therapists had an average of 5.1 (SD = 3.4) years of professional experience.

All therapists were specifically trained in "cognitive behavior therapy (CBT)". According to German law and health regulations this is specified by an official expert panel, as basis for training and reimbursement. CBT is based on behavioral analysis and refers to classical learning theory, coping and social competency, dysfunctional cognitions, and emotional regulation. All treatments in this study were reimbursed by the general health insurance. To ensure treatment integrity, therapists were supervised after every fourth treatment session using audio or video recordings that are routine in the clinics to ensure that treatments follow accepted clinical standards. The recordings were not available for research purposes. On average, therapists had undergone 7.1 (SD = 4.0) sessions of supervision in the course of the treatment about which they reported.

As can be expected in a group of closely supervised therapists, no indicators for gross violations of therapeutic rules or inappropriate therapeutic behavior could be found. Therefore, all ATRs in this report are subsequently called SEs.

The institutional boards of the clinics reviewed and accepted the study protocol. After informing the therapists about the study, written informed consent was obtained. Patients were not personally involved, and data that could identify individuals were not recorded.

Interview

A semi-standardized interview[1] was conducted by the first author, a state licensed clinical psychologist with comprehensive training as a cognitive behavior therapist. Therapists were asked to give at least six examples of potential side effects of psychotherapy that they had experienced or could imagine. Thereafter, they were asked to report on their last treatment case of that day or the last few days in which the patient had attended at least 10 sessions. This minimum number of sessions was required to make sure that the treatment had a chance to show some effect. Sociodemographic indicators and clinical and treatment data were collected, including course of the illness, diagnoses, medication, number of sessions, etc.

Therapists were asked to rate the patient's present state of illness on the Fischer Symptom Checklist (FSCL), an observer rating scale with 41 Items, covering common psychiatric symptoms like mood impairment, anxiety, sleep, thought process and content, or social activity. The Items are scored on a 4 point scale from "not at all" to "severe". The FSCL has shown good reliability and validity and has been used in many clinical trials (Fischer-Cornelssen and Berchier 1982).

To systematically screen for SEs, the interview then followed the Unwanted Events-Adverse Treatment Reactions Checklist (UE-ATR Checklist; Linden 2013) that lists 17 domains of possible UEs (Fig. 2) including, for example, emergence of new symptoms or changes in family relations.

For each UE, it was determined whether the event was an ATR by asking the therapists to describe the UE in detail and explain what might have caused the event in his or her view. The final rating on the "degree of relatedness to psychotherapy" was rendered by the therapist using the UE-ATR Checklist by taking into account the type of UE, the relatedness to the psychotherapeutic processes, or its course and timing. "Relatedness to therapy" was classified as (1) "definitely unrelated", (2) "rather unrelated", (3) "rather related", (4) "most probably related", or (5) "definitely related". Therapists were then asked for information on the "duration" and "severity" of that specific UE. Categories of "duration" were "hours", "days", "weeks", "months", and "persistent". "Severity" was rated on a 5-point Likert-scale with respect to the impact on the patient from (1) "mild, no negative consequences" to (5) "extremely severe".

The UE-ATR Checklist is an instrument for event sampling in combination with a coding scheme. This is so far the only instrument which explicitly discriminates between unwanted events (UE), adverse treatment reactions (ATR) and side effects (SE), and which systematically guides the screening process. Reliability has been tested by having 623 reports of therapists rated by two raters using the 17 domains listed in the UE-ATR Checklist as categories. Cohens Kappa was 0.75 ($p = .000$), which is considered to be good agreement (Wirtz and Caspar 2002). Of the reports, 92% could be assigned to the domains of the checklist (8% were rated as "others"). Furthermore, we tested interrater agreement for the two scales "severity" and "relation to therapy" by comparing therapist ratings with the interviewer judgements. Unadjusted intra-class correlations (Shrout and Fleiss 1979) were 0.68 ($p < .001$) for severity and 0.84 ($p < .001$) for relation.

[1] The interview is available from michael.linden@charite.de.

