Understanding therapeutic change: The impact of early change, dropout and adherence on treatment outcome in internet interventions

Inauguraldissertation zur Erlangung der Doktorwürde (Dr. rer. nat.) im Fach Psychologie, Fachbereich I an der Universität Trier

Vorgelegt im
Oktober 2019

Von
Alice Arndt

Gutachter:
Prof. Dr. Wolfgang Lutz
Prof. Dr. Julian Rubel

Dissertationsort: Trier

DANKSAGUNG

Mein Dank gilt meinem Doktorvater Prof. Dr. Wolfgang Lutz, der mich immer in meiner fachlichen Entwicklung unterstützt und sein Vertrauen in mich gesetzt hat. Weiterhin war er es, der es mir ermöglicht hat an nationalen sowie internationalen Kongressen teilzunehmen, um meine Arbeit vorzustellen, meinen Horizont zu erweitern, die Arbeiten anderer Forschungsgruppen kennenzulernen und mich mit anderen Fachkollegen und Kolleginnen auszutauschen.

Darüber hinaus möchte ich besonders Prof. Dr. Julian Rubel danken, der mich vom Beginn meiner Promotion an immer wieder fachlich unterstützt hat, wenn ich Fragen hatte.

Außerdem möchte ich mich auch besonders bei Frau Kaitlyn Boyle für Ihre Unterstützung bedanken, die geduldig Korrektur gelesen hat und durch ihre Rückmeldungen dadurch die Qualität meiner wissenschaftlichen Arbeiten gesteigert hatte.

Ebenso gilt mein Dank allen anderen Kollegen, mit denen ich viele wichtige Momente bei der Entstehung dieser Arbeit geteilt habe und die mir immer wieder Mut gemacht haben. Ohne euch, eure Anregungen und euren Humor hätte ich nicht so viel Freude auf der Arbeit erlebt.

Weiterhin möchte ich vor allem meiner Familie und meinen Freunden danken, die mir immer beigestanden haben so gut sie konnten und ein offenes Ohr für mich hatten.

Table of contents

List of publications for the cumulative dissertation

1. Study I:

Lutz, W., Arndt, A., Rubel, J., Berger, T., Schröder, J., Späth, C., . . . Moritz, S. (2017). Defining and Predicting Patterns of Early Response in a Web-Based Intervention for Depression. *Journal of Medical Internet Research*, *19*(6), e206. https://doi.org/10.2196/jmir.7367

2. Study II:

Arndt, A., Lutz, W., Rubel, J., Berger, T., Meyer, B., Schröder, J., . . . Moritz, S. (2019). Identifying change-dropout patterns during an Internet-based intervention for depression by applying the Muthen-Roy model. *Cognitive Behaviour Therapy*, 1–19. https://doi.org/10.1080/16506073.2018.1556331

3. Study III:

Arndt, A., Rubel, J., Berger, T. & Lutz, W. Outpatient and self-referred participants: Adherence to treatment components and outcome in an internet intervention targeting anxiety disorders. Submitted *Cognitive Behaviour Therapy*, August 09th, 2019

1. Abstract

Internet interventions have gained popularity and the idea is to use them to increase the availability of psychological treatment. Research suggests that internet interventions are effective for a number of psychological disorders with effect sizes comparable to those found in face-to-face treatment. However, when provided as an add-on to treatment as usual, internet interventions do not seem to provide additional benefit. Furthermore, adherence and dropout rates vary greatly between studies, limiting the generalizability of the findings. This underlines the need to further investigate differences between internet interventions, participating patients, and their usage of interventions. A stronger focus on the processes of change seems necessary to better understand the varying findings regarding outcome, adherence and dropout in internet interventions. Thus, the aim of this dissertation was to investigate change processes in internet interventions and the factors that impact treatment response. This could help to identify important variables that should be considered in research on internet interventions as well as in clinical settings that make use of internet interventions.

Study I (Chapter 5) investigated early change patterns in participants of an internet intervention targeting depression. Data from 409 participants were analyzed using Growth Mixture Modeling. Specifically a piecewise model was applied to model change from screening to registration (pretreatment) and early change (registration to week four of treatment). Three early change patterns were identified; two were characterized by improvement and one by deterioration. The patterns were predictive of treatment outcome. The results therefore indicated that early change should be closely monitored in internet interventions, as early change may be an important indicator of treatment outcome.

Study II (Chapter 6) picked up on the idea of analyzing change patterns in internet interventions and extended it by using the Muthen-Roy model to identify change-dropout patterns. A sligthly bigger sample of the dataset from Study I was analyzed ($N = 483$). Four change-dropout patterns emerged; high risk of dropout was associated with rapid

improvement and deterioration. These findings indicate that clinicians should consider how dropout may depend on patient characteristics as well as symptom change, as dropout is associated with both deterioration and a good enough dosage of treatment.

Study III (Chapter 7) compared adherence and outcome in different participant groups and investigated the impact of adherence to treatment components on treatment outcome in an internet intervention targeting anxiety symptoms. 50 outpatient participants waiting for face-to-face treatment and 37 self-referred participants were compared regarding adherence to treatment components and outcome. In addition, outpatient participants were compared to a matched sample of outpatients, who had no access to the internet intervention during the waiting period. Adherence to treatment components was investigated as a predictor of treatment outcome. Results suggested that especially adherence may vary depending on participant group. Also using specific measures of adherence such as adherence to treatment components may be crucial to detect change mechanisms in internet interventions. Fostering adherence to treatment components in participants may increase the effectiveness of internet interventions.

Results of the three studies are discussed and general conclusions are drawn. Implications for future research as well as their utility for clinical practice and decision-making are presented.

2. Introduction

Research on internet interventions has grown rapidly in the last years (Andrews et al., 2018; Carlbring, Andersson, Cuijpers, Riper, & Hedman-Lagerlöf, 2018). While standard face-to-face psychological treatments like CBT are effective for depression and anxiety disorders, they are also cost intensive and are not broadly available (Andrews & Williams, 2014). In addition, developments such as the broad use of the internet in every-day life, and growing evidence for the effectiveness of internet interventions have led to an increase in the popularity of internet interventions as psychological treatments (Schuster, Berger, & Laireiter, 2018). However the success of implementation efforts as well as the degree of documentation of these efforts varies largely internationally (Andersson, Titov, Dear, Rozental, & Carlbring, 2019). Possible obstacles when implementing internet interventions include differing interests and attitudes of clients, clinicians, and other stakeholders (Andersson et al., 2019). Early attempts to apply internet interventions in general health care date back to the 1990s with the Interapy program in the Netherlands, a tinnitus clinic in Sweden and an intervention targeting depression called moodgym that was applied in Australia (Andersson et al., 2019). More recently, the Improving Access to Psychological Therapies (IAPT) program was introduced in England, which uses a stepped care approach including internet interventions (for details of the program, see Clark, 2018). Treatment provision follows the UK's National Institute for Health and Care Excellence clinical stepped care guidelines (National Collaborating Centre for Mental Health [UK], 2011). As step two of this approach, computerized CBT is recommended for depression, panic disorder as well as generalized anxiety disorder (Clark, 2018). As can be seen, in some countries, internet interventions are a very central part of routine care with many applications having been developed, especially in the Netherlands, Scandinavia, and Anglo-America (Schuster et al., 2018). However, in other countries such as Germany, internet interventions are still almost exclusively offered in research settings (Grünzig, Bengel, Göhner, & Krämer, 2019). It is therefore important to note that the mode of provision differs largely internationally. In addition, depending on the providers of internet interventions, their interests and capacities as well as the implementation context there are further differences to consider between particular interventions that are now shortly described.

1. Treatment

Possible differences refer to aspects such as therapeutic approach, length of treatment, treatment components, and their order. An attempt is often made to mirror face-to-face treatment: An assessment phase is followed by treatment allocation, and specific content (information and treatment components), which is assigned weekly (Carlbring et al., 2018). Exceptions to this typical procedure can be observed in transdiagnostic interventions, which are not necessarily based on an assessment phase.

2. Audience

Differences refer to aspects such as the targeted disorders and conditions as well as special groups of participants, e.g. youths or the elderly. Typically depression and anxiety disorders are treated (Carlbring et al., 2018).

3. Participant support

Differences refer to aspects such as provision of reminders as well as guidance and requirements of qualified professionals (Andersson et al., 2019). It is assumed that more support, including encouragement and the possibility to resolve difficulties, is a sign of more intensive treatment. Support is often provided weekly with a time for support varying between a few minutes, e.g. 10 minutes per week (Andersson, Rozental, Rück, & Carlbring, 2015) and up to four hours per week (Ebert et al., 2018).

4. Symptom monitoring

Differences refer to aspects such as assessments, timing of assessments and potential feedback of results. Progress is monitored in most internet interventions, however this is not true for all available internet interventions (Ebert & Baumeister, 2017).

5. Treatment availability

Differences refer to aspects such as route of access (e.g. part of routine care, source of referral, public or selective access) and costs of the intervention for participants.

6. Flexibility of treatment content and modalities

Differences refer to aspects such as the possibility of treatment adaptation according to a professional or participant choice. In many interventions, participants receive the same content (Ebert et al., 2018), however there are also treatments that offer more possibilities to adapt treatment (e.g. Deprexis, see Schuster et al., 2018).

7. Treatment context

Differences refer to aspects such as association with other interventions and continuity, e.g. stand-alone intervention, parallel use or "blended treatments", integration in stepped-care or post-treatment use to maintain treatment gains.

While there are some aspects in which internet interventions are more likely to differ than others (e.g. target groups, level of guidance, length of treatment, and context of treatment), there are other aspects that do not vary as much between interventions. For instance, there are only a few internet interventions that are based on a psychodynamic therapeutic approach, but a large number of internet interventions use CBT-based treatment components, such as thought protocols (Schuster et al., 2018). Generally, many CBT manuals for different disorders exist that can be easily integrated into internet interventions, which could explain the observed dominace of CBT based approaches.

Furthermore, there are many different internet interventions available, but most target a specific disorder and many are quite standardized with limited possibilities to adapt treatment to the participants' individual needs (Andrews & Williams, 2014; Schuster et al., 2018).

This already illustrates that internet interventions come with particular advantages and disadvantages for the parties involved in their development, implementation, provision and usage. Some of the most important perspectives to be mentioned in this context are:

1. Healthcare providers are interested in the broad and quick availability of psychological treatment that is associated with low costs, while providing a good treatment quality that reduces the risk of future treatment costs.
2. On the one hand, therapists may have a high case load and may not always be able to guarantee quick availability of treatment. However, depending on their treatment approach and their understanding of therapeutic change, they may prefer direct contact when delivering treatment.
3. Participants may profit from advantages such as the easy access and quick availability of internet interventions, a high level of flexibility regarding frequency and timing of usage, as well as anonymity (Andrews et al., 2018). At the same time, depending on participants' ability and motivation to pursue treatment, they may also miss direct contact during treatment or a more individualized treatment.

Taking these different perspectives into account, the implementation of internet interventions into health care is associated with opportunities as well as potential risks. Opportunities of internet interventions are linked to cost reduction and the availability of psychological treatments as well as anonymity and flexibility of usage. Risks include treatment disadvantages such as a lack of options to adapt treatment to participants' individual needs. In addition, despite good possibilities to monitor treatment, practices regarding quality assurance such as monitoring and feedback vary greatly across countries depending on the national regulations (Ebert et al., 2018). Depending on the implementation setting, the lack of continuity wtih other treatment forms could also be a potential pitfall resulting in non-response. These potential risks underline the necessity to continue research on internet interventions in order to find ways to optimize treatment and maximize treatment effects, while minimizing potential risks for participants.

So far, research findings have supported the efficacy of internet interventions (Andrews et al., 2018; Carlbring et al., 2018; Ebert et al., 2018), however, only few findings are consistent regarding predictors of treatment outcome (Schuster et al., 2018). This means knowledge of important variables indicative of treatment outcome is still limited. However, this does not only apply to internet interventions, but also to face-to-face treatments (e.g., Fuertes & Nutt Williams, 2017). Many studies investigating predictors of outcome have focused on variables such as initial impairment, comorbidity of mental disorders, personality disorders, chronicity, and treatment expectations (Delgadillo, Moreea, & Lutz, 2016). While there is good evidence for the impact of the therapeutic relationship on outcome (see Norcross & Lambert, 2018), findings regarding other variables such as gender, age, and socio-economic status remain inconsistent (Bohart & Greaves Wade, 2013). In addition, it is important to note that therapist effects explain a rather small amount of variance (Schiefele et al., 2017) and that there is no indication that one of the empirically-based psychological treatments is generally superior to the other in terms of efficacy (Fuertes & Nutt Williams, 2017). Thus, Fuertes & Nutt Williams (2017) argue that research should focus more on the patient, his views and experiences, as the patient is the one who is working the hardest toward change or improvement in therapy, even when supported by a therapist.

Interestingly, a recent meta-analysis showed that treatments including homework showed higher effect sizes than treatments without homework (Kazantzis et al., 2018), highlighting how important it is for patients to apply what they learn in therapy in daily life. In this context also intersession processes have to be mentioned, which describe patients becoming active outside of the treatment setting, e.g. by seeking out specific experiences, self-disclosing to others about treatment or evaluating therapeutic advice (Bohart & Greaves Wade, 2013). Thus, independent of the treatment itself, the way a patient contributes to treatment, puts effort in and applies what was learned, may be one of the most important factors for treatment outcome (Bohart & Greaves Wade, 2013). The amount and quality of patient engagement may vary depending on individual differences in several domains such as differences in impairment and level of functioning, differences in more stable characteristics such as demographics and life circumstances as well as differences in intrinsic personal factors such as motivation (Fuertes & Nutt Williams, 2017), attitudes, and treatment preferences (Swift, Callahan, Cooper, & Parkin, 2018).

To summarize, despite a large amount of research, it remains unclear, why certain patients benefit from psychological treatments, while others do not (Lambert, 2013). Therefore, improving outcome predictions for individual patients has become a central task of patient-focused research (Delgadillo et al., 2016). In recent years, this area of research has seen strong development: New tools such as the personalized advantage index (DeRubeis et al., 2014) have been developed and new modeling approaches such as network modeling (e.g., Lutz et al., 2018) have been introduced to improve the prediction of treatment outcome for individual patients. These efforts emphasize the need as well as the intention to improve treatment selection and treatment outcome for individual patients.

Especially in the context of internet interventions, knowledge of predictors of treatment outcome may be important: On the one hand, some participants may be more prone than others to benefit from the specific advantages of internet interventions, such as temporal flexibility or anonymity. On the other hand, some participants may be especially at risk of suffering from specific disadvantages such as a lack of direct contact. For example Delgadillo, Huey, Bennett, & McMillan (2017) have pointed out internet interventions can be demanding as the responsibility of improvement lies mostly on the participants themselves. Especially in complex cases patients may be unable to meet this demand, e.g. when they have few resources and are already struggeling with daily life (Delgadillo et al., 2017). To conclude, it is important to develop empirically based recommendations to inform treatment selection (e.g. internet intervention vs. face-to-face treatment) and adaptation. For this reason one focus of this dissertation will be the question what predicts treatment outcome for participants of internet interventions.

In addition, this dissertation will consider a question that is often overlooked in studies of psychological treatments: How is dropout related to treatment change? This seems relevant, as, during nearly every treatment, there are a number of persons who do not complete the scheduled assessments. Still, there are only a limited number of psychological studies that have used specific statistical approaches to investigate dropout and its relation to treatment change (Yang & Maxwell, 2014) and to date none of them have focused on internet interventions.

By investigating these research questions in more depth, it may be possible to gain a better understanding of factors that influence change processes during internet interventions. This, in turn, may clarify how to optimize treatments for different participants. Therefore, this dissertation summarizes three studies that were designed to fill research gaps in the context of internet interventions.

Study I (Chapter 5) investigated the occurrence of early change patterns in an internet intervention targeting depression. Data from 409 participants who had filled out at least three assessments were analyzed via Growth mixture modeling (GMM). A piecewise model was applied to model change from screening to registration (pretreatment) as well as early treatment change (registration to week four of treatment). Three early change patterns were identified with two patterns characterized by improvement and one by deterioration. The findings indicated that different early change patterns also occur in internet interventions and are predictive of treatment outcome as well as adherence. In addition, important patient characteristics that are predictive of early change (initial impairment in depressive symptoms and in physical health) were identified. The results suggest that early change in internet interventions should be closely monitored: If early deterioration occurs, the allocation to another, more intensive treatment may be indicated.

Study II (Chapter 6) investigated how change and dropout are interrelated by applying the Muthen-Roy model, which allows the estimation of change-dropout patterns. When applying this model, it is assumed that dropout may be linked to unobserved variables, such as a latent change pattern characterized by deterioration. In this study, the same dataset from Study I was used including a slightly larger sample of 483 participants. Four different change-dropout patterns were identified with dropout being associated with either very rapid improvement or deterioration. Participants with different change-dropout patterns also differed in adherence. Participant characteristics such as age, initial impairment, physical and mental health as well as attitudes towards internet interventions were significant predictors of change-dropout patterns. The findings suggest that dropout may be linked to deterioration as well as to a "good enough dosage of treatment". The risk of dropout may depend on certain patient characteristics, which should be considered when allocating participants to treatment.

Study III (Chapter 7) compared adherence and outcome in different participant groups and investigated the impact of adherence to treatment components on outcome in an internet intervention targeting anxiety symptoms. The context in which the intervention was provided was considered by comparing self-referred participants (N = 37) to outpatient participants (*N* = 50) waiting for face-to-face treatment. Differences in adherence to treatment components and treatment outcome were investigated. To investigate the effect of the internet intervention on change during the waiting period, outpatient participants were compared to a matched sample of outpatients without access to the internet intervention. Adherence to treatment components was investigated as a predictor of treatment outcome. The results suggest that rather than only using general measures of adherence, studies should also consider adherence to treatment components, which are considered the basis of therapeutic change. Furthermore, the efficacy of internet interventions should be compared more systematically for different participant subpopulations.

All three studies are depicted in Chapters five to seven. Chapter two describes a common theoretical background, which provided the basis of the studies and leads to the deduction of the research questions in chapter three. To facilitate understanding and interpretation of the studies, Chapter four provides the reader with the most important methodological specialties of the studies. Chapters five to seven include Study I, Study II, and Study III, respectively. Finally, a general discussion of the studies is presented in Chapter eight, where future research areas as well as practical implications are considered.

3. Theoretical Background

3.1. Differences in treatment outcome: Early change

Besides the question whether internet interventions are effective considering their general advantages and disadvantages, another important question is which patients specifically benefit from internet interventions and whether outcome can be predicted for different groups of participants. This can be seen as an effort to identify the optimal treatment for each patient and prevent treatment failure. Andrews & Williams (2014) estimated that 50% of participants in internet interventions improve, with the other 50% showing no improvement. Thus, the question is how internet interventions can be improved. Currently, there are two approaches that have received increasing attention in research: The stepped care approach (e.g. Delgadillo, Gellatly, & Stephenson-Bellwood, 2015) and the personalized treatment approach (Fisher, 2015). While stepped-care approaches aim to improve treatment availability and cost effectiveness, personalized treatment approaches try to optimize treatment according to individual needs (Forsell et al., 2019). While one study found that an adapted internet intervention improved outcome for patients at risk of treatment failure (Forsell et al., 2019), research in this area is still sparse, especially regarding internet interventions. It is therefore vital to promote research in this field (see Fisher, 2015). This could facilitate a more precise allocation of participants to internet interventions and increase adherence as well as treatment outcome.

There have been many efforts to predict outcome for participants of internet interventions. Yet, despite these efforts, it remains difficult to predict who will benefit from internet treatment (Schuster et al., 2018). Without reliable information on predictors of treatment outcome, it is difficult to make empirically-based clinical decisions regarding treatment. This implies that either a standardized approach is used (all patients with the same diagnosis are allocated to a similar treatment) or clinical decisions depend on a person such as a clinical psychologist or a nurse (see Gunlicks-Stoessel & Mufson, 2011). Both approaches are clearly unfavorable, as they do not ensure the systematic allocation of participants to their optimal treatment and may introduce an unnecessary risk of treatment failure.

In addition to the investigation of variables that predict pre-to-post change, the investigation of change over time based on multiple measurements can be valuable: It can shed some more light on the point in time when change occurs (Laurenceau, Hayes, & Feldman, 2007). This, in turn, can allow us to more specifically investigate for whom change occurred at which point in time and why or why not. On average, therapeutic change follows a log-linear shape with much change occurring at the beginning of treatment and less change taking place later in treatment. This finding was first reported in 1986 in a study by Howard, Kopta, Krause, & Orlinsky (1986) and was later replicated (Robinson, Delgadillo, & Kellett, 2019; Stulz, Lutz, Kopta, Minami, & Saunders, 2013). However, this does not mean that there are no differences in change between participants. A study by Harnett, O'Donovan, & Lambert (2010) showed that the number of sessions needed for patients to show reliable improvement differed across participants: 50% of participants showed reliable improvement after 8 sessions, while 80% showed reliable improvement after 21 sessions. The dose-response effect in psychotherapy was also investigated in a recent review (Robinson et al., 2019): In routine care settings a dose of 4 to 26 sessions was necessary for patients to reach reliable improvement and in low intensity treatments the optimal dose amounted to four to six sessions. In line with these findings, several studies have reported that the shape of change differs across participants. While some participants show rapid early improvement (early response) and good treatment outcome, there are also participants who do not improve or even deteriorate early in treatment (Lutz et al., 2014; Lutz, Stulz, & Köck, 2009; Nordberg, Castonguay, Fisher, Boswell, & Kraus, 2014; Stulz, Gallop, Lutz, Wrenn, & Crits-Christoph, 2010; Thibodeau et al., 2015).