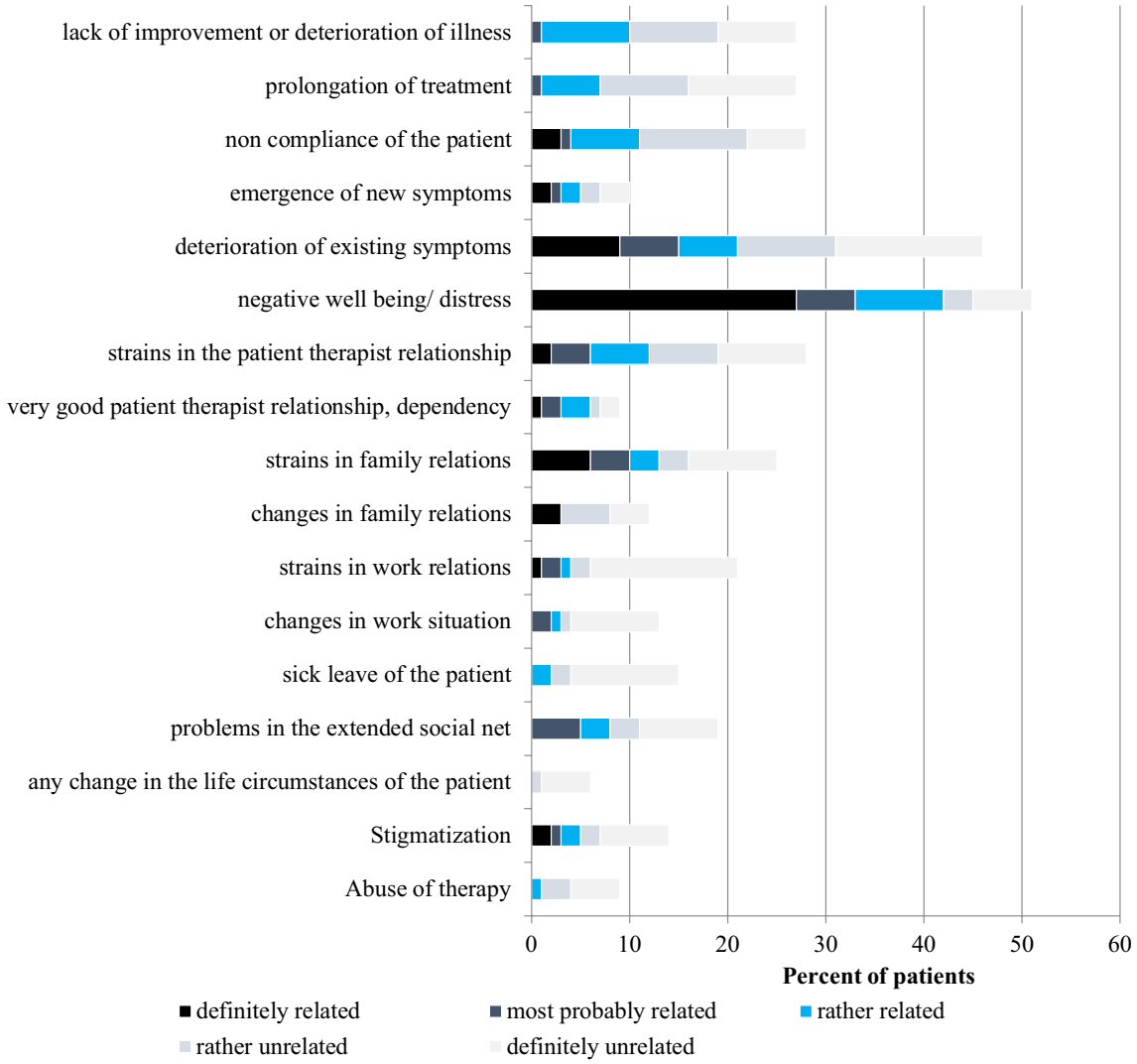

Fig. 2 Percentage of patients with different unwanted events and their relation to therapy

Results

Treatment Cases

Each of the 100 therapists reported on one treatment case (a total of 100 patients). Of the patients, 51% were female, the average age was 38.0 (SD = 11.3) years (range 19–68 years), 53% were married or living in a relationship, and 16% had a university degree.

The mean FSCL total score was 25.1 (SD = 14.2), indicating a degree of psychopathology in the mild to moderate range. The most frequent clinical diagnoses based on *ICD-10* criteria (Dilling 2011) were major depressive episode/recurrent depression (36%), anxiety disorders (16%), and personality disorders (15%). Forty-three percent of patients had been in an inpatient psychiatric hospital at least once, 33% had at least one prior period of psychotherapy, and 42% were receiving some psychotropic medication while CBT was being delivered.

At the time of the interview, patients had completed 10–100 sessions of CBT with a mean of 28.5 (SD = 16.8) sessions over 10.6 (SD = 7.0) months. Prior to conducting the UE-ATR interview, each therapist was asked to make a global judgment on the course of treatment. Five percent expected full remission and 59% expected a good response at the end of treatment.

Unwanted Events (UEs)

When asked for a spontaneous report, 74% of therapists were not aware of any UE or SE in the treatment of their patients. While answering the structured interview, which systematically evokes domains of UEs, therapists reported a total of 372 UEs (average of 3.7 (SD = 2.0) UEs per patient).

Table 1 Frequencies of "severity" and "duration" for unwanted events and side effects

Severity	Mild	Moderate	Severe	Very severe	Extremely severe
Unwanted events (n)	102	142	107	20	1
Percent (n = 372) (%)	27.4	38.2	28.8	5.4	0.3
% of patients	69	94	70	18	1
Side effects (n)	8	26	20	3	–
Percent (n = 57) (%)	14.0	45.6	35.1	5.3	–
% of patients	9	25	18	3	0
Duration[a]	Hours	Days	Weeks	Months	Persisting
Unwanted events (n)	59	52	97	82	62
Percent (n = 372) (%)	15.9	14.0	26.1	22.0	16.7
% of patients	51	40	62	47	40
Side effects (n)	18	18	12	3	5
Percent (n = 57) (%)	31.6	31.6	21.1	5.3	8.8
% of patients	19	17	0	3	5

[a]5.4% missing values in duration of unwanted events; 2 1.8% missing values in duration of side effects

Only for two of the 100 patients no UEs were reported. Three patients had a maximum of 9 UEs. Figure 2 shows the frequencies of the UEs for each of the 17 domains of the UE-ATR Checklist. Most frequent UEs were: "negative well-being/distress" (53% of patients) and "deterioration of existing symptoms" (46% of patients). Least frequent were "change in life circumstances" (6% of patients) and "emergence of new symptoms" (10% of patients). The mean "severity" was 2.13 (SD = 0.89) (i.e., 65.6% of UEs were mild or moderate). Regarding the "duration" of the UEs, 15.9% lasted only for hours, 14.0% for days, 26.1% for weeks, 22.0% for months, and 16.7% were expected to be persistent (Table 1).

Side Effects (SEs)

Of the 372 UEs 35.8% were, according to professional judgement of the interviewer, rated as "definitely unrelated" to treatment, 19.9% as "rather unrelated", 16.4% as "rather related", 9.4% as "most probably related" and 15.3% as "definitely related" to treatment. All events that were rated as "definitely related" to treatment were considered to be SEs.

Of the 100 patients, 43% suffered from at least one SE (average 0.57 (SD = 0.81) SEs per patient). Thirty-one patients had only one SE reported, while one patient had the maximum of 4 SEs reported. Figure 2 shows the frequencies of SEs for each of the 17 domains of the UE-ATR Checklist and Table 1 shows their "severity" and "duration".

The most frequent SEs were "negative wellbeing/distress" (7.2% of all UEs, 27% of patients), "deterioration of symptoms" (2.4% of UEs, 9% of patients) and "strains in family relations" (1.6% of UEs, 6% of patients). Mean severity was 2.32 (SD = 0.78). Fourteen percent of the SEs were rated as mild (in 9% of patients), 45.6% as moderate (in 25% of patients), 35.1% as severe (in 18% of patients) and 5.3% as very severe (in 3% of patients). No SEs were rated as extremely severe. Regarding duration, 31.6% of SEs lasted only for hours, 31.6% for days, 21.1% for weeks, 5.3% for months and 8.8% were expected to be persistent (Table 1).

To provide a better clinical understanding of the nature of UEs and SEs, Table 2 gives some qualitative reports of what is meant by the different dimensions of the UE-ATR Checklist. Examples describe severe and persisting side effects such as suicidality, breakups, negative feedback from family members, withdrawal from relatives, feelings of shame and guilt, or intensive crying and emotional disturbance during sessions.

The longer the treatment lasted, the more UEs, and to a lesser extent SEs, were reported. The correlation between the number of CBT sessions and the number of UEs was $r = .55$ ($p < .001$), and the number of SEs $r = .30$ ($p < .01$). There was no significant correlation between the FSCL score and the number of UEs ($r = .14$, $p > .05$), but there was a modest negative correlation between the FSCL score and the number of SEs ($r = -.22$, $p < .05$). In this respect, there was a medium-sized significant difference in regard to symptom severity as measured with the FSCL [$t(98) = 3.03$, $p < .01$, $d = .65$], with patients with SEs showing less symptom severity than patients without SEs.

To explore whether one could predict which patients would develop SEs, a binary logistic regression analysis was calculated with patient and therapist characteristics as predictors and group (patients with or without SEs) as a dichotomous independent variable. The set of potential predictor variables focused on variables found to impact outcome in other studies such as severity of disease, experience

Table 2 Examples of severe or very severe and persistent side effects

UE-ATR domain	Patient characteristics	Sessions	Narrative description	Severity	Duration
Deterioration of existing symptoms	45 year-old male with phobias, panic disorder, trauma and substance abuse in history	53	Suicidality after intense exposure	Severe	Days
	37 year-old male with narcissistic personality disorder and depression	78	Increased feelings of hopelessness, decline of happiness, increasing social isolation after confrontation with diagnosis and social costs	Severe	Weeks
	33 year-old female with generalized anxiety disorder and personality disorders	75	Increased anxiety and insecurity, emotional outbursts after reflecting her problems, awareness made her feel worse	Severe	Days
	27 year-old female with personality disorders, depression, social phobia and obsessive–compulsive disorder	42	Felt destabilized right at the beginning of therapy, crying and ruminating more often when talking about her problems	Severe	Weeks
	35 year-old female with post-traumatic stress disorder and depression	100	Increased symptoms of dissociation, dissociation in the domestic environment, which had not happened before, patient felt stressed by talking about what happened, insufficient coping strategies and feelings of overburden	Severe	Weeks
Negative wellbeing, distress	47 year-old male with alcohol dependence	18	Felt under pressure and stressed by the question of his therapist, felt angry and tense during sessions	Severe	Hours
	37 year-old male with narcissistic personality disorder and depression	78	Feelings of surprise, shame and guilt after video feedback and discussion of degrading offenses against his therapist	Severe	Days
	39 year-old male with agoraphobia	26	Intensive crying during session when talking about family conflicts	Severe	Days
	35 year-old female patient with depression and agoraphobia	12	While preparing confrontation therapy sessions the patient felt very uneasy and anxious, but did not dare to talk about it in the beginning, confrontation therapy had to be postponed	Severe	Days
	26 year-old male with recurrent depression	30	Felt very uncomfortable during role plays, thought them embarrassing	Severe	Hours
	40 year-old female with recurrent depression and phobia	30	Burdened by exposure therapy, intense fear, blushing, trembling, felt forced by her therapist	Severe	Hours
	68 year-old female with recurrent depression	24	After role play intensive crying, agitated and desolate at the end of session, could not be contained because of time, refused further role plays	Severe	Weeks
	50 year-old female with recurrent depression, somatization disorder and with a history of post-traumatic stress disorder	21	Negative emotions during meditation and relaxation exercises, feelings of weariness of life through unhappy memories	Severe	Days
	33 year-old female with social phobia and adjustment difficulties after cancer diagnosis	18	Discomposed just by talking to her therapist, blushing, nearly crying during sessions, rigid posture, feels so exhausted that sessions must be planned in the afternoons	Severe	Months
	24 year-old female with depression	40	Some issues are very hard for her to discuss, during a family session she felt rejected and criticizes, she cried and blushed and stammered, made her realize that the relationship to her father is not good	Severe	Hours
Emergence of new symptoms	33 year-old male with pathological gambling	46	After talking about biographical background the patient experienced strong feelings of guilt and shame towards his mother, that he had never felt before	Very severe	Days
Very good patient therapist relationship, dependency	45 year-old male with phobias, panic disorder, trauma and substance abuse in history		Patient will have problems to go on without therapy, relationship to therapist is grown and strong, patient finds therapy supportive and is afraid to struggle on his own, may need help in the future as well	Mild	Persistent