Differences in early change in participants of internet interventions may be especially important, because participants have limited access to direct support in case of a crisis and continuity of treatment is not always ensured. Thus, the occurrence of different early change patterns in internet interventions has important implications for treatment adaption (e.g. additional guidance or stepping-up care as responses to early deterioration). However, to date, there has been little research on early change in internet interventions (Schibbye et al., 2014) and low intensity interventions (e.g. Delgadillo et al., 2014), with most studies focusing on pre-to-post comparisons to investigate treatment effectiveness. This is a shortcoming, as considering participants' early change patterns and reacting appropriately could minimize risk for participating patients.

When looking at the findings reported above, it must be considered that different studies use different methods to define early change (see Rubel et al., 2015). Delgadillo et al. (2014) focused on early improvement by using change scores, Schibbye et al. (2014) calculated early change using a regression slope. Several studies that investigated change patterns in face-to-face treatments have used Growth mixture modeling (GMM; Gunlicks-Stoessel & Mufson, 2011; Hunter, Muthén, Cook, & Leuchter, 2010; Lutz et al., 2014) to identify different subpopulations of patients showing similar growth data. However, these models have yet to be applied to investigate change patterns in internet interventions. In contrast to other methods, GMM allow the consideration of multiple measurement points, provide indictors of model fit and allow the identification of unobserved patterns (Johnson, 2015). Therefore, GMM is an interesting approach to gain a clearer picture of change patterns in internet interventions. This type of research that focuses more on individual differences in change patterns could inform clinical decision-making, especially as clinical decisions tend to be more difficult to make in long distance treatment settings.

3.2. Differences in treatment outcome: Dropout

There is a general problem that is often only remotely considered in studies that apply multiple measurements to investigate psychological treatments: The occurrence of missings due to participant dropout (Yang & Maxwell, 2014). Dropout occurs when participants, who have completed one or more assessments before and during treatment, do not complete the remaining scheduled assessments. Instead, assessments are often only available up until some point in time, with no further assessments available (Demirtas & Schafer, 2003). This means that some of the information that was intended to be collected is not available, which hinders the intention of making inferences about the entire data's distribution regarding a specific variable, such as treatment change. To examine treatment change, linear mixed models are often applied for analysis, which estimate an individual specific growth trajectory including fixed (population characteristics) and random effects (individual-specific intercept and slope). Depending on why data is missing, different models have to be taken into account to avoid biased estimates.

If the subsample of participants with missing values at post-treatment is a random subsample of all participants (e.g. participants who moved away or got sick), there is no reason to believe that outcome estimates may differ systematically between the subsample of missing participants and the full sample (Coertjens, Donche, Maeyer, Vanthournout, & van Petegem, 2017). In this case, there is no relationship between the missing value at post-treatment and any observed or unobserved value in the data set (missing completely at random; MCAR). Sometimes, a missing value at a certain point in time is related to the previous measurement, e.g. a patient drops out after having had high impairment scores before. In this case, because dropout depends on observed variables, it is possible to control for this relationship. This mechanism is called MAR (missing at random), meaning that while missingness may depend on observed variables, there is no dependence of dropout on missing values (Coertjens et al., 2017). Both of these missing mechanisms are considered ignorable (Yang & Maxwell, 2014). In contrast, when dropout is outcome-dependent (it depends on the missing value itself, e.g. a high impairment score) or coefficient-dependent (it depends on the unobserved process of change reflected by growth trajectories), it is non-ignorable (not missing at random; NMAR). Importantly, imputation procedures assume missings to be at least MAR. If these imputation procedures such as maximal likelihood approaches and multiple imputation are applied when data is NMAR they may lead to biased estimates, because the participants who drop out are different from other participants (e.g. participants experiencing deterioration or relapse during treatment, see also Yang & Maxwell, 2014). Thus, it is important to apply appropriate models that consider the mechanism of dropout.

There are a number of statistical models that can be applied in such a case (Enders, 2011; Muthén, Asparouhov, Hunter, & Leuchter, 2011). One model that is considered particularly interesting is the Muthen-Roy model (Muthén et al., 2011), which identifies subpopulations of participants, who are similar regarding change and dropout. This is achieved by including two latent variables in the model, one for change and one for dropout. In contrast to other models, which have been criticized for limited interpretability, the Muthen-Roy model allows a clearer interpretation of the emerging patterns (Muthén, Asparouhov, Hunter, & Leuchter, 2011).

Despite these possibilities to model dropout, dropout mechanisms are rarely investigated in psychological intervention studies (Coertjens et al., 2017). This is a shortcoming when one considers the high dropout rates that are reported in some studies investigating internet interventions (Fernandez, Salem, Swift, & Ramtahal, 2015). These varying dropout rates in internet interventions could be the result of implementation differences. Under some conditions (e.g. no obligation to fill out measurements to access the intervention), dropout rates may be high. In addition, as participants of internet interventions fill out questionnaires on their own (see above), some participants may experience the task itself as burdensome, feel monitored or pressured (e.g. Ebert et al., 2018) as well as unsure about how to react if they do not feel they are progressing. These kinds of issues may be potentially difficult to resolve without direct contact, which is not available during internet interventions.

By considering the mechanisms of dropout, a better understanding of dropout and its implications regarding treatment outcome in internet interventions can be achieved. On the one hand, if participants who dropout are more prone to showing negative change, researchers should consider this when estimating treatment effects. On the other hand, clinicians may have to think about ways to identify participants at risk of dropout early and intervene to avoid poor outcome for these participants.

3.3. Differences in treatment outcome: Adherence

To answer the question, who benefits from internet interventions more fully, it is also important to shed more light on the processes that lead to a positive or poor outcome in participants. As has been previously described, engagement or adherence is assumed to be crucial to treatment outcome, as it affects treatment dosage (Cooper et al., 2018). In face-to-face treatments, adherence is often defined via session attendance with only few studies exploring it in more depth. As a consequence, low adherence is often operationalized via premature treatment termination (Koffel, Vitiello, McCurry, Rybarczyk, & Korff, 2018). On average, rates of premature treatment termination in face-to-face treatment range between 20% and 50% (Swift, Greenberg, Tompkins, & Parkin, 2017). Rates vary depending on diagnosis, with higher rates of premature treatment termination reported for patients with eating disorders or substance abuse, and on setting of provision, with higher rates of premature termination in naturalistic settings compared to RCTs (Cooper et al., 2018).

Furthermore, higher rates of premature treatment termination have been reported in treatments targeting PTSD and anxiety symptoms using exposure (Cooper et al., 2018), which could indicate that patients are afraid to engage in it. Interestingly, studies focusing on insomnia and pain also report low adherence, as indicated by low session attendance and a low application of relevant treatment elements (see Koffel et al., 2018; Matsuzawa et al., 2019). Thus, patients' perception of their problems as well as their perception of treatment difficulty and their perceived probability of treatment success may represent crucial factors regarding adherence. Specifically, Matsuzawa et al. (2019) assumes motivation, lack of knowledge regarding mental health problems, locus of control, self-efficacy and coping styles to be important factors that impact adherence. In contrast to this suggestion, many studies have focused more on patient characteristics such as age and gender as predictors of premature treatment termination. However, they have yielded inconsistent results and recently the strong emphasis on these variables has been criticized (Cooper et al., 2018). Instead of focusing on stable characteristics to predict premature termination, Cooper et al. (2018) also suggests examining the impact of more modifiable factors, such as hope and motivation, to increase the ability to target possible risk factors at the beginning of treatment. In practice, it is likely that a patient enters treatment with an initial probability of leaving treatment prematurely based on his perceptions and beliefs (Cooper, Kline, Baier, & Feeny, 2018). Then, during the treatment process, the original risk of premature termination increases or decreases. Treatment-related factors that may be relevant in this context include inherent treatment features such as treatment modality (e.g. group or individual) as well as emergent properties such as the development of a positive therapeutic relationship in face-to-face treatment (Cooper et al., 2018). Internet interventions also represent a special treatment modality and are characterized by a setting where no direct contact to the therapist is provided. This has several implications, with some of the most important being:

1. Participants need a specifically high level of commitment to use internet interventions independently of another person (Ebert & Baumeister, 2017). To form such a commitment, participants must be convinced that it is worth the effort. This may depend on

several variables such as personal motivation for treatment (e.g. experience of symptoms and wish for improvement), treatment expectations, treatment credibility, alliance to treatment (e.g. regarding goals and tasks) as well as self-efficacy (Alfonsson, Olsson, & Hursti, 2016; Beatty & Binnion, 2016; El Alaoui et al., 2015; Johansson, Michel, Andersson, & Paxling, 2015). In addition, acceptability and attitudes towards internet interventions (e.g. regarding anonymity and use of technology) are considered as relevant (see Schröder et al., 2017).

2. As participants cannot receive any direct support in case of a crisis and most internet interventions are rather standardized, they may be less suitable for participants who are highly impaired, have comorbidities or special issues and wish for direct contact (Delgadillo et al., 2017; Ebert et al., 2018).

3. Participants are largely “self-monitoring” and self-administrating therapeutic methods. This may lead to participants having difficult experiences, such as feeling under pressure or feeling overwhelmed (Ebert et al., 2018). Also, when failing to improve, participants may experience reduced self-efficacy and develop negative attitudes towards psychological treatments (Ebert et al., 2018).

While there have been several efforts to limit disadvantages of internet interventions regarding points two and three (e.g. stepping-up participants’ care, varying degrees of monitoring and guidance), point one (participant commitment) is not an issue that can be as easily resolved. This is illustrated by the large amount of variance found regarding participant adherence (Andrews & Williams, 2014; Beatty & Binnion, 2016; Hilvert-Bruce, Rossouw, Wong, Sunderland, & Andrews, 2012). While not all participants may need the same dosage of treatment (see the last section), for some participants, low adherence may be an indicator of treatment failure. It may be that participants do not adhere to treatment, because they have trouble committing to treatment (see point one), or because they experience difficulties during treatment (see points two and three). While many studies report high adherence rates (up to 75%), most participants volunteered and were additionally encouraged by researchers

(Andrews & Williams, 2014). The high level of attention and highly structured nature of RCTs may enhance adherence in RCT participants and lead to an overestimation of efficacy when results from these trials are transferred to routine care (Ebert & Baumeister, 2017). In contrast, adherence rates drop by 50% when a clinician in primary care prescribes the intervention. Therefore, it is important to further investigate how participant adherence is influenced by the referral context.

While several studies have found that adherence predicts outcome (El Alaoui et al., 2015; Hilvert-Bruce et al., 2012; Titov et al., 2013), there are no consistent findings regarding predictors of adherence (Beatty & Binnion, 2016). An important shortcoming of these studies' investigation of adherence is that most studies used general measures of adherence (e.g. Castro et al., 2018; El Alaoui et al., 2015; Klein et al., 2016). Results from a study by Alfonsson, Olsson, & Hursti (2016) indicate that measuring adherence in different ways provides somewhat different results. The authors conclude that it is necessary to carefully define treatment adherence in psychotherapy research.

Furthermore mechanisms of change must be considered: Alongside factors that are common across most treatments (e.g. providing psychoeducation and inducing hope), therapeutic change is assumed to be based on treatment specific factors or treatment components (e.g. Wampold, 2015). Therefore, adherence to these components should be more closely studied. This would allow the identification of potential predictors of adherence to treatment components, which could be vital for good treatment outcome.

4. Research questions

Section two introduced important concepts from patient-focused research and described a lack of research in this field regarding internet interventions. One trend in patient-focused research is to investigate interindividual differences in early change, which is meant to improve prediction of treatment outcome and add to the process of optimal clinical decision-making. The knowledge gained in this field regarding face-to-face treatment still remains to be transferred to research on internet interventions to allow for improvements that benefit participating patients. In this context, in sections two and three, dropout and adherence were described as a potential moderators of treatment outcome in internet interventions.

Regarding dropout, two important points were made in section two: First, studies investigating treatment effects need to consider a NMAR mechanism to avoid the risk of biased estimates. The current neglect of this issue and the lacking application of appropriate models when necessary is an important shortcoming, especially in studies with high dropout rates. A better understanding of dropout mechanisms and the relation of dropout to change may inform researchers as well as clinicians.

Regarding adherence, not much is known about whether adherence and outcome vary between different participants groups. Most studies focusing on internet interventions have either investigated self-selected participants or participants in primary care, without comparing these groups. In addition, to date, most studies have used general measures of adherence only, neglecting the proposed mechanisms of change underlying psychological treatments.

In the following, a short depiction of the theoretical background relevant to each study as well as the research questions addressed in each study is provided.

Study I: Background and research questions

Not much is known about early change patterns of participants in internet interventions and it remains unclear which participants are likely to show early response. Study I investigated early change patterns as well as their associations with adherence and treatment outcome in an internet intervention targeting depressive symptoms.

Thus, Study I's research questions were:

1. Do participants of internet interventions show different early change patterns?
2. Can early change patterns predict adherence and treatment outcome in internet interventions?
3. Which patient characteristics are predictors of early response in internet interventions?

Study II: Background and research questions

The dropout mechanism as well as the relation of dropout and change remains unclear in many clinical studies. Study II investigated the mechanisms of dropout in internet interventions and the relation of dropout to change as well as treatment outcome.

In Study II, the following research questions were investigated in participants of an internet intervention targeting depression:

1. Which change-dropout patterns can be identified using the Muthen-Roy model?
2. How is dropout related to treatment outcome?
3. Which characteristics predict treatment change and dropout in participants?

Study III: Background and research questions

To overcome some limitations of previous studies, Study III compared outpatient participants to self-referred participants and expanded the perspective on adherence by investigating adherence to treatment components and its impact on outcome in an internet intervention targeting anxiety symptoms. Furthermore, the efficacy of the internet intervention for outpatient participants who waited for face-to-face treatment was investigated.

The following research questions were investigated in Study III:

1. Are there differences between self-referred and outpatient participants of an internet intervention regarding adherence and treatment outcome?
2. What is the impact of adherence to treatment components (relaxation, cognitive restructuring, and exposure) on treatment outcome in an internet intervention targeting anxiety disorders?
3. How effective is an internet intervention targeting anxiety for outpatients who access it during the waiting period for face-to-face therapy?
4. Which characteristics predict adherence to treatment components and outcome in participants of an internet intervention targeting anxiety disorders?

To show how the described research questions were investigated in the following section a short depiction of the methodological specialties of the studies is provided.

5. Methodological aspects

To identify early change patterns in Study I, Growth Mixture Modeling (GMM) was used. In Study II, an extension of this model was applied to investigate change-dropout patterns (Muthen-Roy model). In Study III, LASSO regression was used to select the most important predictor variables. Therefore, these methods are described briefly in the following.

5.1. Growth Mixture Modeling (GMM)

GMM can be considered a person-centered analysis (Jung & Wickrama, 2008). Its focus is on the relationships among individuals with the goal of identifying different subgroups of participants based on individual change patterns. Conventionally, a growth model can be described as a multilevel random effect model. This is due to the hierarchical structure of longitudinal growth data with serial measurements clustered within participants: The first level is therefore the measurement occasion, the second the participant (Johnson, 2015). The terminology "mixed model" refers to an analytical procedure in which fixed effects that capture the sample average growth curve and random effects that capture individual departures from that average curve are estimated (Johnson, 2015).

Specifically, GMM is based on latent variables that represent different aspects of individual change: An intercept parameter that captures the level of outcome at the beginning of measurement and a slope that captures change in outcome over time (Berlin, Parra, & Williams, 2014). In GMM, both model parameters, the intercept and the slope, have means that reflect the average of all participants' intercepts and slopes. In addition, for each participant, an individual intercept and slope is estimated, which varies across participants. This variability in parameters is captured by the variance. In addition, to see how the initial level of outcome is related to the rate of change, the covariance of the intercept and the slope can be estimated to reflect this relationship. Furthermore, deviation from the parameter means at each time point is captured by residuals (random effects). For all parameters, significance tests are available to determine whether they are different from zero (Berlin et al., 2014).

By fixing the factor loadings of the intercept and slope, it is possible to test varying hypotheses about initial outcome and the shape of change, such as a log-linear or linear shapes of change. Based on similar growth data patterns, participants are classified into latent classes using probabilistic assignment (similar to a logistic regression, see Berlin et al., 2014). Thus, the observed distribution of all outcome data is assumed to stem from a mixture of different subpopulations. In GMM, all the parameters mentioned above may vary between and within classes. The goal of this analytic approach is to understand and predict the emerging differences between participants regarding parameter estimates, which reflect differences in change over time (Berlin et al., 2014).

The estimation of a GMM is performed in four steps (Ram & Grimm, 2009): First, it is necessary to decide how to model change over time. To do so, an appropriate change function (linear, log-linear, quadratic) is identified based on theory or prior research. This is done by applying single-group models and comparing model fit. Next, model specification takes place. In this step, the number of latent classes is determined and further specifications are made, such as fixing or freeing the mean and variance of latent variables. In a next step, the estimation methods are applied and model fit is evaluated. Finally, the best model is chosen, considering different statistical fit indices (such as the Bayesian Information Criterion; BIC, the entropy, and the Boostraped Likelihood Ratio Test; BLRT).

Following this recommended procedure, in Study I, the best change function was investigated first. In line with the goal of identifying early change patterns only, a limited number of measurements was included (screening, pre-treatment and in-treatment measurements until week four of treatment). Based on the assumption that the rate of change during treatment would differ from the rate of change during the waiting period, a piecewise model was estimated. Two slopes were estimated: one to estimate change during the waiting period for the intervention from screening to pre-treatment, and one from registration until week four of treatment (see Figure 1).

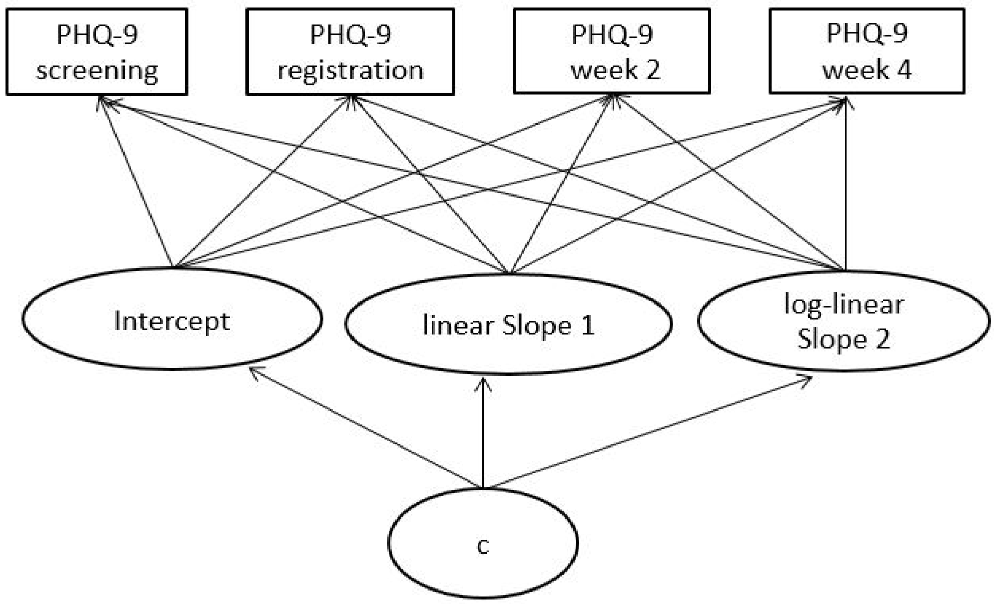

Figure 1. Piecewise GMM with one latent class variable (c) and observations on the Patient Health Questionnaire-9 (PHQ-9) from screening until the fourth week of treatment.

Subsequently, the function of change was identified by applying relative model fit indices (BIC; Schwarz, 1978). In a second step, the model parameters were specified and the number of latent classes was estimated using a combination of BIC and BLRT (Nylund, Asparouhov, & Muthén, 2007).

To investigate overall treatment change in Study II, just one log-linear slope was estimated and all timepoints (screening, pre-treatment, in-treatment measurements from week 2 until week 12) were included in the model.

5.2. The Muthen-Roy model

GMM is based on a maximum likelihood analysis and thus on the assumption that missings are MAR. However, as described above, this does not always have to be true, as dropout may depend on the missing outcome or a latent change trajectory (Yang & Maxwell, 2014). For these situations, different NMAR models have been proposed, one of them being the Muthen-Roy model (Muthén et al., 2011). In addition to a latent class variable that captures change (cy), a second latent class variable that captures dropout (cu) is included in the model. This latent class variable that reflects dropout patterns is estimated based on dummy coded variables d1, d2… that indicate whether dropout occurred at each measurement occasion. The dropout patterns do not directly influence the estimation of the change patterns. Instead, the latent variable that captures patterns of change is estimated the same way as in a conventional growth model: It is based on similar growth data as indicated by the latent growth parameters (intercept and slope, see Figure 2 at the end of this paragraph). As a result of the model, the probability of a participant is calculated to show a combination of a specific dropout pattern and a specific change pattern. This procedure results in a higher number of identified patterns than in conventional GMM and increases the complexity of the model. Nethertheless, the Muthen-Roy model allows the investigation of the combination of dropout and change patterns, while also allowing the separate estimation of latent class variables.