Table 2 (continued)

UE-ATR domain	Patient characteristics	Sessions	Narrative description	Severity	Duration
Strains in family relations	50 year-old female with obsessive–compulsive disorder and agoraphobia	16	Concerned by feedback from family members, who find her irritated and behaving unusual when she tried to deal with negative emotions	Severe	Weeks
	43 year-old female with dysthymia and panic disorder	23	Concerned by feedback of her husband, who finds her more individuating, taking more time for herself and less caressing	Severe	Weeks
	27 year-old female with personality disorders, depression, social phobia and obsessive–compulsive disorder	42	Tries to individuate from her family, in order to work on her social anxiety she tries to distance from her mother, agreed not to talk to her every day and to not see her three times a week, patient feels guilty and cries sometimes	Severe	Hours
	54 year-old male with depression and pedophilia	35	After the discussion about how social support prohibits offenses, outed himself with his wife, confessed possession of illegal material, his wife was shocked, the couple discussed separation	Severe	Persistent
	33 year-old female with social phobia and adjustment difficulties after cancer diagnosis	18	Critical view on her parents, questions their way, relates them to her problems	Very severe	Persistent
Changes in family relations	27 year-old female with personality disorders, depression, social phobia and obsessive–compulsive disorder	42	Broke up with her boyfriend after discussing the disadvantages of an exploitive relationship in therapy, felt very sad about it but hoped for reward in the future	Very severe	Weeks
	35 year-old female with post-traumatic stress disorder and depression	100	Patient decided to end the relationship to her mother after years of problems and conflicts, that was her wish when starting therapy and what she had planned, but she needed support and was sad when she did it	Mild	Persistent
Non-compliance of the patient	51 year-old female with dysthymia and obsessive–compulsive personality disorder	12	Neglects homework, refuses exercises, criticizes interventions, anxious that her therapist might be into Buddhistic religion, complaining about poor match with her therapist, speculated about drop out of therapy	Severe	Weeks
	42 year-old female with generalized anxiety disorder and somatization disorder	26	Patient is not doing her homework or only half of it, she is not working for her goals, it seems we're not working on something important for her, therapy goes in the wrong direction	Moderate	Persistent

Table 3 Logistic regression analysis of group membership (patients with or without SEs)

Independent variable	b	SE	Wald	Probability	Odds
Number of sessions	.018	.014	1.680	.195	1.018
Number of diagnoses	.208	.252	.683	.409	1.232
Sum FSCL	− .054	.019	8.039	.005	.948 1.041
Age of patient	.040	.021	3.603	.058	1.041
Age of therapist	.001	.061	.000	.990	1.001
Years of experience of the therapist	− .005	.086	.003	.955	.995
Gender of therapist	.226	.599	.143	.706	1.254
Gender of patient	.423	.472	.805	.370	1.527
Match of gender	.261	0.303	0.740	.390	1.298
Model χ^2 =	18.417	$p < .05$			
Pseudo R^2 =	.226				
n =	100				

The dependent variable in this analysis is group coded so that 0 = patients without SEs and 1 = patients with SEs

*FSCL*Fischer Symptom Checklist; *SEs* Side effects

of therapist or match of gender (Crits-Christoph et al. 1991; Hamilton and Dobson 2002; Lambert 2011; Leitner et al. 2013; Sotsky et al. 1991). A test of the model against a constant-only model was statistically significant, indicating that the predictors distinguished between patients with and without SEs (chi square = 18.417, $p < .05$ with df = 9). Table 3 shows the variables and their coefficients. With a Nagelkerke´s R^2 of .23, the effect size was very small, resulting in a successful prediction of group in only 63% of cases (instead of 57%).