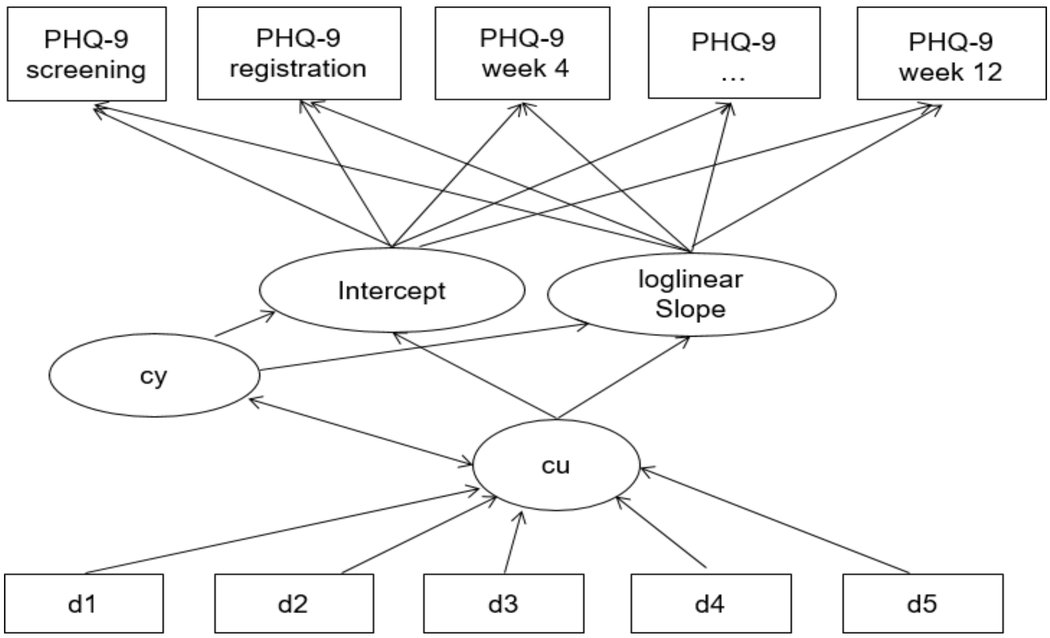

Figure 2. Muthen-Roy model with two latent class variables (cy and cd), biweekly observations on the in-treatment questionnaire (Patient Health Questionnaire-9; PHQ-9) and dropout indicators (d1-d6).

5.3. Least Absolute Shrinkage and Selection Operator (LASSO)

Regression analysis deals with parameter estimation and variable selection, however sometimes problems regarding prediction accuracy arise (Muthukrishnan, 2016). In this context, feature selection can be used to choose only relevant variables, improve interpretability of the model and remove redundant information. Another reason for the application of feature selection is to reduce the risk of overfitting (McNeish, 2015). Different approaches exists such as subset selection, dimension reduction, and shrinkage.

LASSO (Tibshirani, 1996) is a shrinkage approach that includes penalized regression. It penalizes the regression model by putting a constraint on the absolute sum of the coefficients (L1-norm) and applies a shrinking procedure to the coefficients of the regression variables. Thus, it forces some of the coefficient estimates with a minor contribution to the model to be zero. After this process, the variables that still have non-zero coefficients are selected to be part of the model. LASSO improves prediction accuracy and model interpretability (Muthukrishnan & Rohini, 2016 - 2016).

As the current section described important methodological specifics of the studies, in the following three sections, Studies I, II and III are presented.

6. Study I: Defining and Predicting Patterns of Early Response in a Web-Based Intervention for Depression

Lutz, W., Arndt, A., Rubel, J., Berger, T., Schröder, J., Späth, C., . . . Moritz, S. (2017). Defining and Predicting Patterns of Early Response in a Web-Based Intervention for Depression. *Journal of Medical Internet Research*, *19*(6), e206. https://doi.org/10.2196/jmir.7367

Author contributions

W. Lutz was responsible for the concept as well as data management and the final draft. A. Arndt was responsible for the data preparation, data analysis and writing of the manuscript. J. Rubel contributed to the data preparation, data analyses, and revised the manuscript. J. P.Klein was responsible for the study design and the coordination among study centers. He was also responsible for the data management and revised the manuscript. J. Schröder contributed to the study design, developed the attitudes towards psychological interventions questionnaire and contributed to the coordination of study centers as well as the data management. B. Meyer contributed to the study design, was responsible for the cooperation with GAIA AG, which made the intervention available to participants and was also responsible for the data management. C. Späth contributed to the study design and coordination of study centers. T. Berger contributed to the study design, the coordination of study centers and was responsible for the feedback provided to participants.All other authors also contributed to the contributed to the study design, the recruitment of participants, the data collection and the revision of the manuscript. All authors contributed to and have approved the final manuscript.

6.1. Abstract

Background: Web-based interventions for individuals with depressive disorders have been a recent focus of research and may be an effective adjunct to face-to-face psychotherapy or pharmacological treatment.

Objective: The aim of our study was to examine the early change patterns in Web-based interventions to identify differential effects.

Methods: We applied piecewise growth mixture modeling (PGMM) to identify different latent classes of early change in individuals with mild-to-moderate depression (n = 409) who underwent a CBT-based web intervention for depression.

Results: Overall, three latent classes were identified (N = 409): Two early response classes (N = 158, N = 185) and one early deterioration class (N = 66). Latent classes differed in terms of outcome (p<.001) and adherence (p = .03) in regard to the number of modules (number of modules with a duration of at least 10 minutes) and the number of assessments (p<.001), but not in regard to the overall amount of time using the system. Class membership significantly improved outcome prediction by 24.8% over patient intake characteristics (p<.001) and significantly added to the prediction of adherence (p = .04).

Conclusions: These findings suggest that in Web-based interventions outcome and adherence can be predicted by patterns of early change, which can inform treatment decisions and potentially help optimize the allocation of scarce clinical resources.

6.2. Introduction

Web-based interventions for individuals with depressive disorders have been a recent focus of research and may be an effective addition to face-to-face psychotherapy or pharmacological treatment. For example, such interventions may be appropriate for individuals who have difficulty accessing psychological treatment or do not want to utilize face-to-face treatment (Barak & Grohol, 2011; Johansson & Andersson, 2012; Moritz, Schröder, Meyer & Hauschildt, 2013; Ryan, Shochet, & Stallman, 2010; Titov, 2011). Several studies suggest that some forms of Web-based interventions may be as effective as face-to-face therapy (Andrews, Cuijpers, Craske, McEvoy & Titov, 2010), although various methodological limitations of this body of research have also been noted (Arnberg et al., 2014). One limitation of Web interventions is that they are not accepted by all patients and some drop out early or do not adhere to the treatment protocol (Gilbody et al., 2015). Especially in unguided Web interventions, the risk of dropout is high (Melville, Casey, & Kavanagh, 2010; Richards & Richardson, 2012) and results of studies on prereatment predictors of outcome in Web interventions remain inconsistent (Andersson & Hedman, 2013). Additionally, not all Web interventions are equal with regard to their quality or evidence base (Renton et al., 2014). So far, investigations of Web interventions have mainly focused on treatment efficacy and short-term symptom change in comparison with treatment-as-usual control groups, in which participants were only able to access the Web intervention after a delay of several weeks or months (Leykin, Muñoz, Contreras, & Latham, 2014; Richards et al., 2014).

Whereas a good database has been established regarding the general effectiveness of several Web-based interventions for the treatment of psychological problems, there is still a lack of research investigating the process and shape of change (Laurenceau, Hayes, & Feldman, 2007). On the other hand, this area of research has a certain tradition in individual therapy. In recent years, interest in the investigation of early change patterns and their relation to outcome has grown. The basic idea behind this research is to use early change of the target behavior (e.g. depressive symptoms) to predict treatment outcome (Cuijpers, Lier, van Straten, & Donker, 2005; Lutz, Stulz, & Köck, 2009). Early change patterns have been shown to be associated with outcome across different diagnoses (Bradford et al., 2011; ; Lewis, Simons, & Kim, 2012), different treatment approaches (Crits-Christoph et al., 2001; Gunlicks-Stoessel & Mufson, 2011), and different measures (Hunter, Muthén, Cook, & Leuchter, 2010).

For example, a recent study by Lutz and colleagues (2014)investigated early change patterns in patients with panic disorder (*N* = 326), who underwent manualized cognitive behavioral therapy (CBT). Using growth mixture modeling (GMM), 4 latent subgroups were identified, showing clusters of change trajectories over the first 5 sessions. One of the subgroups consisted of patients whose symptoms decreased rapidly and who also showed the best outcomes (early responders). This information on early response improved treatment prediction by 16.1% over patient intake characteristics. Early change patterns also significantly predicted early dropout. Likewise Delgadillo and colleagues (2014) focused on early change patterns during low intensity interventions. This study used data from patients with anxiety disorders or depression, who accessed the Improving Access to Psychological Therapies (IAPT) service in the United Kingdom (*N* = 511). This service comprised between 1 and 8 sessions and was often delivered via telephone. Early response, defined as reliable improvement until session 4, was predictive of clinically significant recovery after treatment termination. It was noted that attrition was highest in early sessions, so that early attempts to engage patients should be made.

This example emphasizes the importance of also studying early change patterns in low intensity and Web interventions. Whereas conclusions drawn from efficacy and effectiveness studies are limited to the average patient after treatment, knowledge about early change patterns may help to answer clinical questions. For example, such knowledge could be used to predict whether the treatment in question will work for a particular subgroup or to decide whether users should continue treatment (Krause, Howard, & Lutz, 1998; Lutz, 2002). Such questions have become increasingly relevant in the face of the recent implementation of stepped-care models, where patients are matched to a treatment with the option of being "stepped-up" to more intensive care (National Institute for Health and Clinical Excellence; NICE, 2011; van Straten, Hill, Richards, & Cuijpers, 2015). Knowledge about predictive determinants may add to the development of empirically based rules that support clinicians in their decisions (van Straten, Hill, Richards, & Cuijpers, 2015) and may help to prevent dropout or low adherence. In addition, investigating change patterns promotes the understanding of change processes, which is necessary for treatment development efforts (Laurenceau et al., 2007).

Although early change patterns are important predictors of treatment outcome (Lutz et al., 2014; Nordberg, Castonguay, Fisher, Boswell, & Kraus, 2014), to date only one study has looked at early change patterns in patients undergoing Web-based interventions. Schibbye and colleagues (2014) examined change patterns during a CBT-oriented Web-based intervention, which was provided to patients with panic disorder, social phobia, or depression ($N = 112$) by the Internet Psychiatry Clinic in Sweden. Outcome of the Web intervention was predicted by estimation of early change. The prediction was best when the rating of a disorder-specific measure at week 4 was used.

In the present study, we analyzed data from a multicenter trial testing the efficacy of a CBT-oriented web-based intervention for individuals with mild to moderate depression. Based on the existing literature on individual therapy, we predicted the existence of distinct early patient response clusters in this Web intervention. We further hypothesized that these clusters would add to the prediction of treatment outcome as well as adherence. This study also examined whether initial impairment, attitudes toward Web-based interventions, and email support predict early change patterns.

6.3. Methods

Participants and Treatment

This study was conducted from January 2012 to December 2013 and approved by the local ethics committee (DGPs, reference number SM 04_2012). Written informed consent was obtained, and the study was registered at ClinicalTrials.gov (identifier: NCT01636752). Several settings were used to recruit participants: (1) In- and outpatient medical and psychological clinics, (2) internet forums for depression, (3) health insurance companies, and (4) the media (eg, newspaper). Participants were directed to the study's website. In total, 2020 participants signed up for the study and were screened for inclusion and exclusion criteria. Inclusion criteria consisted of (besides Internet access) mild-to-moderate depressive symptoms defined by scores between 5 and 14 on the Patient Health Questionnaire-9 (PHQ-9) and ages between 18 and 65 years. Participants who fulfilled these criteria were further screened by telephone using the Mini International Neuropsychiatric Interview (M.I.N.I.; Sheehan et al., 1998). Also, a baseline assessment was conducted using several self-report measures (see below). If PHQ-9 scores were above 14, acute suicidality was determined or a lifetime diagnosis of bipolar disorder or schizophrenia was identified in the interview (Meyer et al., 2015), participants were excluded from the study, and professional help was suggested to them. Included participants (*N* = 1013) were then randomized into either an intervention group (IG), in which a CBT-oriented Web-based intervention (Deprexis) was delivered in addition to care as usual (IG; *N* = 509), or into a control group (CG), which solely consisted of care as usual (CG; *N* = 504). During the study, participants in the care as usual group did not receive any Web intervention. The use of other interventions initiated by the participants in the care as usual group was measured during the course of the study. At posttreatment, participants reported having utilized the following treatments during the course of the study: medication, treatment by a psychotherapist, treatment at an outpatient clinic, and treatment at an inpatient clinic. There were no significant differences between participants in the IG and the CG regarding the use of medication ($p = .54$), treatment by a psychotherapist ($p = .38$), treatment at outpatient clinics ($p = .68$), and treatment at inpatient clinics ($p = .29$).

As incentive, all participants were entered into a lottery for 12 iPods after the last assessment. Furthermore, participants in the CG received access to the Web intervention 1 year after baseline assessments.

In addition to the pre- and post-treatment PHQ-9 assessments, which all participants filled out, participants in the IG filled out PHQ-9 assessments every 2 weeks during the course of the study. Furthermore, participants in the IG who had mild symptoms of depression (PHQ-9 scores between 5-9), received the Web-based intervention without any guidance, whereas participants who had moderate depressive symptoms (PHQ scores between 10-14) received the same Web intervention in combination with weekly email support (Klein et al., 2013; Klein et al., 2016). Studies have shown that unguided Web interventions are also effective in the treatment of depression (e.g. Berger, Hämmerli, Gubser, Andersson, & Caspar, 2011). However, considering safety and efficacy, more intensive support seemed appropriate for patients with moderate depression.

After randomization into the IG, participants had to register on the study's website and were then able to use the Web-based intervention (Deprexis) for a period of 12 weeks. It is based on a cognitive-behavioral approach and consists of 10 modules that are presented in the form of a dialogue or "chat." The modules contain classic CBT elements such as behavioral activation, but also broader therapeutic elements such as mindfulness, emotion-focused interventions, and interpersonal skills. Information as well as advice on the application of the relevant concepts in daily life were combined in the modules, which included text, illustration, and audio. This content was presented in dialogue form, where the user was asked to select one of several response options to the program's explanations. In total, several published randomized controlled trials have provided evidence in support of the program's efficacy, typically with small-to-medium effect sizes (Berger, Hämmerli, Gubser, Andersson, & Caspar, 2011; Meyer et al., 2015, Moritz et al., 2012).

In this study, our main interest was to examine the change patterns of participants who received the intervention. Therefore, we focused on participants for whom not only pre- and post-assessments were available, but also several PHQ-9 assessments from during participation. This condition was fulfilled by participants in the IG (PHQ-9 every 2 weeks), but not by participants in the CG (pre and post PHQ-9 assessments only). In total, 483 of the 509 participants randomized into the IG registered on the study's website. A total of 409 participants filled out at least one PHQ-9 during the first 4 weeks of the intervention (assessment at week 2 or assessment at week 4, see flowchart in Figure 1) and were therefore included in our study sample. Participants without any assessment during the intervention were excluded, as no meaningful course of change could be modeled for those cases. Participants (N=409) included in the study, and participants who did not register (N = 26) or did not complete any assessment during the intervention (N = 74) did not differ with regard to age ($F_{2,506}$ = 1.18, p = .31), gender (χ^2_2 = 2.2, p = .34), or initial impairment (PHQ-9 at screening, $F_{2,506}$ = 1.47, p = .23).

On average, 2 weeks passed between screening and registration (standard deviation, SD = 1.36). The first assessment took place 2 weeks after registration. Most participants were recruited by online forums (N = 82), health insurance companies (N = 134), or learned of the study by other means, commonly by news in media (N = 236). Other participants learned of the study while in treatment (N = 57). Attrition (n = 100) did not differ between the recruiting options described previously (χ^2_{12} = 18.0, p = .12).

On average, participants were 43.16 years old (SD = 11.10, range = 18-65) and approximately 70% of participants (287/409) were women. Close to 50% of participants had a high-school diploma qualifying for university entrance (204/409). Most participants (N = 264) suffered from moderate depression and therefore received the Web intervention as well as additional weekly brief support via email. In this study, early response was estimated based on PHQ-9 score changes between Web-based registration and week 4 of the intervention. In addition to the PHQ-9, other impairment measures and participant attitudes were assessed at screening and posttreatment (see questionnaires below).

Measures

Diagnostic Interview

Diagnoses were made using the M.I.N.I. (Sheehan et al., 1998), and clinician-rated severity of depression was assessed with the 24-item version of the Hamilton Depression Rating Scale (HDRS-24). The M.I.N.I. is a short structured diagnostic interview for Diagnostic and Statistical Manual of Mental Disorders, 4th Edition (DSM-IV) and International Classification of Diseases, Tenth Edition (ICD-10) disorders that has been translated into multiple languages. In several studies, it has shown good interrater reliability (e.g. Rossi et al., 2004). Acute suicidality was assessed based on current suicidal ideation and past suicide attempts. In this study, trained raters (postgraduate students) conducted the interviews via telephone. Before they were permitted to rate trial participants, raters were trained to conduct the interview either face-to-face or via telephone modules and had to demonstrate adequate interrater reliability on an audiotaped interview.

Patient Health Questionnaire-9 (PHQ-9)

The PHQ-9 consists of 9 items that reflect the criteria of depression in DSM-IV (Kroenke & Spitzer, 2002). Answers are provided on a 4-point Likert scale (0-“not at all” and 3-“nearly every day”). Thus total scores range from 0-27, with scores between 5 and 9 indicating mild depression and scores between 10 and 14 indicating moderate depression. The instrument has a good test-retest reliability (r_{tt} = .84) and internal consistency (Cronbach alpha = .86-.89; [37]). To operationalize reliable improvement, the reliable change index (RCI), which reflects the pre- post treatment difference ΔRC large enough to not be attributable to measurement error, was calculated following Jacobson and Truax (1991):

where r is the reliability of the PHQ-9 (r = .86) and SD the standard deviation of the PHQ-9 intake score (SD = 2.37). The RCI score for the PHQ-9 was 2.46 total points.

Short-Form Health Survey-12 (SF-12)

The SF-12 assesses limitations in role functioning with 12 items. It consists of two subscales measuring physical health ($\text{SF-12}_{\text{Physical Health Scale}}$) and mental health ($\text{SF-}_{\text{12Mental Health Scale}}$) (Burdine, Felix, Abel, Wiltraut, & Musselman, 2000). Presence and severity of different impairments over the last 4 weeks are rated. Subscale scores can vary between 0-100, with higher scores indicating less impairment. Reliability is good with a Cronbach alpha of .76 (King, Horowitz, Kassam, Yonas, & Roberts, 2005) and test-retest correlations of r_{tt} = .76 for the physical component and r_{tt} = .89 for the mental component (Burdine, Felix, Abel, Wiltraut, & Musselman, 2000).

Questionnaire for the Evaluation of Psychotherapeutic Progress-2 (FEP-2)

The FEP-2 comprises 40 items and measures 4 dimensions of therapeutic progress and outcome (well-being, symptom distress, incongruence, and interpersonal problems; Lutz et al., 2009). Answers are provided on a 5-point Likert scale (1-“never” and 5-“very often”) with higher scores indicating higher impairment. Reliability is high for the global scale (Cronbach alpha = .96; Retest between r_{tt} = .69-.77) and sensitivity to change has been demonstrated [41].

Attitudes Toward Psychological Online Interventions (APOI) Questionnaire

The attitudes toward psychological online interventions (APOI; Schröder et al., 2015)measures attitudes toward online-based interventions with 16 items. The following subscales are assessed: (1) Confidence in Effectiveness, (2) Skepticism and Perception of Risks, (3) Technologization Threat, and (4) Anonymity Benefits. Answers are provided on a 5-point Likert scale (1-“I disagree entirely” and 5-“I agree entirely”) and subscale scores range from 4-20. Higher values on the $\text{APOI}_{\text{total}}$ score indicate a more positive attitude toward psychological online interventions (POI). Reliability is good with a Cronbach alpha of .77.

Adherence

Adherence to the intervention was defined as the extent to which participants used the intervention. A number of modules were calculated by summing up all modules that were accessed for at least ten minutes. Usage time was defined as the number of hours participants spent using the Web intervention. At screening, at registration, every 2 weeks during the 12-week Web intervention period, and after the intervention, participants were asked to fill out the PHQ-9. The number of completed PHQ-9 assessments after week 4 of the intervention was used as an additional indicator of adherence.

Data Analytic Strategy

Patterns of early change in depressive symptoms, measured by the PHQ-9 over the first 4 weeks of the Web intervention, were identified using piecewise growth mixture modeling (PGMM). GMMs are considered a conservative method of identifying early change patterns in comparison with rational definitions such as reliable or clinical significant change (Rubel, et al., 2014). Individual variance of intercepts (intake scores) and slopes (change) are captured in terms of a latent class variable that is added to the growth model (Wang & Bodner, 2007), which allows the identification of subpopulations of participants with similar growth curves. In contrast to conventional growth models, which assume that there is only one underlying population with a single change pattern, GMMs allow the investigation of an a priori unknown number of latent subpopulations, which can differ with regard to intercepts and slopes (in the case of a linear model) as well as class specific variations around these parameters. In GMM, cases with a missing value in the PHQ-9 over the first 4 weeks were not excluded, but rather all available data was used to estimate growth curves within clusters.

In this study, we applied a PGMM, modeling the change pattern as 2 distinct phases (phase 1: time between screening and registration; phase 2: time between registration and assessment at week 4 of the intervention). Therefore, we used a model with 3 latent growth factors: an intercept indicating initial impairment and 2 slopes (one for each phase of change).

To model change before the intervention (phase 1), the first slope loadings, which represent change in phase 1, were fixed to 0 at screening and to 1 at registration and later assessments. To model change during the first 4 weeks of the intervention (phase 2), the second slope loadings, which represent change in phase 2, were fixed to 0 at screening and registration and for the following 2 assessments, the log-linear transformation (base 10) of 2 and 3 were used respectively. According to the Bayesian information criterion (BIC; Schwartz, 1978), the log-linear transformation of factor loadings for the second slope improved the model fit compared with the linear transformation and was therefore used in subsequent analyses (linear: 6826.98, log-linear: 6820.20).

In order to model early response while taking potential spontaneous remission into account, we implemented one categorical latent class factor based on the 3 growth parameters (intercept, first slope, and second slope). Several fit criteria had been discussed to determine the optimal number of latent trajectory classes. In this study, we applied the BIC and the bootstrapped likelihood ratio test (BLRT) as proposed by Nylund et al (2007) to determine the optimal number of latent trajectory classes. Thus, the model determination process was 2-fold. To identify the model with the lowest BIC value, the estimation procedure started with a 1-class solution and then one more class was added in each subsequent run. As mixture models are sensitive to class overextraction, in the second step an additional criterion (BLRT) was used to balance against this potential bias. Once the BIC value no longer decreased from a model with k classes to a model with k+1 classes, this solution was then tested against a solution with k-1 classes using the BLRT. If the BLRT revealed a significant P value ($p<.05$), the model was chosen as the best solution. If, however, the BLRT was not significant, the model was rejected and the solution with one class less ($k-1$) was tested against a model with two classes less ($k-2$). This procedure was repeated until the BLRT resulted in a significant P value.

In the final analysis, we fixed the variances around the class-specific slopes to zero in both phases, whereas intercept variances were freely estimated but constrained to be constant between classes. Therefore, heterogeneity in change had to be captured by the difference in mean slopes of different latent classes completely. Thus, in line with our main interest, we forced the estimation procedure to be more sensitive to patterns of change over time rather than to differences of initial levels of impairment. This approach can be seen as a hybrid of models in which all parameters' variances are fixed to zero (latent class growth models) and in which the free estimation of all parameters is allowed (for similar approaches, also see Hunter, Muthen, Cook, & Leuchter, 2010; Uher et al., 2010).

As the purpose of this study was to evaluate the impact of early change on overall treatment response, we examined the effect of early change patterns on change from pre- to post-treatment in terms of effect sizes as well as reliable change. To evaluate change on the PHQ-9, within-group effect sizes were calculated by subtracting the PHQ-9 score at post from PHQ-9 at screening and dividing the result by the SD of the PHQ-9 score at screening. As described previously, reliable change criteria were applied to the change scores to classify patients into 3 groups: reliably improved (pre to post improvement larger than the RCI of the PHQ-9, which equals 2.46 total points), reliably deteriorated (pre to post deterioration larger than the RCI), and not reliably changed (pre to post change remained under the RCI).

Subsequently, the identified latent change patterns were used to predict outcome and adherence, while controlling for initial impairment (PHQ-9 at screening, HRSD-24 at screening), patient characteristics (FEP-2, SF-12$_{\text{Physical Health Status}}$ and SF-12$_{\text{Mental Health Status}}$), and attitudes toward the online intervention (APOI$_{\text{total}}$) in stepwise regression analysis. Finally, we examined whether initial impairment, patient characteristics, and attitudes toward the online intervention predicted early change patterns using analysis of variances (ANOVAs) and multinomial regression analysis.

6.4. Results

Patterns of Early Change

Following the 2-old model determination process, the 3-class solution showed the best model fit, as suggested by the BIC and BLRT (see Table 1). As a result, the 3-class solution was used for further analyses.

Graphical inspection revealed 2 early response groups and 1 early deterioration group. As shown in Figure 2, patients in classes 1 and 3 were characterized by higher PHQ-9 scores at screening that were above the cut-off score of 9 for clinical samples (C1: M = 12.08, C3: M = 11.27). Class 2 started treatment with lower (mild) depressive symptom severity (C2: M = 8.44).

The first subgroup labeled “early response after registration” (C1: 38.6%, 158/409) showed rapid early decrease in depressive symptom severity after registration. The early change effect size (between screening and week 4) in this group was d = 1.35, reflecting rapid improvement. In the second subgroup which was labeled “early response after screening” (C2: 45.2%, 158/409), depressive symptoms decreased significantly not only during phase 2, but already during phase 1. The early response effect size within this latent class was large (d = 0.98). In contrast to these 2 groups, a third subgroup of participants (C3: 16%, 66/409) showed a significant increase of depressive symptoms from screening to registration and from registration to assessment at week 4. This was the only class with a negative early change effect size (d = −1.78) and was therefore labeled “early deterioration.”

As some participants received additional treatment during the Web intervention and some were provided with email support, we compared these variables between classes to control for differential influences. The number of patients who were in therapy at the beginning of the Web intervention did not differ across classes (χ^2_2 = 4.4, p = .11), and there was no difference between classes in regard to reported change in additional treatment status at the end of the Web intervention (χ^2_{10} = 10.6, p = .39). Also, at the beginning of treatment, there was no difference between classes with regard to the number of patients receiving medication (χ^2_2 = 1.4, p = .50). At the end of treatment, classes did not differ with regard to number of patients reporting change in medication (χ^2_2 = 0.9, p = .64) or use of medication ($\chi^2 2$ =4.1, p = .13). Also there was no difference regarding use of psychotherapy (χ^2_2 = 0.2, p = .92). Only 9 patients reported being treated in outpatient clinics, and only 5 reported being treated in inpatient clinics, so no meaningful difference between classes could be established.

Furthermore, the number of patients receiving email support during the Web intervention differed significantly between classes (p<.001). Whereas almost all participants in C1 (96.2%, 152/158) and C3 (83%, 55/66) received email support, only 31% (57/185) of participants in C2 exceeded the cut-off of 10 on the PHQ-9 at screening and thus received email support. Therefore, email support was included as a predictor variable in the following analyses.

Patterns of Early Change and Treatment Outcome

Table 2 shows the relative frequency of reliable improvement and pre-post effect sizes on the PHQ-9 depending on class membership. The relationship between reliable improvement of depressive symptoms and class membership was analyzed with a chi-square test, which revealed a significant association (χ^2_9= 74.8, p<.001).

As can be seen in Table 2, 62% (99/158) of participants in C1 (early response after registration) showed reliable change (standardized residual = 1.5) and, on average, participants in this group showed the largest pre-post effect size (d = 1.63). Only 3% (5/158) showed a reliable negative development at the end of treatment. In C2 (early response after screening) the rate of reliable improvement (56%, 104/185) and the pre-post effect size (d = 1.25) were slightly lower than in C1. The rate of reliable deterioration was 7% (12/185) for this class. In C3 (early deterioration) only 27% (18/66; standardized residual = −3.00) of participants showed reliable improvement, yet in 39% of cases (26/66; standardized residual = 7.2), participants in this class showed reliable deterioration. This was also the only class with a negative pre-post effect size ($d = -0.47$). All 3 classes had similar rates of lacking reliable change (C1: 34.2%, 54/158; C2: 37.2%, 69/185; C3: 33%, 22/66).

To identify potential relevant predictors beyond impairment at screening (PHQ-9, HRSD-24), we first correlated patient characteristics (FEP-2, SF-12$_{\text{Physical Health Status}}$ and SF-12$_{\text{Mental Health Status}}$) and attitudes toward online interventions (APOI$_{\text{total}}$) with PHQ-9 postscores. Only initial attitudes toward online interventions ($r = -.12$, $p = .01$) were significantly correlated with outcome.

Subsequently, the estimation of the additional predictive power of early change patterns beyond variables at screening was conducted via a stepwise regression analysis. Impairment at screening (PHQ-9) was added into the model first, followed by HRSD-24 and APOI$_{\text{total}}$. The dummy coded class membership variables were added to the model in the last stage of the analysis (see Table 3). The inclusion of the PHQ-9 score at screening explained 7.6% of the variance of PHQ-9 at post (p<.001). Participants with higher scores at the beginning of treatment tended to end with higher scores after treatment. HRSD-24 was also included and significantly increased the amount of explained variance by 3.4% (p<.001). Similarly, participants with higher impairment scores tended to end with higher scores.

The addition of APOI$_{\text{total}}$ significantly increased the amount of explained variance by a further 1.5% ($p = .008$) resulting in a total of 12.5% of explained variance. A higher score at intake, indicating a more positive attitude toward the intervention, was significantly associated with lower PHQ-9 scores after treatment (see Table 3). Email support was not significant and therefore excluded in the next step ($t_{407} = -0.05$, $p = .96$). Adding the dummy coded variables for class membership resulted in a further increase of 21.5% explained variance of treatment outcome ($p \leq .001$). Thus, in total, 34% of variability of PHQ-9 change during the course of treatment was able to be explained by the model that contained initial PHQ-9, HRSD-24, and APOI$_{\text{total}}$ scores, as well as early change patterns (see Table 3).

Patterns of Early Change and Adherence

Adherence assessed via the number of modules and number of assessments (number of completed PHQ-9s) was significantly associated with class membership (see Table 2). A 1-way ANOVA revealed significant associations between mean number of modules of the Web intervention and class membership ($F_{2,406} = 3.65$, $p = .03$). A post hoc test using Bonferroni correction showed that participants in C1 had accessed significantly more modules of the Web intervention than participants in C2 ($d = 0.28$; pooled SDs between clusters were used to calculate between group effect-sizes). Concerning the number of assessments ($F_{3,405} = 6.91$, $p = .001$), similar results were found using Bonferroni corrected P values: Participants in C1 filled out more assessments than participants in C2 ($d = 0.36$) and C3 ($d = 0.44$). There was no significant difference between the subgroups regarding usage time ($F_{2,406} = 2.32$, $p = .10$).

For adherence measured by the number of modules of the Web intervention used and the number of assessments completed, the predictive power of early change patterns was examined using stepwise regression analysis (see Table 3). In the first step of the analysis, impairment at screening (PHQ-9) was included in the equation. It was significantly associated with number of assessments ($F_{1,406} = 15.21$, $p < .001$) and explained 3.6% of variance. HRSD-24 at screening was excluded ($t_{407} = -0.44$, $p = .66$) and neither attitudes toward Web interventions ($t_{407} = -0.03$, $p = .98$) nor email support ($t_{407} = 0.83$, $p = .41$) enhanced predictability. In the second step, class membership was entered in the model. The addition of class membership explained an additional 1.5% of variance of number of assessments ($p = .04$). Thus, a total of 5.1% of variability of number of assessments was able to be explained by the final model, which included PHQ-9 at screening and class membership

Prediction of Early Change Based on Patient Intake Characteristics

Next, we investigated the relationships between class membership, initial impairment, participants' intake characteristics and attitudes toward Web interventions via separate ANOVAs ($APOI_{total}$, $SF\text{-}12_{Physical\ Health\ Scale}$, $SF\text{-}12_{Mental\ Health\ Scale}$, and FEP-2). Using Bonferroni corrected P values, baseline scores on the FEP-2, SF-$12_{\text{Physical Health Scale}}$, and SF-$12_{\text{Mental Health Scale}}$ showed significant relationships with class membership. With regard to the SF-$12_{\text{Physical Health Scale}}$, C3 participants showed significantly lower values than C1 (d = 0.40) and C2 (d = 0.43) participants, indicating a higher level of physical impairment in C3 participants. On the SF-$12_{\text{Mental Health Scale}}$, C2 participants reached significantly higher values than C1 (d = 0.76) and C3 (d = 0.49) participants, indicating that participants in C2 were less mentally impaired than participants in C1 and C3. Impairment measured by the FEP-2 differed significantly between C2 and C1 (d = 0.42) as well as C2 between and C3 (d = 0.44), with C2 showing the lowest values, indicating the lowest level of impairment.

When adding these significant variables and email support to multinomial logistic regressions, depressive symptoms measured by the PHQ-9 ($\chi^2 2 = 75.4$; p<.001) and HRSD-24 ($\chi^2 2 = 34.8$; p<.001) as well as physical health (SF-$12_{\text{Physical Health Scale}}$; $\chi^2 2 = 6.6$; p = .04) demonstrated specific predictive power for class membership. Results of multinomial logistic regression analyses with patient characteristics as predictors of class membership are presented in Table 4.

PHQ-9 intake scores significantly discriminated between classes. Higher scores were associated with a lower probability of belonging to C3 or C2 compared with C1 and a higher probability of belonging to C3 compared with C2. In addition, HRSD-24 intake scores also discriminated between classes with higher scores associated with a lower probability of belonging to C2 compared with C1 and a higher probability of belonging to C3 compared with C2. Higher SF-$12_{\text{Physical Health Scale}}$ intake scores, indicating lower impairment, were associated with a lower probability of membership in C3 compared with C1.

6.5. Discussion

This study examined patterns of early change during the first 4 weeks of a 12-week CBT-oriented Web-based intervention for depression by applying a PGMM analysis. We were able to identify 3 early change patterns: The first was characterized by early improvement after screening, the second by early improvement after registration, and the third by early deterioration. Furthermore, latent classes differed with regard to outcome and adherence measured by the number of assessments (number of completed PHQ-9s) and number of modules used (for a duration of at least ten minutes), but not with regard to the overall amount of time spent using the system. Class membership improved outcome prediction by 21.5% over impairment at intake (PHQ-9 at screening, HRSD-24) and attitudes toward online interventions (APOI). In addition, initial impairment on the PHQ-9 and class membership significantly predicted the number of assessments. Furthermore, group membership of patients was significantly predicted by initial impairment on the PHQ-9 and HRSD-24 as well as by impairment on the SF-12 scale physical health.

The early response and deterioration patterns identified in this study of a CBT-oriented Web-based intervention have also been found in studies of individual face-to-face therapy (Lutz et al., 2014). In this study, a more differentiated investigation of early response was made possible by including the phase from screening to registration in the analysis. The identification of a subgroup that improved before treatment started may be indicative of a regression to the mean or “spontaneous remission” effect for some patients. Yet spontaneous remission may only explain part of the effect in this group. The decision to start treatment and the knowledge of being screened and accepted for the Web intervention may have already created a positive effect by inducing hope and positive treatment expectations, therefore leading to continuous positive changes in outcome, reaching to the end of treatment. Interestingly, participants in this class had significantly less email support than the other 2 classes. Clearly, more studies, which consider the pretreatment phase in the study of early change patterns, are necessary.

In contrast, early response after registration may correspond to a response pattern, which has recently been described as the pliant response pattern (DeRubeis et al., 2014). For this patient group, the impact of the specific treatment is essential: response to treatment is excellent, if the treatment provided is excellent and poor and if the treatment provided is poor. In line with the results of previous studies (Delgadillo et al., 2014; Lutz et al, 2014), participants with early positive change were likely to be improved (reliably) at the end of the treatment and, on average, showed a higher mean effect size than other participants.

In our study, the rate of participants showing early deterioration (16.1%, 66/409) was somewhat higher than in other studies (4.6%; Lutz et al, 2014); 2.4%; Rubel et al, 2014), yet more studies are required before conclusions regarding the risk of deterioration during Web interventions can be drawn. Participants who deteriorate early may be facing crisis and be in need of more immediate help than can be provided by an Web intervention. They may especially benefit from treatment selection and the combination of face-to-face and Web interventions (Lutz, de Jong, & Rubel, 2015). In any case, email support was not lower in this group than in the early response after registration group.

One possible explanation of the mixed findings regarding the frequency of early negative response patterns could be the varying settings, with different early response rates in face-to-face, medication, and Web interventions. However, further studies must also investigate the influence of varying (outcome?) instruments and definitions of early response within this area of investigation (Lutz et al, 2014).

Although email support did not predict outcome, initial PHQ-9 and HRSD-24 scores as well as attitudes toward Web interventions remained significant predictors of outcome after controlling for class membership. The finding of an association between attitudes toward Web interventions and outcome fits well with findings concerning the contribution of treatment expectation to treatment outcome (Greenberg, Constantino, & Bruce, 2006). In Web interventions, one of the first aims should therefore be to promote positive attitudes and motivation with regard to the intervention, making improvement more likely, while preventing dropout.

With regard to adherence, it could be shown that participants with early symptom deterioration completed fewer modules of the intervention and fewer assessments than the early response after registration group. However, class membership predicted number of assessments only. Participants who experience improvement may feel more inclined to track their progress and to make maximum use of the limited time available (12 weeks), whereas participants who show less early improvement may be discouraged from using the intervention more intensively. While early response is often associated with shorter treatment length (Stulz, Lutz, Leach, Lucock, & Barkham, 2007), it has also already been reported that in time-limited treatment protocols, early response participants tend to complete the protocol and are less likely to drop out of treatment (Lutz et al, 2014).

Somewhat surprisingly, there was no difference between early response groups with regard to usage time. A possible explanation could be that the Web intervention was tailored to patients, resulting in individual patients taking different paths within the Web intervention. These paths varied in length, presenting participants with critical problems with more content and longer paths, which took more time.

Physical health was associated with a higher probability of belonging to the early response after registration group compared with the early deterioration group, indicating that physical health may be an important factor not only in face-to-face treatment but also in Web interventions. In addition, Web interventions may be needed, which take poor physical health into account, for example, by providing psychoeducation and/or special coping strategies for patients with symptoms of pain. This could increase adherence by addressing a possibly important concern of some participants who may otherwise feel like the intervention is not adequately targeting their problems (Rozental, Boettcher, Andersson, Schmidt, Carlbring, 2015). Furthermore, poor physical health may decrease motivation and increase negative expectations such as “nothing is going to change” or “I can’t do this” leading to dropout or lack of improvement. In this case, the initiation of motivation and hope may be especially crucial. Similar to face-to-face settings, early change patterns during Web interventions may have important implications for treatment selection, the continuation and adaptation of treatment, as well as the development of new Web or blended interventions. Early response monitoring may support the decision-making process with regard to the addition of special content (eg, coping with physical impairment and enhancing positive treatment expectations) or the necessity of higher intensity treatments. Furthermore, physical health and attitudes toward Web interventions may be important factors that influence early response or early deterioration and may be useful indicators when deciding whether a specific Web intervention should be applied. Although some interventions target multiple problems (Deady, Mills, Teesson, & Kay-Lambkin, 2016), it is still unclear whether such interventions can raise the early positive response rate. Also, given that the participants in the early response after screening and early response after registration groups showed improvement, it may be that varying factors contribute to early response.

Conclusions

Clearly, hope and positive expectations have an impact on early response; however, we don’t yet know much about specific personal characteristics such as self-efficacy. It would be interesting to investigate whether participants that improve or show an early positive change differ with regard to self-efficacy and whether high or low self-efficacy influences outcome in the long-term.

To summarize, more research is still necessary to understand which factors contribute to early response, which factors are indicate risk of early deterioration or dropout, and how clinicians or developers of Web interventions can best adapt interventions, particularly in routine care settings (Deady, Mills, Teesson, & Kay-Lambkin, 2016; Drozd, Vaskinn, Bergsund, Haga, Slinning, & Bjørkli, 2016).

In summary, identifying patterns of early change can have implications for treatment outcome and treatment completion rates. Session-by-session monitoring and feedback of this information may increase awareness of these early change patterns and be applied as part of a stepped-care treatment approach (Lutz, de Jong, & Rubel, 2015).

6.6. Limitations

The following limitations of this study should be considered when interpreting the results. Unfortunately, number of modules of the Web intervention and usage time could not be assessed on a weekly basis, limiting what can be said about the progress of adherence in relation to the progress of symptoms. Also, due to economic considerations, the PHQ-9 was used as the sole outcome measure over the course of treatment. In future studies, a broader range of outcome measures (eg, an anxiety measure) and usage variables could be regularly monitored, improving the estimation and investigation of outcome and adherence. In future studies, additional predictors of adherence should also to be studied.

In addition, only participants with at least one assessment during the intervention were included in the analyses. We addressed this issue by testing for differences between included and excluded participants. Although we did not find any differences, these results should not be generalized to participants who, for whichever reason, did not complete any assessments during the first weeks of treatment. In addition, it must be mentioned that the application of GMM and the associated selection of an optimal number of groups is not without disadvantages (Bauer, 2011; Nagin & Odgers, 2010). One disadvantage is the possibility of specification errors, which can result in the overextraction of trajectory classes through GMM (Bauer, 2007). For this reason, after examining 2 common fit indices (BIC and BLRT), we decided to take the 3-class model into account only. When interpreting early change patterns extracted using GMM, it should not be forgotten that the result is a simplification of a more complex reality, which warrants caution (Bauer, 2011). GMM remains just one possibility to identify early change patterns and other identification possibilities should be considered.

Despite these limitations, this study underlines the potential of early change patterns as predictors of treatment outcome as well as adherence, which, in the future, may guide treatment decisions regarding the content and continuation of Web interventions.

6.7. Figures and Tables

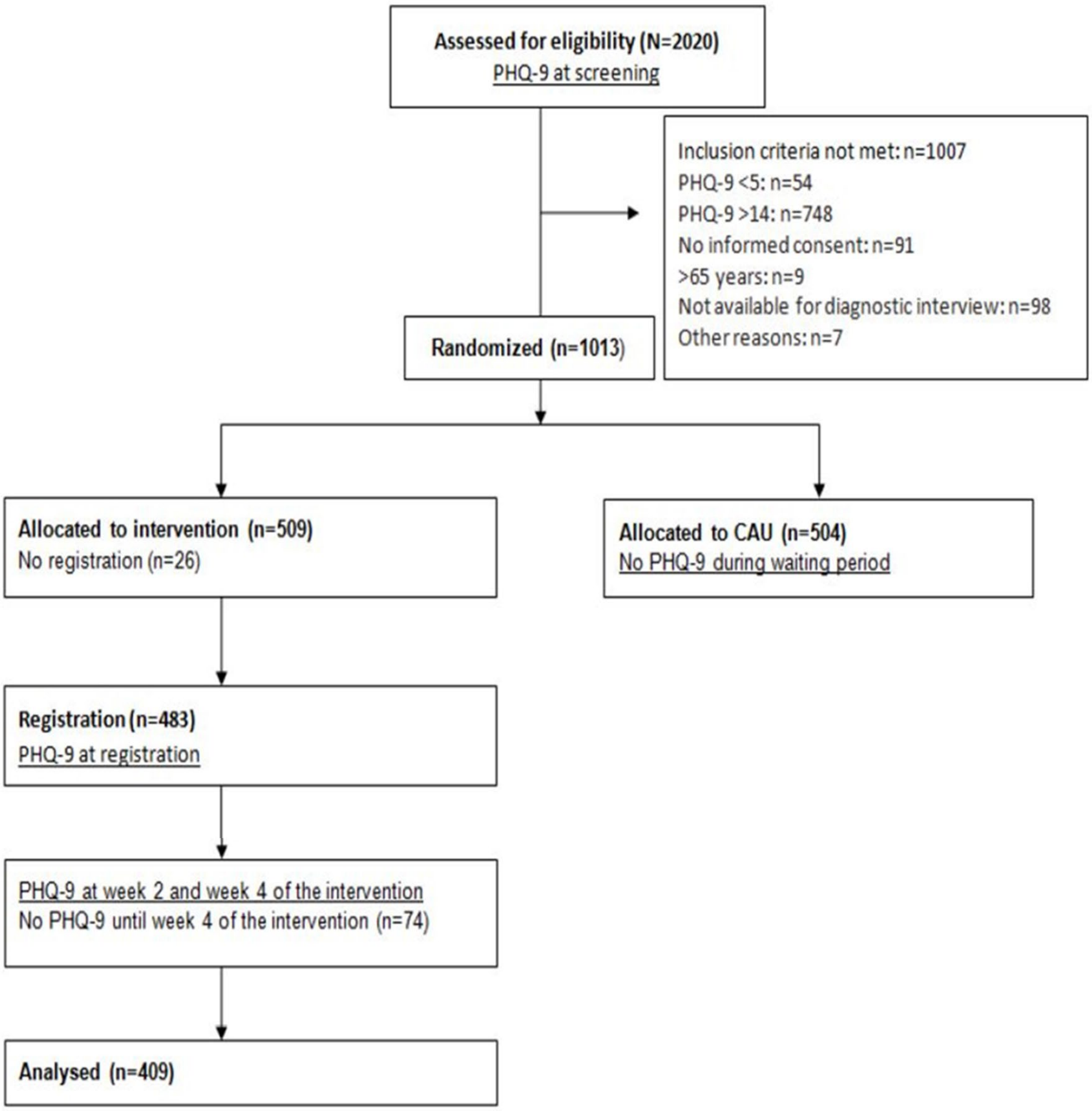

Figure 1. Flowchart of participants.