Further analyses showed that the FSCL score was the only variable that could by itself significantly predict the classification of patients with or without SEs. Patients with a lower FSCL score tended to be patients with SEs. No significant prediction was found with respect to number of sessions, number of diagnoses, age of patient, age of therapist, years of experience of the therapist, gender of therapist, gender of patient or match of gender.

Discussion

To our knowledge, this is the first study to attempt to separate unwanted events (UE), adverse treatment reactions (ATR), and side effects (SE) in a large sample of psychotherapy patients. It demonstrates that one can make this distinction, and it confirms the concept and the tool. Looking for unwanted events independent of any early judgment as to causality helps to overcome the non-recognition bias of therapists. We found that 74% of therapists were not aware of any side effects in their treatment before the systematic evaluation began. The present findings suggest that the use of structured assessment methods like the UE-ATR Checklist can improve the recognition of side effects.

A separate judgment about whether an unwanted event is an adverse treatment reaction helps to avoid incorrectly inflating the rate of side effects. In our sample, 98% of patients had experienced some unwanted events, but only 43% were found to suffer from adverse treatment reactions or side effects. It is an open question who is in the best position to report side effects: the patient, the therapist or other professionals. The patient is positioned to identify negative feelings or outcomes, but does not have the professional knowledge or perspective to say whether these are side effects. The therapist can make a judgment about side effects, especially those of clinical relevance, but may have a recognition or judgement bias. Our method of having an independent professional interview the therapist is perhaps best, though it might be improved if a separate interview of the patient had been possible. Future research should combine both perspectives.

We found 372 unwanted events in 98 of 100 patients. Unwanted events that were "definitely related" to treatment were found in 43% of cases. This is a conservative estimate of side effects. Had we included those unwanted events that were rated as "most probably" and "rather related" to the treatment, 63% of cases would have been classified as suffering from side effects. This rate of side effects is somewhat higher than in most other studies (Barlow 2010; Berk and Parker 2009; Leitner et al. 2013; Lilienfeld 2007; Linden and Strauß 2013; Parker et al. 2013), which can be explained by the use of the structured UE-ATR procedure.

A point of discussion is whether ordinary reactions to CBT which may be indispensable for the success of treatment, such as distress during exposure treatment should be called "side effects". We argue that they are side effects although they may be unavoidable, justified, or even needed and intended. If there were an equally effective treatment that did not promote anxiety in the patient, the present form of exposure treatment would become unethical as it is a burden to the patient. It is important to make a distinction between unavoidable and possibly even intended negative effects on one side, and desired ones on the other. This is a general rule in medicine, like in surgery, where in earlier times it may have been necessary, unavoidable and intended to remove a breast to fight cancer. But still, it was not desired and so surgeons developed new treatments without this burden to the patient. Unavoidable and intended negative effects are burdens to the patient and therefore undesired. To acknowledge this is important for the improvement of psychotherapy in the individual case as this can help to avoid unnecessary distress for the patient and select the best treatment option. It is also important for the development of

treatment alternatives that are better tolerated, similar to the development of strategies in surgery which allow to keep a breast.

Most of the side effects in this study were rated as mild or moderate (59.6%) and transient (89.6%). However, more than 40% of side effects were rated as severe (i.e., countermeasures are necessary) or very severe (i.e. enduring negative consequences) and 8.8% as persistent. Psychotherapy is not harmless.

Apart from prevalence, our data also show that side effects are very multi-facetted in regard to type and content. Deterioration of symptoms can be caused by increased feelings of hopelessness or despair when looking at existing problems. Feelings of dependency and lack of self-efficacy can be caused by a very close and supportive therapeutic relation. Break up of relations with partners or parents, or avoidance of work can result from explicit or implicit incrimination of living situations. Problems in therapy cooperation can follow feelings of shame after not doing homework assignments. Those reports have to be interpreted carefully and may not be applicable to other samples and settings. Still, these examples show that regular and appropriate therapeutic interventions can have negative consequences, which may even be enduring like quitting a job, getting into trouble with close persons, undermining of self-confidence, or aggravation of problem perception.

The positive correlation between number of sessions and number of unwanted events and side effects gives validity to the assessment, as more unwanted events should emerge over the course of time. An interesting and somewhat unexpected finding was that patients with side effects had lower scores on the FSCL, which suggests that side effects are not an expression of severity of illness. It can be assumed that mild side effects are better recognized in less severe cases. The FSCL score was the only significant predictor of side effects. Other variables, such as number of diagnoses or experience of therapists, did not contribute to the classification of patients with or without SEs. This finding may be due to a relatively small sample size concerning the number of variables in the model. Further research is needed to identify patients at risk for developing side effects and to replicate the association between symptom severity and side effects.