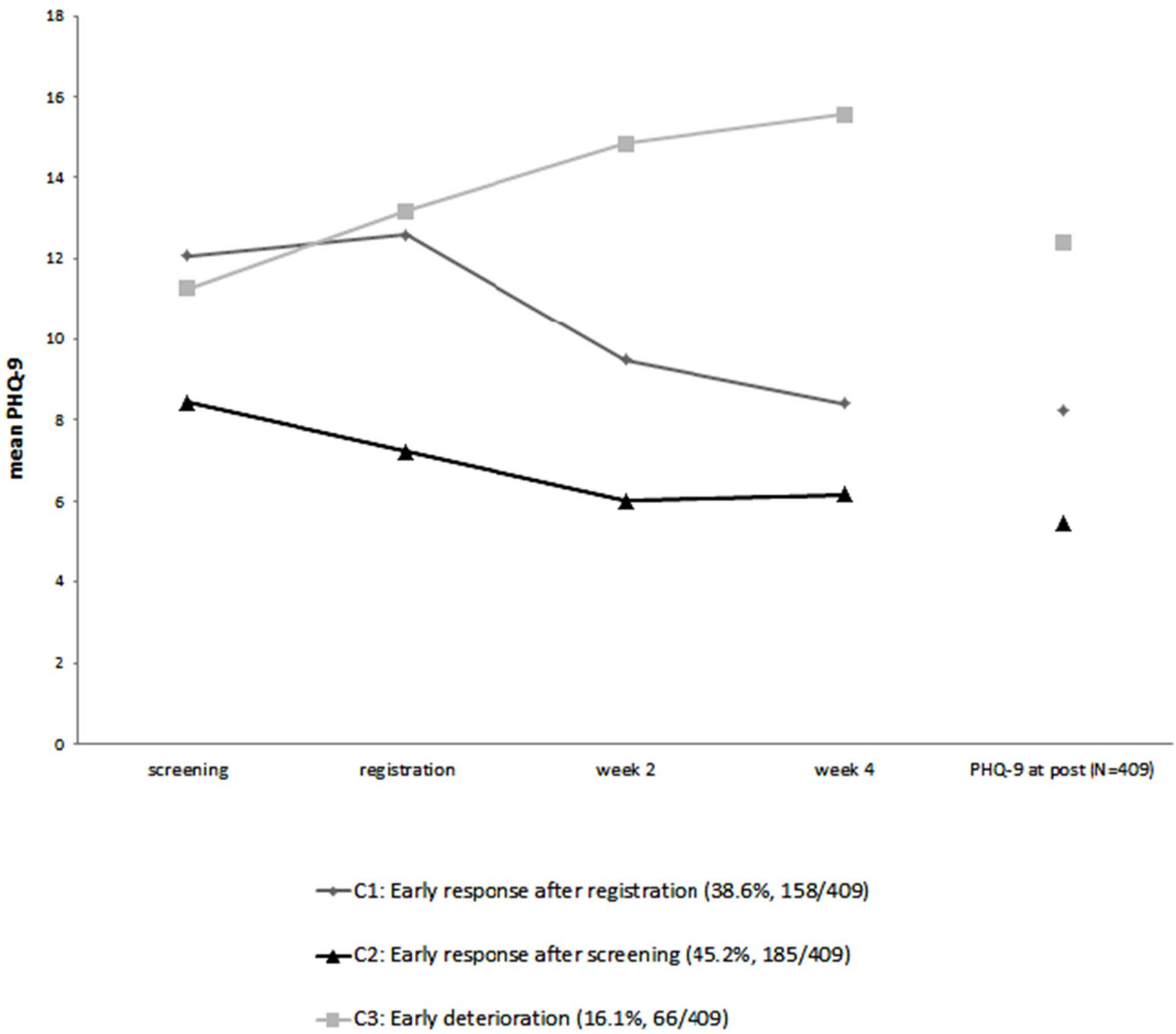

Figure 2. Mean latent growth curves for PGMM solution with three latent classes within the first five weeks and observed mean scores (PHQ-9) in the respective classes after the internet intervention.

Table 1

Information criteria, entropy and p- values in a bootstrapped likelihood ratio test for up to four latent classes in a 2-piece model.

# Classes	BIC[a]	SABIC[b]	AIC[c]	Entropy	BLRT[d] *p* value
1	6855.68	6830.30	6823.57		
2	6782.91	6744.84	6734.75	0.74	<.001
3	6773.41	6722.64	6709.19	0.65	<.001
4	6777.33	6713.87	6697.06	0.66	<.001

Note. BIC = Bayesian Information Criterion; SABIC = Sample Size Adjusted BIC; AIC = Akaike Information Criterion; BLRT = Bootstrapped Likelihood Ratio Test

Table 2

Relative frequencies of reliable change at final treatment outcome, change on the PHQ-9 during treatment (effect sizes) and adherence in patient groups of early change.

		Outcome		Adherence		
		Reliable improvement on PHQ-9	Pre-post ES[b] on PHQ-9	Usage time (in hours)	No. modules intervention	No. assessments
Sample	*N*	*N* (%)	(*d*) (95% CI)	*M* (*SD*)	*M* (*SD*)	*M* (*SD*)
All patients	409	221 (54)	1.12 (0.94-1.30)	7.89 (4.81)	9.10 (4.38)	2.52 (1.25)
Class 1[e]	158	99 (62)	1.63 (1.39-1.86)	8.36 (4.08)	9.84 (3.90)	2.80 (1.16)
Class 2[f]	185	104 (56)	1.25 (1.04-1.47)	7.32 (4.69)	8.64 (4.52)	2.37 (1.25)
Class 3[g]	66	18 (27)	−0.47 (−1.05 to 0.12)	8.33 (6.39)	8.65 (4.85)	2.27 (1.36)
p		<.001[a]	<.001[b]	.10[b]	.03[b]	<.001[b]

Note. ES: Effect size; CI: Confidence interval, PHQ-9: Patient Health Questionnaire, Number assessments: Number of PHQ-9s after week 4.

Class 1: Early response after registration. Class 2: Early response after screening, Class 3: Early deterioration.

[a] χ^2- tests were performed, testing the association between class membership and categorized treatment outcome.

[b] One-way ANOVAs were performed, testing the association between class membership and mean d for pre- to post change, usage time, number of modules of the internet intervention and number of assessments.

Table 3

Stepwise multiple regression analyses predicting outcome and adherence by patient characteristics, email-support, and patterns of early change.

		Outcome			Adherence					
		PHQ-9 at post			No. assessments			No. modules intervention		
Step	Predictor	ΔR^2	Beta	*p*	ΔR^2	Beta	*p*	ΔR^2	Beta	*p*
1		.125		.008	.036		<.001	.024		.002
	PHQ-9		.216	<.001		.190	<.001		.155	.002
	HRSD-24		.199	<.001		−.043	.41		.051	.33
	APOI		−.124	.008		.005	.92		.082	.10
	$\text{Support}_{\text{Email}}$		−.004	.96		.066	.45		.002	.98
2		.215			.015		.04	.006		.31
	PHQ-9		−.040	.50		.192	.007		.146	.04
	HRSD-24		.045	.32		−.024	.66		.068	.22
	APOI		−.129	.002		.001	.98		.080	.11
	$\text{Support}_{\text{Email}}$		−.068	.36		.073	.41		.003	.97
	C2-dummy		−.346	<.001		−.030	.69		−.033	.66
	C3-dummy		.342	<.001		−.132	.01		−.082	.13
Total R^2		.34			.05			.03		
N		409			409			409		

Note. Predictors included impairment on Patient Health Questionnaire-9 (PHQ-9) at screening, impairment on 24-item Hamilton Rating Scale for Depression (HRSD-24) at screening, Attitudes towards Psychological Online Interventions (APOI), email support, and early change patterns. No. assessments: Number of PHQ-9s after week four. C2 and C3 dummy: Dummy coded class membership variable with Class 1 (early response after registration) used as reference class.

Table 4.

Prediction of class membership by patient intake characteristics via multinomial logistic regression analyses.

Variables	Beta (*SE*)	*p*	95% CI for odds ratio Lower	*OR*	Upper
Class 1 vs. class 2					
Intercept	14.84 (2.96)	<.001			
PHQ-9[c]	−1.31 (0.19)	<.001	0.18	0.27	0.39
HRSD-24	−0.12 (0.03)	<.001	0.84	0.89	0.94
FEP-2	0.19 (0.67)	.78	0.32	1.20	4.46
$SF\text{-}12_{Physical\ Health}$	−0.01 (0.02)	.73	0.95	0.99	1.03
$SF\text{-}12_{Mental\ Health}$	0.02 (0.03)	.44	0.97	1.02	1.08
$Support_{Email}$	0.07 (0.65)	.91	0.30	1.08	3.83
Class 1 vs. class 3					
Intercept	3.65 (2.72)	.18			
PHQ-9	−0.26 (0.13)	.04	0.60	0.77	0.98
HRSD-24	0.04 (0.02)	.08	0.99	1.04	1.09
FEP-2	0.37 (0.60)	.54	0.45	1.45	4.68
$SF\text{-}12_{Physical\ Health}$	−0.05 (0.02)	.01	0.92	0.96	0.99
$SF\text{-}12_{Mental\ Health}$	0.01 (0.03)	.70	0.96	1.01	1.06
$Support_{Email}$	−0.77 (0.68)	.26	0.12	0.47	1.75
Class 2 vs. class 3					
Intercept	−11.19 (3.12)	<.001			
PHQ-9	1.05 (0.20)	<.001	2.55	2.87	4.21
HRSD-24	0.16 (0.03)	<.001	1.07	1.18	1.25
FEP-2	0.19 (0.73)	.80	0.29	1.21	5.07
$SF\text{-}12_{Physical\ Health}$	−0.04 (0.02)	.07	0.92	0.96	1.00
$SF\text{-}12_{Mental\ Health}$	−0.01 (0.03)	.70	0.93	0.99	1.05
$Support_{Email}$	−0.84 (0.65)	.20	0.12	0.43	1.56

Note. R^2=.51 (Cox & Snell) and 0.588 (Nagelkerke). Model χ^2_8=284.3. For each comparison, the class mentioned first is used as the reference class in the multinomial logistic regression. PHQ-9: Patient Health Questionnaire. HRSD-24: 24-item Hamilton Rating Scale for Depression. FEP-2: Questionnaire for the Evaluation of Psychotherapeutic Progress-2. SF-12: 12-item Short Form Health Survey. Class 1: Early response after registration. Class 2: Early response after screening. Class 3: Early deterioration.

6.8. References

Andersson, G., & Hedman, E. (2013). Effectiveness of Guided Internet-Based Cognitive Behavior Therapy in Regular Clinical Settings. *Verhaltenstherapie*, *23*(3), 140–148. https://doi.org/10.1159/000354779

Andrews, G., Cuijpers, P., Craske, M. G., McEvoy, P., & Titov, N. (2010). Computer therapy for the anxiety and depressive disorders is effective, acceptable and practical health care: A meta-analysis. PloS One, 5(10), e13196. https://doi.org/10.1371/journal.pone.0013196

Arnberg, F. K., Linton, S. J., Hultcrantz, M., Heintz, E., & Jonsson, U. (2014). Internet-delivered psychological treatments for mood and anxiety disorders: A systematic review of their efficacy, safety, and cost-effectiveness. PloS One, 9(5), e98118. https://doi.org/10.1371/journal.pone.0098118

Barak, A., & Grohol, J. M. (2011). Current and Future Trends in Internet-Supported Mental Health Interventions. Journal of Technology in Human Services, 29(3), 155–196. https://doi.org/10.1080/15228835.2011.616939

Bauer, D. J. (2007). Observations on the use of Growth Mixture Models in psychological research. Multivariate Behavioral Research, 42, 757–786. http://dx.doi.org/10.1080/00273170701710338

Bauer, D. J. (2011). Evaluating individual differences in psychological processes. Current Directions in Psychological Science, 20, 115–118. http://dx.doi.org/10.1177/0963721411402670

Berger, T., Hämmerli, K., Gubser, N., Andersson, G., & Caspar, F. (2011). Internet-based treatment of depression: A randomized controlled trial comparing guided with unguided self-help. Cognitive Behaviour Therapy, 40(4), 251–266. https://doi.org/10.1080/16506073.2011.616531

Bradford, A., Cully, J., Rhoades, H., Kunik, M., Kraus-Schuman, C., Wilson, N., & Stanley, M. (2011). Early response to psychotherapy and long-term change in worry symptoms in older adults with generalized anxiety disorder. The American Journal of Geriatric Psychiatry : Official Journal of the American Association for Geriatric Psychiatry, 19(4), 347–356. https://doi.org/10.1097/JGP.0b013e3181f18061

Burdine, J. N., Felix, M. R., Abel, A. L., Wiltraut, C. J., & Musselman, Y. J. (2000). The SF-12 as a population health measure: An exploratory examination of potential for application. Health Services Research, 35(4), 885–904.

Crits-Christoph, P., Connolly, M. B., Gallop, R., Barber, J. P., Tu, X., Gladis, M., & Siqueland, L. (2001). Early improvement during manual-guided cognitive and dynamic psychotherapies predicts 16-week remission status. The Journal of Psychotherapy Practice and Research, 10(3), 145–154.

Cuijpers, P., van Lier, P. A. C., van Straten, A., & Donker, M. (2005). Examining differential effects of psychological treatment of depressive disorder: An application of trajectory analyses. Journal of Affective Disorders, 89(1-3), 137–146. https://doi.org/10.1016/j.jad.2005.09.001

Deady, M., Mills, K. L., Teesson, M., & Kay-Lambkin, F. (2016). An Online Intervention for Co-Occurring Depression and Problematic Alcohol Use in Young People: Primary Outcomes From a Randomized Controlled Trial. Journal of Medical Internet Research, 18(3), e71. https://doi.org/10.2196/jmir.5178

Delgadillo, J., McMillan, D., Lucock, M., Leach, C., Ali, S., & Gilbody, S. (2014). Early changes, attrition, and dose-response in low intensity psychological interventions. The British Journal of Clinical Psychology, 53(1), 114–130. https://doi.org/10.1111/bjc.12031

DeRubeis, R. J., Cohen, Z. D., Forand, N. R., Fournier, J. C., Gelfand, L. A., & Lorenzo-Luaces, L. (2014). The Personalized Advantage Index: Translating research on prediction into individualized treatment recommendations. A demonstration. PloS One, 9(1), e83875. https://doi.org/10.1371/journal.pone.0083875

Drozd, F., Vaskinn, L., Bergsund, H. B., Haga, S. M., Slinning, K., & Bjørkli, C. A. (2016). The Implementation of Internet Interventions for Depression: A Scoping Review. Journal of Medical Internet Research, 18(9), e236. https://doi.org/10.2196/jmir.5670

Gilbody, S., Littlewood, E., Hewitt, C., Brierley, G., Tharmanathan, P., Araya, R., . . . White, D. (2015). Computerised cognitive behaviour therapy (cCBT) as treatment for depression in primary care (REEACT trial): Large scale pragmatic randomised controlled trial. BMJ (Clinical Research Ed.), 351, h5627. https://doi.org/10.1136/bmj.h5627

Greenberg, R. P., Constantino, M. J., & Bruce, N. (2006). Are patient expectations still relevant for psychotherapy process and outcome? Clinical Psychology Review, 26(6), 657–678. https://doi.org/10.1016/j.cpr.2005.03.002

Gunlicks-Stoessel, M., & Mufson, L. (2011). Early patterns of symptom change signal remission with interpersonal psychotherapy for depressed adolescents. Depression and Anxiety, 28(7), 525–531. https://doi.org/10.1002/da.20849

Hunter, A. M., Muthén, B. O., Cook, I. A., & Leuchter, A. F. (2010). Antidepressant response trajectories and quantitative electroencephalography (QEEG) biomarkers in major depressive disorder. Journal of Psychiatric Research, 44(2), 90–98. https://doi.org/10.1016/j.jpsychires.2009.06.006

Jacobson, N. S., & Truax, P. (1991). Clinical significance: A statistical approach to defining meaningful change in psychotherapy research. Journal of consulting and clinical psychology, 59(1), 12–19. https://doi.org/10.1037/0022-006X.59.1.12

Johansson, R., & Andersson, G. (2012). Internet-based psychological treatments for depression. Expert Review of Neurotherapeutics, 12(7), 861-9; quiz 870. https://doi.org/10.1586/ern.12.63

King, J. T., Horowitz, M. B., Kassam, A. B., Yonas, H., & Roberts, M. S. (2005). The short form-12 and the measurement of health status in patients with cerebral aneurysms:

Performance, validity, and reliability. Journal of Neurosurgery, 102(3), 489–494. https://doi.org/10.3171/jns.2005.102.3.0489

Klein, J. P., Berger, T., Schröder, J., Späth, C., Meyer, B., Caspar, F., . . . Moritz, S. (2013). The EVIDENT-trial: Protocol and rationale of a multicenter randomized controlled trial testing the effectiveness of an online-based psychological intervention. BMC Psychiatry, 13, 239. https://doi.org/10.1186/1471-244X-13-239

Klein, J. P., Berger, T., Schröder, J., Späth, C., Meyer, B., Caspar, F., . . . Hohagen, F. (2016). Effects of a Psychological Internet Intervention in the Treatment of Mild to Moderate Depressive Symptoms: Results of the EVIDENT Study, a Randomized Controlled Trial. Psychotherapy and Psychosomatics, 85(4), 218–228. https://doi.org/10.1159/000445355

Krause, M. S., Howard, K. I., & Lutz, W. (1998). Exploring individual change. Journal of consulting and clinical psychology, 66(5), 838–845. https://doi.org/10.1037/0022-006X.66.5.838

Kroenke, K., Spitzer, R. L., & Williams, J. B. (2001). The PHQ-9: Validity of a brief depression severity measure. Journal of General Internal Medicine, 16(9), 606–613. https://doi.org/10.1046/j.1525-1497.2001.016009606.x

Laurenceau, J.-P., Hayes, A. M., & Feldman, G. C. (2007). Some methodological and statistical issues in the study of change processes in psychotherapy. Clinical Psychology Review, 27(6), 682–695. https://doi.org/10.1016/j.cpr.2007.01.007

Lewis, C. C., Simons, A. D., & Kim, H. K. (2012). The role of early symptom trajectories and pretreatment variables in predicting treatment response to cognitive behavioral therapy. Journal of Consulting and Clinical Psychology, 80(4), 525–534. https://doi.org/10.1037/a0029131

Leykin, Y., Muñoz, R. F., Contreras, O., & Latham, M. D. (2014). Results from a trial of an unsupported internet intervention for depressive symptoms. Internet Interventions, 1(4), 175–181. https://doi.org/10.1016/j.invent.2014.09.002

Lutz, W. (2002). Patient-Focused Psychotherapy Research and Individual Treatment Progress as Scientific Groundwork for an Empirically Based Clinical Practice. Psychotherapy Research, 12(3), 251–272. https://doi.org/10.1080/713664389

Lutz, W., Hofmann, S. G., Rubel, J., Boswell, J. F., Shear, M. K., Gorman, J. M., . . . Barlow, D. H. (2014). Patterns of early change and their relationship to outcome and early treatment termination in patients with panic disorder. Journal of Consulting and Clinical Psychology, 82(2), 287–297. https://doi.org/10.1037/a0035535

Lutz, W., Jong, K. de, & Rubel, J. (2015). Patient-focused and feedback research in psychotherapy: Where are we and where do we want to go? Psychotherapy Research : Journal of the Society for Psychotherapy Research, 25(6), 625–632. https://doi.org/10.1080/10503307.2015.1079661

Lutz, W., Schürch, E., Stulz, N., Böhnke, J. R., Schöttke, H., Rogner, J., & Wiedl, K. H. (2009). Entwicklung und psychometrische Kennwerte des Fragebogens zur Evaluation von Psychotherapieverläufen (FEP). Diagnostica, 55(2), 106–116. https://doi.org/10.1026/0012-1924.55.2.106

Lutz, W., Stulz, N., & Köck, K. (2009). Patterns of early change and their relationship to outcome and follow-up among patients with major depressive disorders. Journal of Affective Disorders, 118(1-3), 60–68. https://doi.org/10.1016/j.jad.2009.01.019

Melville, K. M., Casey, L. M., & Kavanagh, D. J. (2010). Dropout from Internet-based treatment for psychological disorders. The British Journal of Clinical Psychology, 49(Pt 4), 455–471. https://doi.org/10.1348/014466509X472138

Meyer, B., Bierbrodt, J., Schröder, J., Berger, T., Beevers, C. G., Weiss, M., . . . Klein, J. P. (2015). Effects of an Internet intervention (Deprexis) on severe depression symptoms: Randomized controlled trial. Internet Interventions, 2(1), 48–59. https://doi.org/10.1016/j.invent.2014.12.003

Moritz, S., Schröder, J., Meyer, B., & Hauschildt, M. (2013). The more it is needed, the less it is wanted: Attitudes toward face-to-face intervention among depressed patients undergoing online treatment. Depression and Anxiety, 30(2), 157–167. https://doi.org/10.1002/da.21988

Nagin, D. S. & Odgers, C. L. (2010). Group-based trajectory modeling in clinical research. Annual Review of Clinical Psychology, 6, 109–138. http://dx.doi.org/10.1146/annurev.clinpsy.121208.131413

National Institute for Health and Clinical Excellence. NICE. 2011. Common mental health disorders: identification and pathways to care URL: https://www.nice.org.uk/guidance/cg123 [accessed 2017-05-22] [WebCite Cache ID 6qeO5Prfl]

Nordberg, S. S., Castonguay, L. G., Fisher, A. J., Boswell, J. F., & Kraus, D. (2014). Validating the rapid responder construct within a practice research network. Journal of Clinical Psychology, 70(9), 886–903. https://doi.org/10.1002/jclp.22077

Nylund, K. L., Asparouhov, T., & Muthén, B. O. (2007). Deciding on the Number of Classes in Latent Class Analysis and Growth Mixture Modeling: A Monte Carlo Simulation Study. Structural Equation Modeling: A Multidisciplinary Journal, 14(4), 535–569. https://doi.org/10.1080/10705510701575396

Renton, T., Tang, H., Ennis, N., Cusimano, M. D., Bhalerao, S., Schweizer, T. A., & Topolovec-Vranic, J. (2014). Web-based intervention programs for depression: A scoping review and evaluation. Journal of Medical Internet Research, 16(9), e209. https://doi.org/10.2196/jmir.3147

Richards, D., & Richardson, T. (2012). Computer-based psychological treatments for depression: A systematic review and meta-analysis. Clinical Psychology Review, 32(4), 329–342. https://doi.org/10.1016/j.cpr.2012.02.004

Richards, D., Timulak, L., Doherty, G., Sharry, J., Colla, A., Joyce, C., & Hayes, C. (2014). Internet-delivered treatment: Its potential as a low-intensity community intervention for adults with symptoms of depression: Protocol for a randomized controlled trial. BMC Psychiatry, 14, 147. https://doi.org/10.1186/1471-244X-14-147

Rossi, A., Alberio, R., Porta, A., Sandri, M., Tansella, M., & Amaddeo, F. (2004). The Reliability of the Mini-International Neuropsychiatric Interview-Italian Version. Journal of Clinical Psychopharmacology, 24(5), 561–563. https://doi.org/10.1097/01.jcp.0000139758.03834.ad

Rozental, A., Boettcher, J., Andersson, G., Schmidt, B., & Carlbring, P. (2015). Negative effects of internet interventions: A qualitative content analysis of patients' experiences with treatments delivered online. Cognitive Behaviour Therapy, 44(3), 223–236. https://doi.org/10.1080/16506073.2015.1008033

Rubel, J., Lutz, W., Kopta, S. M., Köck, K., Minami, T., Zimmermann, D., & Saunders, S. M. (2015). Defining early positive response to psychotherapy: An empirical comparison between clinically significant change criteria and growth mixture modeling. Psychological Assessment, 27(2), 478–488. https://doi.org/10.1037/pas0000060

Ryan, M. L., Shochet, I. M., & Stallman, H. M. (2010). Universal online interventions might engage psychologically distressed university students who are unlikely to seek formal help. Advances in Mental Health, 9(1), 73–83. https://doi.org/10.5172/jamh.9.1.73

Schibbye, P., Ghaderi, A., Ljótsson, B., Hedman, E., Lindefors, N., Rück, C., & Kaldo, V. (2014). Using early change to predict outcome in cognitive behaviour therapy: Exploring timeframe, calculation method, and differences of disorder-specific versus general measures. PloS One, 9(6), e100614. https://doi.org/10.1371/journal.pone.0100614

Schröder, J., Sautier, L., Kriston, L., Berger, T., Meyer, B., Späth, C., . . . Moritz, S. (2015). Development of a questionnaire measuring Attitudes towards Psychological Online

Interventions-the APOI. Journal of Affective Disorders, 187, 136–141. https://doi.org/10.1016/j.jad.2015.08.044

Schwarz, G. (1978). Estimating the Dimension of a Model. The Annals of Statistics, 6(2), 461–464. https://doi.org/10.1214/aos/1176344136

Sheehan, D. V., Lecrubier, Y., Sheehan, K. H., Amorim, P., Janavs, J., Weiller, E., . . . Dunbar, G. C. (1998). The Mini-International Neuropsychiatric Interview (M.I.N.I.): The development and validation of a structured diagnostic psychiatric interview for DSM-IV and ICD-10. The Journal of Clinical Psychiatry, 59 Suppl 20, 22-33;quiz 34-57.

Stulz, N., Lutz, W., Leach, C., Lucock, M., & Barkham, M. (2007). Shapes of early change in psychotherapy under routine outpatient conditions. Journal of Consulting and Clinical Psychology, 75(6), 864–874. https://doi.org/10.1037/0022-006X.75.6.864

Titov, N. (2011). Internet-delivered psychotherapy for depression in adults. Current Opinion in Psychiatry, 24(1), 18–23. https://doi.org/10.1097/YCO.0b013e32833ed18f

Uher, R., Muthén, B., Souery, D., Mors, O., Jaracz, J., Placentino, A., . . . McGuffin, P. (2010). Trajectories of change in depression severity during treatment with antidepressants. Psychological Medicine, 40(8), 1367–1377. https://doi.org/10.1017/S0033291709991528

van Straten, A., Hill, J., Richards, D. A., & Cuijpers, P. (2015). Stepped care treatment delivery for depression: A systematic review and meta-analysis. Psychological Medicine, 45(2), 231–246. https://doi.org/10.1017/S0033291714000701

Wang, M., & Bodner, T. E. (2007). Growth Mixture Modeling. Organizational Research Methods, 10(4), 635–656. https://doi.org/10.1177/1094428106289397

7. Study II: Identifying change-dropout patterns during an Internet-based intervention for depression by applying the Muthen-Roy model

Arndt, A., Lutz, W., Rubel, J., Berger, T., Meyer, B., Schröder, J., . . . Moritz, S. (2019). Identifying change-dropout patterns during an Internet-based intervention for depression by applying the Muthen-Roy model. *Cognitive Behaviour Therapy*, 1–19. https://doi.org/10.1080/16506073.2018.1556331

Author contributions

A. Arndt was responsible for the concept, the data preparation, data analysis and writing of the manuscript. W. Lutz was responsible for the data management and the final draft. J. Rubel contributed to the data analyses, and revised the manuscript. J. P.Klein was responsible for the study design and the coordination among study centers. He was also responsible for the data management and revised the manuscript. J. Schröder contributed to the study design, developed the attitudes towards psychological interventions questionnaire and contributed to the coordination of study centers as well as the data management. B. Meyer contributed to the study design, was responsible for the cooperation with GAIA AG, which made the intervention available to participants and was also responsible for the data management. C. Späth contributed to the study design and coordination of study centers. T. Berger contributed to the study design, the coordination of study centers and was responsible for the feedback provided to participants.All other authors also contributed to the contributed to the study design, the recruitment of participants, the data collection and the revision of the manuscript. All authors contributed to and have approved the final manuscript.

7.1. Abstract

To date, only few studies have attempted to investigate non-ignorable dropout during Internet-based interventions by applying a NMAR model, which includes missing data indicators in its equations. Here the Muthen-Roy model was used to investigate change and dropout patterns in a sample of patients with mild to moderate depression symptoms (N = 483) who were randomized to a 12-week Internet-based intervention (deprexis, identifier: NCT01636752). Participants completed the PHQ-9 biweekly during the treatment. We identified four change-dropout patterns: Participants showing high impairment, improvement and low dropout probability(C3, $N = 134$) had the highest rate of reliable change at 6- and 12-month follow-up. A further pattern was characterized by high impairment, deterioration and high dropout probability (C2, $N = 32$), another by low impairment, improvement and high dropout probability (C1, $N = 198$). The last pattern was characterized by high impairment, no change and low dropout probability (C4, $N = 119$). In addition to deterioration, also rapid improvement may lead to dropout as a result of a perceived “good enough” dosage of treatment. This knowledge may strengthen sensitivity for the mechanisms of dropout and help to consider its meaning in efforts to optimize treatment selection.

Keywords: change, dropout, Internet-based interventions, NMAR, prediction

7.2. Introduction

Individual differences in treatment change may reflect different change mechanisms and may be clinically meaningful (cf. Kazdin, 2017). Several studies have investigated patterns of change using Growth mixture modeling (GMM) to identify subgroups with similar change patterns (Nordberg, Castonguay, Fisher, Boswell, & Kraus, 2014; Rubel et al., 2015). Based on several measurement points GMM estimate change trajectories as a function of two latent growth factors, a latent intercept (initial impairment) and a latent slope (rate of change). Depending on specific restrictions those latent factors are allowed to vary between and within different latent subpopulations (Gottfredson, Bauer, & Baldwin, 2014). However, when several points of measurement are assessed, often dropout can occur, where participants miss one assessment and do not return later (Yang & Maxwell, 2014). While modern methods like GMM do not require complete data at each measure point, they are based on a missing at random assumption (MAR; (Gottfredson et al., 2014). MAR assumes that, the probability of a dropout depends only on those variables which are observed in the model, e.g. the last observed value or covariates like patients' attitudes towards interventions (Fernandez, Salem, Swift, & Ramtahal, 2015; Schröder et al., 2015). However in some cases, dropout may be non-ignorable, because it is likely to dependent on unobserved variables (cf. Gottfredson et al., 2014). For example, patients may deteriorate or improve, which could result in a feeling the treatment is either not useful or not needed, causing patients to terminate participation (random coefficient dependent missing). In these cases, MAR does not hold, because the probability of dropout is not dependent on observed variables, but depends on the unobserved latent change pattern (Yang & Maxwell, 2014).

These cases illustrate when it is important to consider the mechanisms behind dropout. For that reason, several NMAR models have been introduced and successfully applied (Schafer & Graham, 2002). However only few psychological studies have used NMAR models which remains an important shortcoming, because patients that provide data may differ from those who stop doing so, e.g. by showing deterioration. When this is not considered, it may lead to an overestimation of the treatment effect. Therefore, it is important to investigate missing mechanisms using models that consider non-ignorable dropout. This might be especially important in relation to psychological Internet-based interventions that, while effective (see Andrews et al., 2018) often report high dropout rates (e.g. Fernandez et al., 2015).

In this context, the goal of this study is to investigate change and dropout during an Internet-based intervention. We applied the Muthen-Roy model, which estimates different change trajectories including one latent class variable representing change and also identifies dropout patterns by using another latent class variable that summarizes missing data patterns. In addition, we investigated differences in treatment outcome depending on change-dropout patterns and explored possible predictors of change-dropout patterns. With this study, we hope to contribute to a better understanding of dropout and its mechanisms.

7.3. Methods

Participants and Treatment

This study was registered at ClinicalTrials.gov (identifier: NCT01636752) and was conducted between January 2012 and December 2013. It was approved by the ethics committee of the German Psychological Society (DGPs, reference number SM 04_2012) and used several mental health care settings and media to recruit participants. After informed consent was obtained, participants filled out several self-administered questionnaires and were screened via telephone using the Mini International Neuropsychiatric Interview (M.I.N.I, Sheehan et al., 1998). The inclusion criteria did not comprise a clinical diagnoses, only the presence of mild to moderate depressive symptoms (a total score of 5-14 on the Patient Health Questionnaire-9, PHQ-9) as well an age of between 18 and 65 years. In addition, participants were excluded from the study if acute suicidality or a lifetime diagnosis of bipolar disorder or schizophrenia was reported.

After inclusion and randomization (N = 1013 participants), participants in the intervention group (IG, N = 509) were able to access an Internet-based intervention (Deprexis) for a period of 12 weeks. In total, 483 of the 509 participants randomized into the IG registered on the study's website. The intervention consisted of 10 modules and was primarily based on a cognitive-behavioral approach. The Internet-based intervention included weekly email support for participants with moderate depression as indicated by a PHQ-9 total score between 10-14. Participants in the control group (CG) received the Internet-based intervention after one year. All participants were free to seek out treatment in routine care (treatment as usual). A recent meta-analysis showed that this intervention typically achieves effects of moderate magnitude (i.e., effect size g = 0.54; Twomey, O'Reilly, & Meyer, 2017). For further details on the study protocol, see Klein et al. (2016).

In the IG, the PHQ-9 was assessed every two weeks during the intervention (in-treatment questionnaires), at post (12 weeks after the end of the intervention) and at 6- and 12-month follow-up. In total 222 of 483 (46%) participants terminated the in-treatment questionnaires before week 12 of the intervention. These participants were treated as dropouts in the NMAR analysis. On average participants were 43 years old (SD = 11.07, range = 18-65) and approximately 69% of them (N = 335) were women. Close to 50% of participants had a high-school diploma qualifying for university entrance (N = 236). Most participants (N = 305; 63%) reached PHQ-9 scores above 10 (indicating moderate depression) and therefore received the Internet-based intervention as well as additional weekly brief support via email.

Diagnostic Instruments

As described above, the M.I.N.I was used as part of the diagnostic procedure. It is a short structured diagnostic interview (Sheehan et al., 1998), based on the Diagnostic and Statistical Manual of Mental Disorders, 4th Edition (DSM-IV) and the International Classification of

Diseases, Tenth Edition (ICD-10). Good interrater reliability has been reported (Rossi et al., 2004). The diagnostic interviews were conducted by trained raters who also assessed suicidal ideation and past suicide attempts.

Measures

The following measures were used in the analyzes (for more details, see (Klein et al., 2016).

Patient Health Questionnaire-9 (PHQ-9)

The items of the PHQ-9 (Kroenke, Spitzer, & Williams, 2001) reflect the criteria of depression based on the DSM-IV. Higher scores indicate more severe depressive symptoms (mild: 0-9; moderate: 10-14; severe: 15-27). Items are answered on a four-point Likert scale (0: “not at all” and 3: “nearly every day”). Good test-retest reliability (r_{tt}=.84) has been reported (Kroenke et al., 2001) and in our data, internal consistency was good (Cronbach’s alpha = 0.83).

Attitudes Toward Psychological Online Interventions Questionnaire (APOI)

To measure participant attitudes, the Attitudes Toward Psychological Online Interventions questionnaire (APOI, Schröder et al., 2015) was used. It consists of 16 items that are rated on a 5-point Likert scale (ranging from 1: “I disagree entirely” to 5: “I agree entirely”) and summed up to the following subscales consisting of positive and negative attitudes: Positive attitudes are represented by ‘confidence in effectiveness’ ($APOI_{confidence\ in\ effectiveness}$) and ‘anonymity benefits’ ($APOI_{anonymity\ benefits}$). Negative attitudes are represented by ‘skepticism and perception of risks’ ($APOI_{skepticism,\ perception\ of\ risks}$) and ‘technologization threat’ ($APOI_{technologization\ threat}$). Subscale scores range from 4 to 20, while total score ranges from 16 to 80. A Cronbach’s alpha of .77 has been reported indicating good reliability (Schröder et al., 2015). In our data, a Cronbach’s alpha of .76 was estimated, indicating good reliability.

Short-Form Health Survey-12 (SF-12)

The SF-12 consists of two subscales measuring physical health (SF-12Physical Health Scale) and mental health ($\text{SF-12}_{\text{Mental Health Scale}}$) that are rated over the last four weeks (Burdine, Felix, Abel, Wiltraut, & Musselman, 2000). Presence and severity of different impairments are assessed with subscale higher scores (between 0-100) indicating less impairment. A Cronbach alpha of .76 (King, Horowitz, Kassam, Yonas, & Roberts, 2005) and test-retest correlations of r_{tt}=.76 for the physical component and r_{tt}=.89 for the mental component indicate good reliability (Burdine et al., 2000).

Questionnaire for the Evaluation of Psychotherapeutic Progress-2 (FEP-2)

The FEP-2 measures therapeutic progress and outcome in four dimensions well-being, symptom distress, incongruence, and interpersonal problems (Lutz et al., 2009). It consists of 40 items that are answered on a 5-point Likert scale (1-“never” and 5-“very often”) with higher scores indicating higher impairment. For the global scale sensitivity to change and a high reliability were reported (Cronbach alpha = .96 and Retest between r_{tt} = .69-.77) (Lutz et al., 2009).

Data Analytic Strategy

In a first step to investigate how change progressed when modeled under the assumption that dropout is ignorable (MAR) a growth mixture model (GMM) was conducted. Then we investigated dropout by applying the Muthen-Roy model. As part of a sensitivity analysis Muthén, Asparouhov, Hunter, and Leuchter (2011) recommended comparing different NMAR models based on model fit as indicated by the lowest value of the Bayesian information criterion (BIC, Schwarz, 1978). Therefore, we compared the Muthen-Roy model to a conventional pattern mixture model and a latent class version of a pattern mixture model as proposed by Roy (2007). Subsequently the Muthen-Roy model was compared to the results of a GMM in order to investigate the effect of dropout on parameter estimates. In the following section, a conceptual overview of the applied models is given.

GMM

In GMM latent subpopulations are identified, which can differ with regard to the growth parameters of the model. Model fit as indicated by the BIC was compared for a linear and a loglinear model with the loglinear model showing a better model fit (linear: BIC = 11379,943; loglinear: BIC = 11310,517). Thus for subsequent estimations a loglinear model was used (see Figure 1 for a diagram of the model). The intercept was set at the screening timepoint. The variances of slope and intercept were allowed to vary freely, but were held equal between classes. To determine the number of classes in this study, we applied the BIC and the bootstrapped likelihood ratio test (BLRT) as described by Nylund, Asparouhov, and Muthén (2007). First models with 1 latent class were calculated, then subsequently one class was added to the model until model fit deteriorated. Then the BLRT was used to test the solution against a solution with k-1 classes. At last the 3-class solution showed the best model fit, as suggested by the BIC and BLRT (see Table 1).

NMAR models

In Internet-based interventions dropout may depend on latent change patterns with a negative change pattern leading to higher dropout rates. Thus the MAR assumption is not necessarily fulfilled. When this is the case, it is necessary to specify a distribution for the missing data. For this purpose several different NMAR models have been proposed that are based on a joint distribution of observations and dropout indicators. There are six dropout indicators in this study (*d1*, *d2*, *d3*, *d4*, *d5*, *d6*), which describe the point in time (PHQ-9 at 2, 4…12 weeks in treatment) at which dropout occurred for each participant (see Figure 1). If a participant drops out at the week 2 assessment, the indicators are coded as follows: *d1* = 1, *d2* = 0, *d3* = 0, *d4* = 0, *d5* = 0, *d6* = 0. Participants with intermittent missings are treated like participants with complete data.

In line with the results from the classical GMM which indicate loglinear growth, we used loglinear models here. All models were calculated by Mplus and included the eight assessments as well as the six dropout indicators (see Figures 1-4). No further variables were included in the models. For the Roy and the Muthen-Roy model, the number of classes was determined by comparing model fit where sequentially one class was added to the model until model fit deteriorated.

In conventional pattern mixture models change across time on the PHQ-9 is modeled as a function of time and the interaction of Dropout and Time (for example, see Hedeker & Gibbons, 1997)[1].By means of this procedure, it is possible to examine if and how groups with specific missing data patterns differ with regard to parameter estimates (see Figure 2 for a diagram of the model). For some patterns with missing data model parameters need to be estimated. To achieve this often restrictions to the model are applied which requires assumptions about the unobserved data (Coertjens, Donche, Maeyer, Vanthournout, & van Petegem, 2017). As values at screening and registration were available for every participant, no additional constraints needed to be made to allow for model identification. The variances of intercept and slope were allowed to vary freely within classes, but were held equal between classes. The residual variances of the outcome variables were estimated and allowed to be different across time.

Roy (2007) pointed out that it is difficult to interpret pattern mixture models if many unique patterns emerge, because it is unclear which patterns are meaningful. He proposed a modified pattern mixture model with a latent class variable. The model implies a growth mixture model with the intercept β_{0i} and the slope β_{1i} varying as a function of the latent trajectory class c with k values (see [2]). Dropout indicators are included in the model as covariates of class membership probability (Dantan, Proust-Lima, Letenneur, & Jacqmin-

---

1 $y_{ik} = \text{ß}_0 + \text{ß}_1(\text{Time}_{ik}) + \text{ß}_2(\text{Dropout}_i) + \text{ß}_3(\text{Dropout}_i\text{xTime}_{ir}) + v_{0i} + v_{1i}(\text{Time}_{ir}) + E_{ik}$
withk = time point, i = subject , ß_0= intercept, ß_1 = slope, ß_2 =difference in intercept between dropout and completers, ß_3 = difference in slope between dropout and completers, v_{0i} = deviationintercept, v_{1i} = deviation slope, E_{ik} =model residuals
2 $y_{ti\, ci=k} = \text{ß}_{0i} + \text{ß}_{1i}(\text{Time}_i) + E_{ti}$ with $\text{ß}_{0i\vee ci=k} = \text{ß}_{0k} + v_{0i}$ and $\text{ß}_{1i\vee ci=k} = \text{ß}_{1k} + v_{1i}$

Gadda, 2008), see schematic Figure 3. Again, variances of intercept and slope were allowed to vary freely, but were held equal between classes.