Our study has several implications. First, therapists are often not aware of side effects in their treatment as they primarily focus on positive changes (Tomba et al. 2017). An awareness and recognition of unwanted events and side effects in all therapies will benefit patients, improve therapy or reduce attrition, analogous to the benefit of measurement-based monitoring of treatment progress (Castonguay et al. 2010). It is probably not realistic to expect that therapists use checklists in their daily practice, as is similar with other clinical scales. Still, it may be helpful to know the dimensions of the UE-ATR Checklist, as this can help therapists to not forget important aspects in the clinical assessment. The scale could and should be used in the training of therapists, so that they learn what a side effect is, how to recognize side effects, and how to ask patients about them (Nolan et al. 2004). The therapist bias of not seeing side effects must be overcome during training. Finally, the checklist may be used in research and especially in clinical trials, where there is often a lack of focus on side effects (Jonsson et al. 2014). The research question is not only to describe side effects but also to develop strategies on what to do should they occur, to identify patients at risk, and to develop therapies without certain side effects (Creed et al. 2014). Finally, the data describe the SE profile of CBT under routine clinical conditions in a heterogenous patient population. This is what can occur with some probability in patients undergoing this mode of treatment. Therapist can use this knowledge when informing their patients about the treatment and also for risk monitoring in the course of treatment.

Limitations

The identification of SEs in this report depends on our judgment as to the appropriateness of the therapy and the unobserved but presumably ethical behavior of the therapists. We can say that all therapists were trained in CBT, but that does not necessarily verify treatment integrity and competency (Lambert 2011). However, since the therapists were closely supervised, gross malpractice is very unlikely. The cross-sectional study design has limitations, including that of recall. Prospective, longitudinal data would have allowed a more fine-grained analysis and potentially a more accurate assessment of the events during the course of therapy. Another limitation is that we do not have data on the relationship between SEs and the outcome of therapy. This report focused only on the nature, frequency, severity and duration of UEs, but not their consequences. Since about half of the patients were receiving additional psychotropic medication, we cannot attribute the presence of SEs to the psychotherapy alone. The instructions for the therapists were very clear as to report unwanted events and subsequently rate whether those events were caused by psychotherapy and psychotherapy only. But since those ratings can be flawed some SEs may have been related to the pharmacotherapy itself or an interaction between the pharmacotherapy and CBT. To use therapists ratings only, limits the generalizability of the results per se. It may be possible that patients would consider the same events noted by therapists as wanted and beneficial instead of unwanted (e.g. a break up) or not caused by psychotherapy but some aspects outside of therapy (e.g. job loss). A combination of the patient, therapist and a separate interviewer may indeed be the best way to approach this problem (Smith et al. 2003). Since we investigated a

heterogeneous sample of patients with different conditions treated by a range of therapists with a range of skills and experiences and different methods of CBT, the results could be different in different samples of patients, conditions, therapists and even types of therapy.

Conclusions

A structured assessment, using the UE-ATR-Checklist allows to measure side effects of psychotherapy (SE), and discriminate between SEs and other unwanted events (UEs). SEs include symptoms, but also additional problems in many areas of life. They are common even in the course of rigorously supervised CBT. The data describe the SE profile of CBT under routine clinical conditions in a heterogenous patient population. Such information is needed in clinical practice, training, and research.

Acknowledgements The authors thank the Institute for Behavior Therapy Berlin, the Centre for Psychotherapy at the Humboldt-University Berlin and the Institute for Behavior Therapy Brandenburg for participating in the study. We would also like to acknowledge the editorial assistance of Jon Kilner, MS, MA (Pittsburgh, PA).

Compliance with Ethical Standards

Conflict of Interest Marie-Luise Schermuly-Haupt, Michael Linden and A. John Rush declare that they have no conflict of interest.

Human and Animal Rights All procedures performed in studies involving human participants were in accordance with the ethical standards of the institutional and/or national research committee and with the 1964 Helsinki declaration and its later amendments or comparable ethical standards.

Informed Consent Informed consent was obtained from all individual participants included in the study.

References

Anderson, E. M., & Lambert, M. J. (2001). A survival analysis of clinically significant change in outpatient psychotherapy. *Journal of Clinical Psychology, 57*(7), 875–888.

Barlow, D. H. (2010). Negative Effects from psychological treatments. A perspective. *American Psychologist, 65*(1), 13–20. https://doi.org/10.1037/a0015643.

Bergin, A. E. (1963). The effects of psychotherapy: Negative results revisited. *Journal of Consulting Psychology, 10*, 244–250.

Berk, M., & Parker, G. (2009). The elephant on the couch: Side effects of psychotherapy. *Australian & New Zealand Journal of Psychiatry, 43*(9), 787–794. https://doi.org/10.1080/00048670903107559.

Brainerd, C. J., & Reyna, V. F. (2005). *The science of false memory*. New York: Oxford University Press.

Buckley, P., Karasu, T. B., & Charles, E. (1981). Psychotherapists view their personal therapy. *Psychotherapy: Theory, Research & Practice, 18*(3), 299–305. https://doi.org/10.1037/h0088377.

Castonguay, L. G., Boswell, J. F., Constantino, M. J., Goldfried, M. R., & Hill, C. E. (2010). Training implications of harmful effects of psychological treatments. *American Psychologist, 65*(1), 34–49. https://doi.org/10.1037/a0017330.

Creed, T. A., Wolk, C. B., Feinberg, B., Evans, A. C., & Beck, A. T. (2014). Beyond the Label: Relationship Between Community Therapists' Self-Report of a Cognitive Behavioral Therapy Orientation and Observed Skills. *Administration and Policy in Mental Health and Mental Health Services Research, 43*(1), 1–8. https://doi.org/10.1007/s10488-014-0618-5.

Crits-Christoph, P., Baranackie, K., Kurcias, J., Beck, A., Carroll, K., Perry, K., Luborsky, L., McLellan, A., Woody, G., Thompson, L., Gallagher, D., & Zitrin, C. (1991). Meta-analysis of therapist effects in psychotherapy outcome studies. *Psychotherapy Research, 1*(2), 81–91.

Dilling, H. (2011). *Internationale Klassifikation psychischer Störungen: ICD-10 Kapitel V (F); klinisch-diagnostische Leitlinien (8., überarb. Aufl. unter Berücksichtigung der Änderungen entsprechend ICD-10-GM 2011.)*. Bern: Huber.

Fischer-Cornelssen, K. A., & Berchier, P. (1982). Validität und Reliabilität einer Symptom-Checkliste (FSCL): Anwendung in der Psychogeriatrie. *Zeitschrift für Gerontologie und Geriatrie, 15*, 31–37.

Fisher, P. L., & Durham, R. C. (1990). Recovery rates in generalized anxiety disorder following psychological therapy: An analysis of clinically significant change in the STAI-T across outcome studies since 1990. *Psychological Medicine, 29*(6), 1425–1434.

Foa, E. B., Zoellner, L. A., Feeny, N. C., Hembree, E. A., & Alvarez-Conrad, J. (2002). Does imaginal exposure exacerbate PTSD symptoms? *Journal of Consulting Clinical Psychology, 70*(4), 1022–1028.

Hafner, R. J. (1984). The marital repercussions of behavior therapy for agoraphobia. *Psychotherapy: Theory, Research, Practice, Training, 21*(4), 530–542. https://doi.org/10.1037/h0085999.

Hamilton, K. E., & Dobson, K. S. (2002). Cognitive therapy of depression: Pretreatment patient predictors of outcome. *Clinical Psychology Review, 22*(6), 875–893.

Hannan, C., Lambert, M. J., Harmon, C., Nielsen, S. L., Smart, D. W., Shimokawa, K., & Sutton, S. W. (2005). A lab test and algorithms for identifying clients at risk for treatment failure. *Journal of Clinical Psychology, 61*(2), 155–163. https://doi.org/10.1002/jclp.20108.

Harnett, P., O´Donovan, A., & Lambert, M. J. (2010). The dose response relationship in psychotherapy: Implications for social policy. *Clinical Psychology, 14*(2), 39–44. https://doi.org/10.1080/13284207.2010.500309.

Hatfield, D., McCullough, L., Frantz, S. H., & Krieger, K. (2010). Do we know when our clients get worse? An investigation of therapists' ability to detect negative client change. *Clinical Psychology and Psychotherapy, 17*(1), 25–32. https://doi.org/10.1002/cpp.656.

Henry, W. P., Strupp, H. H., Butler, S. F., Schacht, T. E., & Binder, J. L. (1993). Effects of training in time-limited dynamic psychotherapy: Changes in therapist behavior. *Journal of consulting and clinical psychology, 61*(3), 434–440.

Hoffmann, S. O., Rudolf, G., & Strauß, B. (2008). Unerwünschte und schädliche Wirkungen von Psychotherapie. *Der Psychotherapeut, 53*, 4–16.

Jonsson, U., Alaie, I., Parling, T., & Arnberg, F. K. (2014). Reporting of harms in randomized controlled trials of psychological interventions for mental and behavioral disorders: A review of current practice. *Contemporary Clinical Trials, 38*(1), 1–8. https://doi.org/10.1016/j.cct.2014.02.005.

Lambert, M. J. (2010). Predicting negative treatment outcome: Methods and estimates of accuracy. In M. J. Lambert (Ed.), *Prevention of treatment failure: The use of measuring, monitoring, and*

feedback in clinical practice (pp. 83–105). Washington, DC: American Psychological Association.

Lambert, M. J. (2011). What have we learned about treatment failure in empirically supported treatments? Some suggestions for practice. *Cognitive and Behavioral Practice, 18*(3), 413–420. https://doi.org/10.1016/j.cbpra.2011.02.002.

Lambert, M. J., & Ogles, B. M. (2004). The efficacy and effectiveness of psychotherapy. In M. J. Lambert (Ed.), *Bergin & Garfield's handbook of psychotherapy and behavior change* (5th edn., pp. 139–193). New York: Wiley.

Leitner, A., Märtens, M., Koschier, A., Gerlich, K., Liegl, G., Hinterwallner, H., & Schnyder, U. (2013). Patients' perceptions of risky developments during psychotherapy. *Journal of Contemporary Psychotherapy, 43*(2), 95–105. https://doi.org/10.1007/s10879-012-9215-7.

Lilienfeld, S. O. (2007). Psychological treatments that cause harm. *Perspectives on Psychological Science, 2*(1), 53–70. https://doi.org/10.1111/j.1745-6916.2007.00029.x.

Linden, M. (2013). How to define, find and classify side effects in psychotherapy: From unwanted events to adverse treatment reactions. *Clinical Psychology and Psychotherapy, 20*(4), 286–296. https://doi.org/10.1002/cpp.1765.