Finally, proposed a further modified pattern mixture model the Muthen-Roy model (Muthén et al., 2011). In addition to a latent variable cc representing change, a latent variable *cd* is introduced that represents dropout and is a function of the dropout indicators only. In this model, slope and intercept vary as a function of latent change patterns and latent dropout patterns[3]. Observed scores are dependent on intercept and slope, which are class specific for change and dropout classes. Missing indicators are not directly related to growth factors. Rather, the class variable representing change patterns and the class variable representing dropout patterns are interrelated (see Figure 4). One limitation of this model is that when a high number of parameters hast to be estimated this may lead to identification problems, especially if there are only few assessments available. Including eight observations a loglinear model was applied and identified. As before, the variance and covariance of the growth factors were allowed to vary freely within classes, but were held equal between classes.

Reliable Change

Furthermore, we were interested in the association between change dropout patterns and overall treatment outcome. To estimate how clinically meaningful change on the PHQ-9 was, we used the reliable change index (RCI, Jacobson & Truax, 1991): For the PHQ-9, the RCI was calculated to amount to RCI = 2.5[4]. For participants that had no measure at post, 6-month follow-up or 12-month follow-up, the last observed value of the PHQ-9 was used.

---

3 $ß_{0i \vee cd_i=k; cy_i=l} ß_{0kl} + v_{0i}$ and $ß_{1i \vee di=k; cyi=l} = ß_{1kl} + v_{1i}$

4 The RCI is defined as the pre-treatment–post-treatment difference ΔRC that is large enough to be not attributable to measurement error and is calculated as:
$\Delta_{RC} = 1.96 \cdot \sqrt{2\left(SD\ \sqrt{1-r}\right)^2} = 1.96 \cdot \sqrt{2\left(2.37 \cdot \sqrt{1-.86}\right)^2} = 2.46$ where *r* is the reliability (internal consistency) of the PHQ-9 and *SD* the standard deviation of the PHQ-9 intake score.

Predictor variable selection procedure

Fifteen baseline patient variables were considered as potential predictors of change-dropout patterns: Impairment as indicated by the PHQ and the FEP, mental and physical health as indicated by the SF-12, education (having a university entrance diploma), marital status (being married), diagnoses of dysthymia and depression as assessed by the M.I.N.I., sickness, using medication, number of depressive episodes (one or more), having previous experience with psychological treatments, internal motivation (versus external reference), daily usage of internet (versus less frequent usage) and employment status (fully employed versus not fully employed). In addition, two treatment related variables, email support (yes/no) and usage of treatment as usual (yes/no) were considered as potential predictors. To identify important predictors, we used a popular algorithm called xgBoost that is implemented in R. It includes a tree learning algorithm and is considered a very good choice for accuracy (Chen & Guestrin, 2016). XgBoost is an implementation of gradient boosting machine, an ensemble of learning algorithms in which the prediction of several base estimators is combined in order to improve robustness over a single estimator. In a first step, categorical predictors were recoded into dummy variables (see variable description above). Then, to increase the likelihood of reproducibility in other data, a training data set was randomly sampled comprising 80% of participants from the full sample. Resampling was conducted in the training data set via the SMOTE algorithm (see Chawla, Bowyer, Hall, & Kegelmeyer, 2002) due to highly unbalanced classes. Following this procedure, xgBoost was applied with tenfold cross-validation that was repeated five times. Model performance was assessed by testing the model on the test data set (see Table 5). Finally, the 10 most important variables (indicated by Gain, a measure of improvement in accuracy brought by a feature) were entered as predictors in the multinomial regression to verify the obtained results.

7.4. Results

Characteristics of Change-Dropout Patterns

Even when using NMAR models, it is difficult to establish whether dropout is non-ignorable (NMAR). We conducted some analyses to gain initial insight into the participants' dropout mechanism. In a clinical study, the post assessment is used to estimate treatment effect. To verify whether participants with or without missings at post differed from each other with regard to PHQ-9 scores at different time points, t-tests were performed. Participants with missings at post had significantly higher PHQ-9 values in the fourth week of treatment. Furthermore to verify whether participants who also dropped out of treatment (here indicated by the completion of less than five out of ten modules) differed with regard to PHQ-9 scores, t-tests were performed. Participants who completed just four modules or less had significantly higher PHQ-9 scores at week four of treatment.

The Muthen-Roy model with 4 classes (2 change classes and 2 dropout classes) showed the best fit as suggested by the BIC (see Table 3). For a graphical depiction of the identified change-dropout patterns, see Figure 5, the different characteristics of change-dropout patterns are shown in Table 4.

Class 1 participants had a low initial PHQ-9 score (C1: $M = 7.98$, $SD = 1.42$) and, compared to participants in other classes, significantly fewer participants in C1 were diagnosed with depression (standardized residual = -3.9, $\chi^2_2 = 42.447$, $p < .001$). Participants in this class improved, but showed a high dropout probability for in-treatment questionnaires: Only 40.4% of participants continued the in-treatment questionnaires until week 10 (see Figure 6). Participants in this class completed an average of seven modules with an average usage time of six hours.

In contrast to C1 participants, participants in class 2 (C2) were more severely impaired at screening and deteriorated over the course of the intervention (C2: $M = 10.16$, $SD = 1.87$). Among C2 participants 43.8% fulfilled the criteria for depression. With a rate of 50% participants in C2 had the highest rate of dysthymia. Like C1 participants, participants in C2 showed high dropout risk for in-treatment questionnaires. Only 37.5% of participants continued until week 10, the rest of participants terminated early with a peak of 18.8% terminating in the fourth week of the intervention (see Figure 6). Participants in this class completed an average of seven modules with an average usage time of six hours.

In class 3 (C3), participants were characterized by rather high PHQ-9 scores at screening (C3: $M = 11.87$, $SD = 1.34$). Rates of depression and dysthymia were lower than in C2, but did not differ significantly from other classes. Participants improved and showed low dropout probability for in-treatment questionnaires: 69.4% of participants continued the in-treatment questionnaires until week 10 of the intervention (see Figure 6). Class 3 participants completed an average of nine modules and had an usage time of eight hours. They completed significantly more modules than C1 and C2 participants ($F_{3,479} = 7.619$, $p < .001$) and showed a significantly higher usage time than C1 participants ($F_{3,479} = 8.112$, $p < .001$).

Participants in class 4 (C4) were characterized by rather high PHQ-9 scores at screening (C4: $M = 12.34$, $SD = 1.30$). In total, the rate of participants with a depressive episode was significantly higher than in other classes (standardized residual = 3.5). Participants showed no change and low dropout probability for in-treatment questionnaires: 63.9% of participants continued the in-treatment questionnaires until week 10 (see Figure 6). On average, C4 participants completed nine modules and had an average usage time of eight hours. C4 participants also completed significantly more modules than C1 and C2 participants ($F_{3,479} = 7.619$, $p < .001$)and showed a significantly higher usage time than C1 participants ($F_{3,479} = 8.112$, $p < .001$).

In addition, we also examined how many participants reported using medication. We found no significant difference between groups ($\chi^2_2 = 4.836$, $p = .186$).

To better estimate the effect of dropout on parameter estimates, the results of a conventional GMM were compared with the results obtained by applying the Muthen-Roy model (see Figures 5 and 7). The estimated overall mean change represented by the slope differed in both models (Muthen-Roy estimate: $s_{mean} = -0.684$, GMM estimate: $s_{mean} = -0.849$). The estimate was smaller in the Muthen-Roy model.

Change-Dropout Patterns in Relation to Long Term Treatment Outcome

At post, 102 (21%) participants, at 6-month follow-up, 119 (25%) participants, and at 12-month follow-up, 146 (30%) participants did not return questionnaires. To examine outcome across different change-dropout patterns, we estimated the rate of participants showing reliable change (see Table 4). Participants in C3 had the highest rate of reliable improvement and differed significantly from the other classes at 6-month (χ^2_2= 89.312, standardized residual = 4.5) and 12-month follow-up (χ^2_2= 92.497, standardized residual= 4.6). In C2, only very few participants reached reliable improvement at 6-month (standardized residual = -3.8) and 12-month follow-up (standardized residual = -3.8). C4 participants were significantly less likely to show reliable improvement at the 12-month follow-up (standardized residual = -1.7).

Prediction of Change-Dropout Patterns based on Intake Characteristics

Initial impairment on the PHQ-9, mental and physical health measured with the SF-12, age, initial impairment on the FEP and attitudes towards Internet-based interventions ($APOI_{confidence\ in\ effectiveness}$, $APOI_{anonymity\ benefits}$, $APOI_{skepticism,\ perception\ of\ risks}$, $APOI_{technologization\ threat}$) as well as email support were the 10 most important variables identified by xgboost (see Tables 5 and 6). To further investigate the resulting predictors, we applied a multinomial regression with these variables as predictor and class membership as dependent variable with C3 (high impairment, improvement, low dropout probability) used as the reference class.

Baseline impairment as measured by the PHQ-9, mental as well as physical health as measured by the SF-12, age and negative attitudes towards Internet-based interventions ($APOI_{technologization\ threat}$) were significant predictors of C3 class membership (see Table 7).

Higher scores on the PHQ-9 were associated with a higher probability to belong to C3 compared to C1 (p <.001) and C2 (p <.001), but with a lower probability to belong to C3 compared to C4 (p = .032). With every unit increase in PHQ-9, the probability of belonging to C1 vs. C3 decreased by a factor of 0.10, and the probability to belong to C2 vs. C3 decreased by a factor of 0.45. With every unit the probability to belong to C4 vs. C3 increased by a factor of 1.276.

Age discriminated significantly between C2 and C3 (p = .022) and between C4 and C3 (p = .044). With one unit increase in age, the probability of belonging to C2 compared to C3 decreased by a factor of 0.94 and the probability to belong to C4 compared to C3 decreased by 0.97.

Perception of technologization threat ($APOI_{technologization\ threat}$) discriminated C3 from C1 (p = .016). Higher skepticism and perception of risk was associated with a higher probability to belong to C1. With every unit increase on $APOI_{technologization}$ threat, the probability of belonging to C1 vs C3 increased by a factor of 1.35.

Physical health discriminated between C2 and C3 (p <.001) as well as between C3 and C4 (p = .020). With each unit increase on physical health scores indicating better physical health the probability to belong to C2 versus C3 decreased by a factor of 0.88 and the probability to belong to C4 compared to C3 decreased by a factor of 0.96.

In addition mental health discriminated between C2 and C3 (p = .004) with higher scores on mental health indicating better mental health decreasing the probability to belong to C2 compared to C3 by a factor of 0.88.

7.5. Discussion

This study investigated non-ignorable dropout by examining change-dropout patterns among 483 participants undergoing a 12-week CBT-oriented Internet-based intervention for depression. Using the Muthen-Roy model, four change-dropout patterns were identified: C1: low impairment, improvement, high dropout probability (*N* = 198); C2: high impairment, deterioration, high dropout probability (*N* = 32); C3: high impairment, improvement, low dropout probability (*N* = 134); C4: high impairment, no change, low dropout probability (*N* = 119). Participants with different change-dropout patterns differed with regard to long term outcomes with C3 participants (high impairment, improvement, low dropout probability) showing the best outcome at all time points. C3 class membership was predicted by initial impairment, age, physical and mental health and attitudes towards Internet-based interventions.

Most participants showed a change-dropout pattern that was characterized by low impairment, improvement and high dropout probability (C1). These participants also completed fewer modules than participants in C3 and C4. In C1, a low rate of participants fulfilled the criteria for a depressive episode. It is possible that mildly impaired participants seek help for a very specific symptom like sleep problems or may need short term help while coping with a difficult situation. These participants might not have been as motivated, possibly because they were less severely affected. Also, because of their low initial impairment, those participants did not receive email-support. Although guided Internet-based interventions report higher treatment completion (e.g. Berger, Hämmerli, Gubser, Andersson, & Caspar, 2011) it is unclear, whether these participants would have improved even more after completing more modules, as they were already only mildly impaired. Furthermore, results at follow-up assessments at 6 and 12 months indicate that improvement in this group reflects lasting changes in participants' impairment.

A small group of participants (C2: high impairment, deterioration, high dropout probability) was also characterized by high dropout probability and relatively high early termination rates already in the fourth week of treatment. Notably, in contrast to C1 participants, more C2 participants fulfilled criteria for depression. Also C2 participants showed the lowest rates of reliable change. This may indicate that a pattern of early deterioration and early dropout is predictive of unfavorable outcomes in the long-term.

Another identified pattern was characterized by high impairment, improvement and low dropout probability (C3). For this group, the positive outcome as indicated by rates of reliable change was also stable at 6- and 12-month follow-ups. A second pattern was also characterized by high initial impairment and low dropout probability (C4), yet participants' symptoms did not improve. In C4 (high impairment, no change and low dropout probability) the rate of participants fulfilling the criteria for a depressive episode was higher than in the other groups. Participants in this group did not seem to have motivation problems concerning participation as they had a low dropout probability and completed a similar number of modules as participants in C3. It seems possible that participants in this group represent a pattern of patients that is at risk of staying depressed, even when in treatment. They may be in need of treatment more specifically tailored to their needs. This includes, but is not limited to, a more intensive treatment, e.g. weekly sessions of face-to-face psychotherapy (van Straten, Hill, Richards, & Cuijpers, 2015).

In contrast to C3 participants, C2 and C4 participants had lower levels of physical health. This may indicate that they may have specific physical health issues, which would make it more difficult to respond to a treatment not targeting these issues. Also, participants in these classes were slightly younger than C3 participants. This is in line with findings that report better treatment outcome for older participants (see Karyotaki et al., 2018), however this does not necessarily imply a causal effect. In comparison to C3 participants, C1 participants (low impairment, improvement, high dropout probability) reported a higher perception of technologization threat. Interestingly, they did not only have a higher probability of dropping out of in-treatment questionnaires, but also completed less modules of the intervention. When combined with low impairment and probably less intense need of treatment, negative attitudes towards Internet-based interventions may promote dropout and hinder treatment completion. At the same time, it is possible that these participants may have received their “good enough” dosage of treatment (Barkham et al., 2006; Stulz, Lutz, Leach, Lucock, & Barkham, 2007). To investigate this further, future trials are required that include repeated measurements for control groups.

In conclusion, in this study, we investigated non-ignorable dropout in Internet-based interventions by examining change-dropout patterns. The findings of this study indicate that dropout in Internet-based interventions may be NMAR, with estimates of change slightly lower when a NMAR model is applied. The results of our study suggest that dropout is linked to negative change, but can also occur when initial impairment is low. Unexpectedly, a non-improving pattern (C4) was characterized by low dropout probability and relatively high treatment completion as indicated by the number of modules. Participants with this pattern had higher rates of depression and therefore might have been in need of more intensive treatment. In addition, our findings suggest that some baseline variables like low physical health may promote the risk for deterioration while negative attitudes may promote dropout. For participants at risk of deterioration or no change, it is advisable to select treatment based on the estimated benefit that a patient will have from a certain treatment (DeRubeis et al., 2014; Lorenzo-Luaces, DeRubeis, van Straten, & Tiemens, 2017; Lutz, Zimmermann, Müller, Deisenhofer, & Rubel, 2017). However, it is not easy to decide on the correct procedure: Even when C1 participants were mildly impaired, it may have been important to them just to have access to the intervention. In a similar vein, it is possible that C4 participants may have deteriorated if they had not been provided with the intervention. It would be of value to explore how patients that do not experience improvement perceive the interventions offered to them. In face-to-face therapies, research on feedback tools has progressed (e.g. Lambert & Shimokawa, 2011; Lutz, Jong, & Rubel, 2015) this research could be transferred to Internet-based interventions, as findings suggest that Internet-based interventions can reach similar effects as face-to-face therapies (e.g. see Carlbring, Andersson, Cuijpers, Riper, & Hedman-Lagerlöf, 2018). In the future, more research using sophisticated prediction models is necessary to improve our ability to identify important therapeutic processes and optimize treatment selection and delivery.

7.6. Limitations

In this study we compared several NMAR models by applying the BIC. However, the BIC cannot account for model selection uncertainty (see Lubke et al., 2017). In addition, for the computation of models as complex as the Muthen-Roy model, a high number of observations is often desirable. Also a comparison of patterns based on two randomized samples (see Muthen & Brown, 2009) would have made it possible to better estimate the effect of the intervention. In general, patterns identified by these types of models are only an approximation and simplification of a far more complex reality.

Moreover, the improvement of C3 participants may be the result of regression to the mean. Furthermore, participants in this study showed little variation regarding attitudes towards Internet-based interventions, which may be due to the self-selected sample. Also, the small number of participants in C2 makes it difficult to identify important characteristics of this high risk group. Further studies with larger samples are necessary.

In general, participants were able to use treatment as usual in addition to the Internet-based intervention so that the treatment effects cannot be attributed to the Internet-based intervention alone. Also, due to ethical considerations, we only included participants with mild to moderate depression as measured by the PHQ-9. It is necessary to examine whether findings can be replicated in a sample of more severely impaired participants. The inclusion of a broader range of measures at multiple time points could provide further insight into the therapeutic process. Despite these limitations, this study contributes to a more nuanced view of the link between change and dropout with implications for analyses procedures. Further, it sheds light on the role of important patient variables such as attitudes towards Internet-based interventions that may affect study participation and treatment progress.

7.7. Figures and Tables

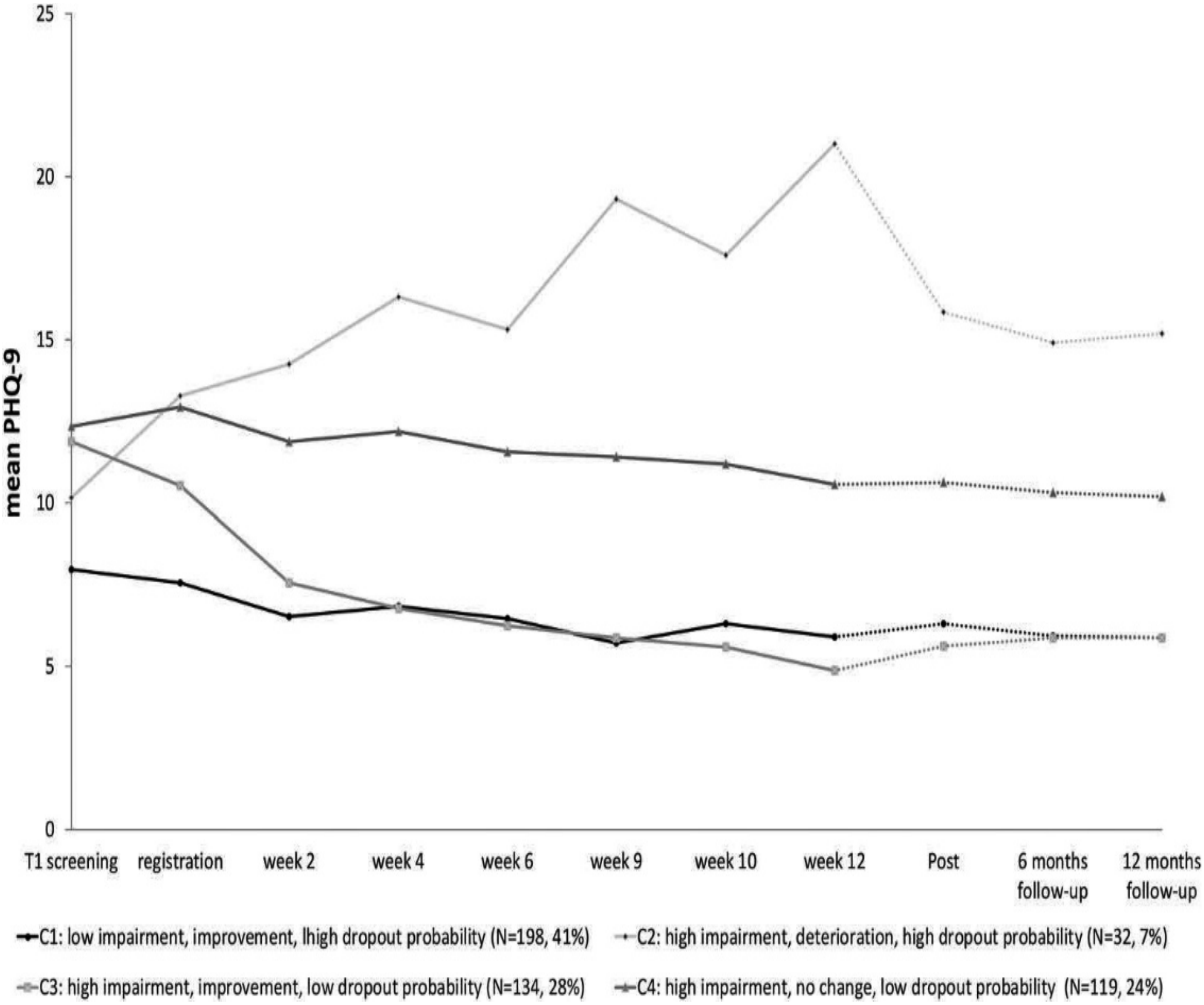

Figure 1. Muthen-Roy model with 4 latent classes: Observed mean scores on the PHQ-9 during the intervention.