Linden, M., & Schermuly-Haupt, M. L. (2014). Definition, assessment and rate of psychotherapy side effects. *World Psychiatry, 13*(3), 306–309. https://doi.org/10.1002/wps.20153.

Linden, M., & Strauß, B. (2013). *Risiken und Nebenwirkungen von Psychotherapie. Erfassung, Bewältigung, Risikovermeidung*. Berlin: Medizinisch-Wissenschaftliche Verlagsgesellschaft.

Macaskill, N. (1992). Psychotherapists-in-training evaluate their personal therapy: Results of a UK Survey. *British Journal of Psychotherapy, 9*(2), 133–138. https://doi.org/10.1111/j.1752-0118.1992.tb01211.x.

Mayou, R. A., Ehlers, A., & Hobbs, M. (2000). Psychological debriefing for road traffic accident victims: Three-year follow-up of a randomized controlled trial. *British Journal of Psychiatry, 176*(6), 589–593. https://doi.org/10.1192/bjp.176.6.589.

Milton, F., & Hafner, R. J. (1979). The outcome of behavior therapy for agoraphobia in relation to marital adjustment. *Archives of General Psychiatry, 36*(7), 807–811.

Nolan, S. A., Strassle, C. G., Roback, H. B., & Binder, J. L. (2004). Negative treatment effects in dyadic psychotherapy: A focus on prevention and intervention strategies. *Journal of Contemporary Psychotherapy, 34*(4), 311–330. https://doi.org/10.1007/s10879-004-2526-6.

Parker, G., Fletcher, K., Berk, M., & Paterson, A. (2013). Development of a measure quantifying adverse psychotherapeutic ingredients: The Experiences of Therapy Questionnaire (ETQ). *Psychiatry Research, 206*(2–3), 293–301. https://doi.org/10.1016/j.psychres.2012.11.026.

Roback, H. B. (2000). Adverse outcomes in group psychotherapy. Risk factors, prevention, and research directions. *Journal of Psychotherapy Practice Research, 9*(3), 113–122.

Rose, S. C., Bisson, J., Churchill, R., & Wessely, S. (2002). Psychological debriefing for preventing post traumatic stress disorder (PTSD). *The Cochrane database of systematic reviews, 2002*(2), 1–47. https://doi.org/10.1002/14651858.CD000560.

Rosen, M. I., Rounsaville, B. J., Ablondi, K., Black, A. C., & Rosenheck, R. A. (2010). Advisor-teller money manager (ATM) therapy for substance use disorders. *Psychiatry Services, 61*(7), 707–713. https://doi.org/10.1176/appi.ps.61.7.707.

Scheeringa, M. S., Weems, C. F., Cohen, J. A., Amaya-Jackson, L., & Guthrie, D. (2011). Trauma-focused cognitive-behavioral therapy for posttraumatic stress disorder in three-through six year-old children: A randomized clinical trial. *Journal of Child Psychology and Psychiatry, 52*(8), 853–860. https://doi.org/10.1111/j.1469-7610.2010.02354.x.

Shrout, P. E., & Fleiss, J. L. (1979). Intraclass correlations: Uses in assessing rater reliability. *Psychology Bulletin, 86*(2), 420–428.

Sijbrandij, M., Olff, M., Reitsma, J. B., Carlier, I. V. E., & Gersons, B. P. R. (2006). Emotional or educational debriefing after psychological trauma. Randomised controlled trial. *British Journal of Psychiatry, 189*, 150–155. https://doi.org/10.1192/bjp.bp.105.021121.

Smith, S. R., Hilsenroth, M. J., Baity, M. R., & Knowles, E. S. (2003). Assessment of patient and therapist perspectives of process: A revision of the Vanderbilt Psychotherapy Process Scale. *American Journal of Psychotherapy, 57*(2), 195–205.

Sotsky, S. M., Glass, D. R., Shea, M. T., Pilkonis, P. A., Collins, J. F., Elkin, I., Watkins, J. T., Imber, S. D., Leber, W. R., Moyer, J., & Oliveri, M. E. (1991). Patient predictors of response to psychotherapy and pharmacotherapy: Findings in the NIMH Treatment of Depression Collaborative Research Program. *American Journal of Psychiatry, 148*(8), 997–1008. https://doi.org/10.1176/ajp.148.8.997.

Tarrier, N., Pilgrim, H., Sommerfield, C., Faragher, B., Reynolds, M., Graham, E., & Barrowclough, C. (1999). A randomized trial of cognitive therapy and imaginal exposure in the treatment of chronic posttraumatic stress disorder. *Journal of Consulting and Clinical Psychology, 67*(1), 13–18.

Tomba, E., Tecuta, L., Schumann, R., & Balladini, D. (2017). Does psychological well-being change following treatment? An exploratory study on outpatients with eating disorders. *Comprehensive Psychiatry, 74*, 61–69. https://doi.org/10.1016/j.comppsych.2017.01.001.

Wirtz, M. A., & Caspar, F. (2002). *Beurteilerübereinstimmung und Beurteilerreliabilität: Methoden zur Bestimmung und Verbesserung der Zuverlässigkeit von Einschätzungen mittels Kategoriensystemen und Ratingskalen*. Göttingen: Hogrefe.