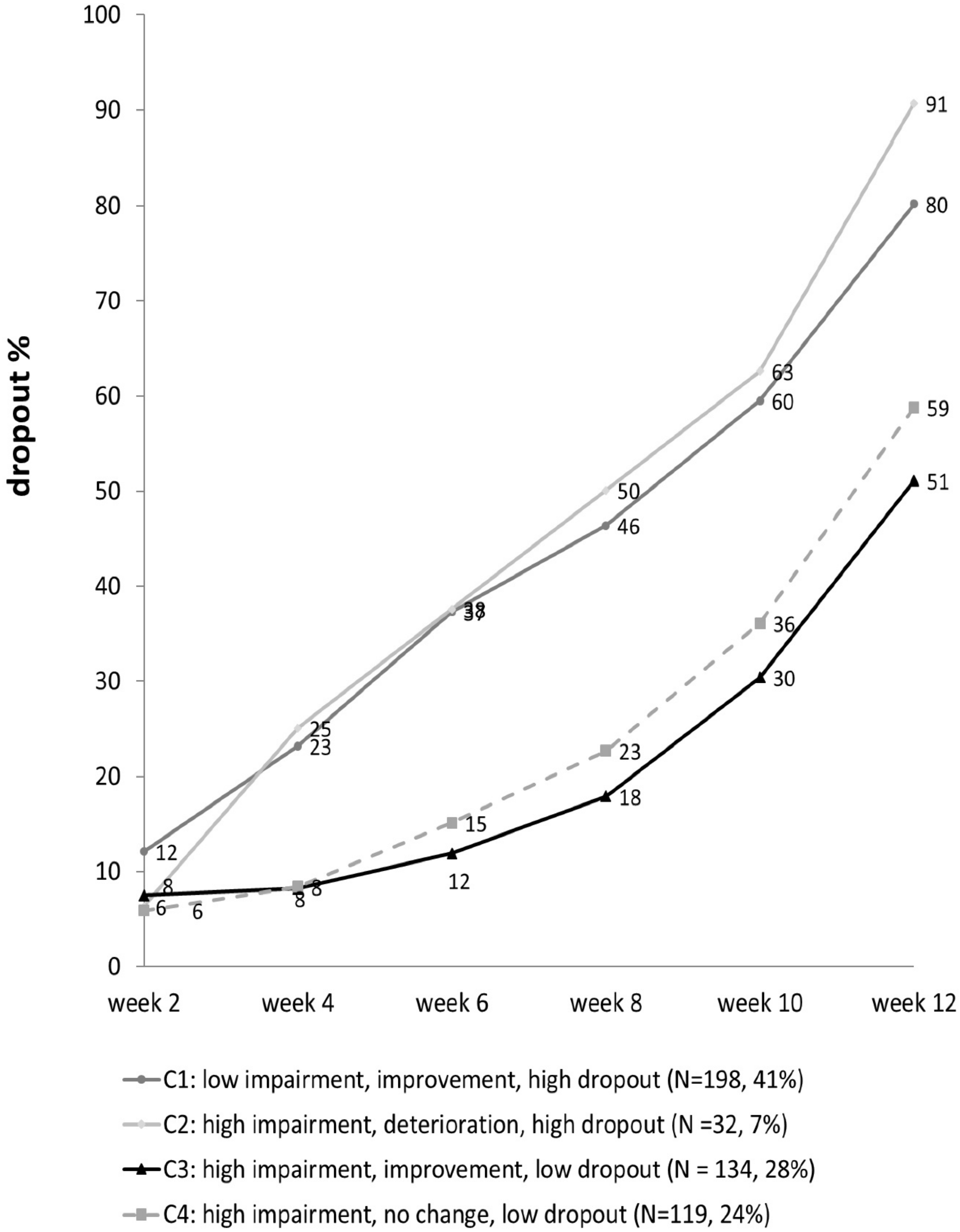

Figure 2. Rates of dropout of in-treatment questionnaires for different change-dropout patterns.

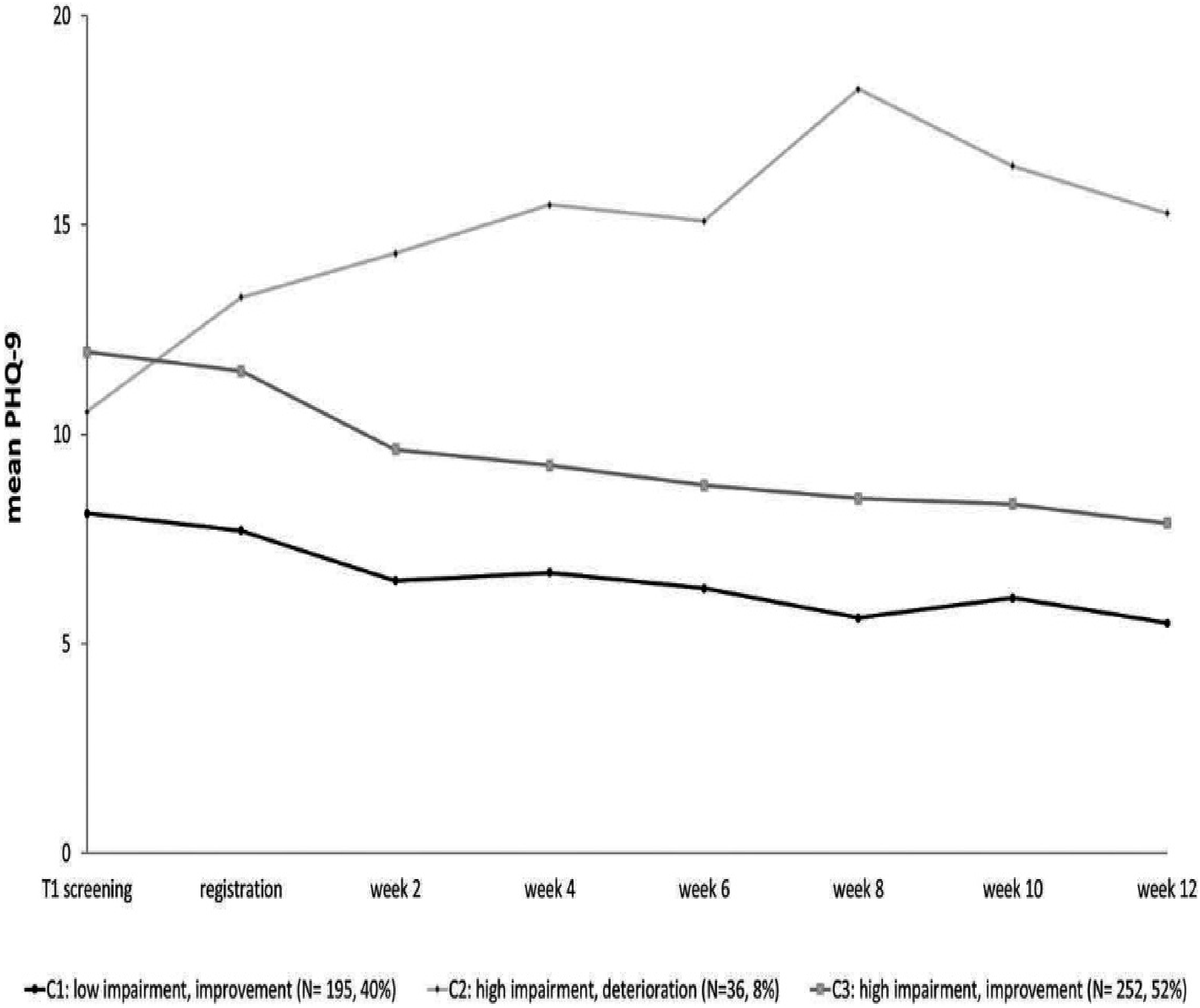

Figure 3. GMM model with 3 latent classes: Observed mean scores on the PHQ-9 during the intervention.

Table 1

Information criteria and entropy for three different NMAR models

Models	BIC	Entropy
Pattern mixture model	11,364.963	-
Roy model	11,311.657	0.652
Muthen-Roy model (2cy, 2cd)	11,295.218	0.618
Muthen-Roy model (2cy, 3cd)	11,327.760	0.751
Muthen-Roy model (3cy, 2cd)	11,302.106	0.808

Note. BIC: Bayesian information criterion. cy: change classes. cd: dropout classes.

Table 2

Rate of depression and dysthymia diagnoses, medication usage, initial impairment and treatment progress depending on different change-dropout patterns

		Diagnoses		Impairment at screening				Progress in treatment	
Pattern	*N*	*N* with depression (%)	*N* with dysthymia (%)	Age *M* (*SD*)	PHQ-9 *M* (*SD*)	SF-12 Physical health scale *M* (*SD*)	SF-12 Mental health scale *M* (*SD*)	No. of modules *M* (*SD*)	Hours of usage *M* (*SD*)
All	483	146 (30.2)	172 (35.6)	42.85 (11.07)	10.28 (2.42)	47.71 (9.26)	31.56 (7.64)	8.35 (4.69)	7.19 (4.89)
C1[a]	198	30 (15.2)	58 (29.3)	43.74 (11.05)	7.98 (1.42)	48.7 (8.73)	34.4 (7.73)	7.40 (4.60)	6.05 (4.41)
C2[b]	32	14 (43.8)	16 (50)	39.66 (11.08)	10.16 (1.87)	45.3 (9.91)	28.9 (5.85)	6.78 (3.60)	6.24 (3.92)
C3[c]	134	45 (33.6)	53 (39.6)	43.82 (10.68)	11.87 (1.34)	47.8 (9.72)	29.3 (7.37)	9.19 (4.51)	7.98 (4.57)
C4[d]	119	57 (47.9)	45 (37.8)	41.14 (11.31)	12.34 (1.30)	46.6 9.28	28.9 (6.43)	9.40 (4.90)	8.45 (5.74)
p		<.001	.058	.05	<.001	0.118	<.001	<.001	<.001

[a]C1: Low impairment, improvement, high dropout probability.

[b]C2: High impairment, deterioration, high dropout probability.

[c]C3: High impairment, improvement, low dropout probability.

[d]C4: High impairment, no change, low dropout probability.

Table 3

Rate of patients who were reliably improved in the PHQ-9 at follow-up assessments depending on change-dropout patterns

Pattern	*N*	Pre-6 months *N* (%) reliably improved	Pre-12 months *N* (%) reliably improved
All	483	280 (58.0)	276 (57.1)
C1[a]	198	102 (51.5)	103 (52.0)
C2[b]	32	2 (6.3)	2 (6.3)
C3[c]	134	117 (87.3)	117 (87.3)
C4[d]	119	59 (49.6)	54 (45.4)
p		<.001	<.001

Note: χ^2 tests were performed.

[a]C1: Low impairment, improvement, high dropout probability.

[b]C2: High impairment, deterioration, high dropout probability.

[c]C3: High impairment, improvement, low dropout probability.

[d]C4: High impairment, no change, low dropout probability.

Table 4

Significant predictors of change-dropout patterns estimated with multinomial logistic regression analyses

				95% CI for odds ratio		
Patterns	Variables	Beta (*SE*)	*p*	Lower	*OR*	Upper
C3[c] vs. C1[a]	PHQ-9 at screening	-2.325 (0.441)	<.001*	0.041	0.097	0.231
	$\text{APOI}_{\text{technologization threat}}$	0.296 (0.136)	.030*	1.028	1.345	1.759
C3 vs. C2[b]	PHQ-9 at screening	-0.794 (0.244)	.001*	0.280	0.452	0.731
	$\text{SF-12}_{\text{physical health scale}}$	-0.127 (0.033)	<.001*	0.825	0.881	0.939
	age	-0.064 (0.028)	.022*	0.889	0.939	0.991
	$\text{SF-12}_{\text{mental health scale}}$	-0.129 (0.045)	.004*	0.804	0.879	0.960
C3vs. C4[d]	PHQ-9 at screening	0.244 (0.114)	.032*	1.021	1.276	1.596
	$\text{SF-12}_{\text{physical health scale}}$	-0.044 (0.019)	.020*	0.920	0.956	0.993
	age	-0.031 (0.016)	.044*	0.940	0.969	0.999

Note: Likelihood ratio test $\chi^{2}=$ 461.59, p <.001; McFadden R^2 = 0.476
[a]C1: Low impairment, improvement, high dropout probability.

[b]C2: High impairment, deterioration, high dropout probability.

[c]C3: High impairment, improvement, low dropout probability.

[d]C4: High impairment, no change, low dropout probability.

7.8. References

Andrews, G., Basu, A., Cuijpers, P., Craske, M. G., McEvoy, P., English, C. L., & Newby, J. M. (2018). Computer therapy for the anxiety and depression disorders is effective, acceptable and practical health care: An updated meta-analysis. *Journal of Anxiety Disorders*, *55*, 70–78. https://doi.org/10.1016/j.janxdis.2018.01.001

Barkham, M., Connell, J., Stiles, W. B., Miles, J. N. V., Margison, F., Evans, C., & Mellor-Clark, J. (2006). Dose-effect relations and responsive regulation of treatment duration: The good enough level. Journal of Consulting and Clinical Psychology, 74, 160–167. https://doi.org/10.1037/0022-006X.74.1.160

Berger, T., Hämmerli, K., Gubser, N., Andersson, G., & Caspar, F. (2011). Internet-based treatment of depression: a randomized controlled trial comparing guided with unguided self-help. Cognitive Behaviour Therapy, 40, 251–266. https://doi.org/10.1080/16506073.2011.616531

Burdine, J. N., Felix, M. R., Abel, A. L., Wiltraut, C. J., & Musselman, Y. J. (2000). The SF-12 as a population health measure: an exploratory examination of potential for application. Health Services Research, 35, 885–904.

Carlbring, P., Andersson, G., Cuijpers, P., Riper, H., & Hedman-Lagerlöf, E. (2018). Internet-based vs. face-to-face cognitive behavior therapy for psychiatric and somatic disorders: an updated systematic review and meta-analysis. Cognitive Behaviour Therapy, 47, 1–18. https://doi.org/10.1080/16506073.2017.1401115

Chawla, N. V., Bowyer, K. W., Hall, L. O., & Kegelmeyer, W. P. (2002). SMOTE: Synthetic Minority Over-sampling Technique. Journal of Artificial Intelligence Research, 16, 321–357. https://doi.org/10.1613/jair.953

Chen, T., & Guestrin, C. (2016). XGBoost. In B. Krishnapuram, M. Shah, A. Smola, C. Aggarwal, D. Shen, & R. Rastogi (Eds.), Proceedings of the 22nd ACM SIGKDD

International Conference on Knowledge Discovery and Data Mining - KDD '16 (pp. 785–794). New York, New York, USA: ACM Press. https://doi.org/10.1145/2939672.2939785

Coertjens, L., Donche, V., Maeyer, S. de, Vanthournout, G., & van Petegem, P. (2017). To what degree does the missing-data technique influence the estimated growth in learning strategies over time? A tutorial example of sensitivity analysis for longitudinal data. PLoS ONE, 12, e0182615. https://doi.org/10.1371/JOURNAL.PONE.0182615

Dantan, E., Proust-Lima, C., Letenneur, L., & Jacqmin-Gadda, H. (2008). Pattern Mixture Models and Latent Class Models for the Analysis of Multivariate Longitudinal Data with Informative Dropouts. The International Journal of Biostatistics, 4. https://doi.org/10.2202/1557-4679.1088

DeRubeis, R. J., Cohen, Z. D., Forand, N. R., Fournier, J. C., Gelfand, L. A., & Lorenzo-Luaces, L. (2014). The Personalized Advantage Index: Translating Research on Prediction into Individualized Treatment Recommendations. A Demonstration. PLoS ONE, 9, e83875. https://doi.org/10.1371/journal.pone.0083875

Fernandez, E., Salem, D., Swift, J. K., & Ramtahal, N. (2015). Meta-analysis of dropout from cognitive behavioral therapy: Magnitude, timing, and moderators. Journal of Consulting and Clinical Psychology, 83, 1108–1122. https://doi.org/10.1037/ccp0000044

Gottfredson, N. C., Bauer, D. J., & Baldwin, S. A. (2014). Modeling Change in the Presence of Non-Randomly Missing Data: Evaluating A Shared Parameter Mixture Model. Structural Equation Modeling : a Multidisciplinary Journal, 21, 196–209. https://doi.org/10.1080/10705511.2014.882666

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

Jacobson, N. S., & Truax, P. (1991). Clinical significance: A statistical approach to defining meaningful change in psychotherapy research. Journal of Consulting and Clinical Psychology, 59, 12–19. https://doi.org/10.1037/0022-006X.59.1.12

Karyotaki, E., Ebert, D. D., Donkin, L., Riper, H., Twisk, J., Burger, S., . . . Cuijpers, P. (2018). Do guided internet-based interventions result in clinically relevant changes for patients with depression? An individual participant data meta-analysis. Clinical Psychology Review, 63, 80–92. https://doi.org/10.1016/j.cpr.2018.06.007

Kazdin, A. E. (2017). Addressing the treatment gap: A key challenge for extending evidence-based psychosocial interventions. Behaviour Research and Therapy, 88, 7–18. https://doi.org/10.1016/j.brat.2016.06.004

King, J. T., Horowitz, M. B., Kassam, A. B., Yonas, H., & Roberts, M. S. (2005). The short form-12 and the measurement of health status in patients with cerebral aneurysms: performance, validity, and reliability. Journal of Neurosurgery, 102, 489–494. https://doi.org/10.3171/jns.2005.102.3.0489

Klein, J. P., Berger, T., Schröder, J., Späth, C., Meyer, B., Caspar, F., . . . Hohagen, F. (2016). Effects of a Psychological Internet Intervention in the Treatment of Mild to Moderate Depressive Symptoms: Results of the EVIDENT Study, a Randomized Controlled Trial. Psychotherapy and Psychosomatics, 85, 218–228. https://doi.org/10.1159/000445355

Kroenke, K., Spitzer, R. L., & Williams, J. B. (2001). The PHQ-9: validity of a brief depression severity measure. Journal of General Internal Medicine, 16, 606–613.

Lambert, M. J., & Shimokawa, K. (2011). Collecting client feedback. Psychotherapy (Chicago, Ill.), 48, 72–79. https://doi.org/10.1037/a0022238.

Lorenzo-Luaces, L., DeRubeis, R. J., van Straten, A., & Tiemens, B. (2017). A prognostic index (PI) as a moderator of outcomes in the treatment of depression: A proof of concept combining multiple variables to inform risk-stratified stepped care models. Journal of Affective Disorders, 213, 78–85. https://doi.org/10.1016/j.jad.2017.02.010

Lubke, G. H., Campbell, I., McArtor, D., Miller, P., Luningham, J., & van den Berg, S. M. (2017). Assessing Model Selection Uncertainty Using a Bootstrap Approach: An update. Structural Equation Modeling : a Multidisciplinary Journal, 24, 230–245. https://doi.org/10.1080/10705511.2016.1252265

Lutz, W., Jong, K. de, & Rubel, J. (2015). Patient-focused and feedback research in psychotherapy: Where are we and where do we want to go? Psychotherapy Research : Journal of the Society for Psychotherapy Research, 25, 625–632. https://doi.org/10.1080/10503307.2015.1079661

Lutz, W., Schürch, E., Stulz, N., Böhnke, J. R., Schöttke, H., Rogner, J., & Wiedl, K. H. (2009). Entwicklung und psychometrische Kennwerte des Fragebogens zur Evaluation von Psychotherapieverläufen (FEP). Diagnostica, 55, 106–116. https://doi.org/10.1026/0012-1924.55.2.106

Lutz, W., Zimmermann, D., Müller, V. N. L. S., Deisenhofer, A.-K., & Rubel, J. A. (2017). Randomized controlled trial to evaluate the effects of personalized prediction and adaptation tools on treatment outcome in outpatient psychotherapy: study protocol. BMC Psychiatry, 17, 133. https://doi.org/10.1186/s12888-017-1464-2

Muthén, B., Asparouhov, T., Hunter, A. M., & Leuchter, A. F. (2011). Growth modeling with nonignorable dropout: Alternative analyses of the STAR*D antidepressant trial. Psychological Methods, 16, 17–33. https://doi.org/10.1037/a0022634

Nordberg, S. S., Castonguay, L. G., Fisher, A. J., Boswell, J. F., & Kraus, D. (2014). Validating the Rapid Responder Construct Within a Practice Research Network. Journal of Clinical Psychology, 70, 886–903. https://doi.org/10.1002/jclp.22077

Nylund, K. L., Asparouhov, T., & Muthén, B. O. (2007). Deciding on the Number of Classes in Latent Class Analysis and Growth Mixture Modeling: A Monte Carlo Simulation Study. Structural Equation Modeling : a Multidisciplinary Journal, 14, 535–569. https://doi.org/10.1080/10705510701575396

Rossi, A., Alberio, R., Porta, A., Sandri, M., Tansella, M., & Amaddeo, F. (2004). The Reliability of the Mini-International Neuropsychiatric Interview-Italian Version. Journal of Clinical Psychopharmacology, 24, 561–563. https://doi.org/10.1097/01.jcp.0000139758.03834.ad

Roy, J. (2007). Latent class models and their application to missing-data patterns in longitudinal studies. Statistical Methods in Medical Research, 16, 441–456. https://doi.org/10.1177/0962280206075311

Rubel, J., Lutz, W., Kopta, S. M., Köck, K., Minami, T., Zimmermann, D., & Saunders, S. M. (2015). Defining early positive response to psychotherapy: An empirical comparison between clinically significant change criteria and growth mixture modeling. Psychological Assessment, 27, 478–488. https://doi.org/10.1037/pas0000060

Schafer, J. L., & Graham, J. W. (2002). Missing data: Our view of the state of the art. Psychological Methods, 7, 147–177. https://doi.org/10.1037//1082-989X.7.2.147

Schröder, J., Sautier, L., Kriston, L., Berger, T., Meyer, B., Späth, C., . . . Moritz, S. (2015). Development of a questionnaire measuring Attitudes towards Psychological Online Interventions–the APOI. Journal of Affective Disorders, 187, 136–141. https://doi.org/10.1016/j.jad.2015.08.044

Schwarz, G. (1978). Estimating the Dimension of a Model. The Annals of Statistics, 6, 461–464. https://doi.org/10.1214/aos/1176344136

Sheehan, D. V., Lecrubier, Y., Sheehan, K. H., Amorim, P., Janavs, J., Weiller, E., . . . Dunbar, G. C. (1998). The Mini-International Neuropsychiatric Interview (M.I.N.I.): the development and validation of a structured diagnostic psychiatric interview for DSM-IV and ICD-10. The Journal of Clinical Psychiatry, 59 Suppl 20, 22-33;quiz 34-57.

Stulz, N., Lutz, W., Leach, C., Lucock, M., & Barkham, M. (2007). Shapes of early change in psychotherapy under routine outpatient conditions. Journal of Consulting and Clinical Psychology, 75, 864–874. https://doi.org/10.1037/0022-006X.75.6.864

Twomey, C., O'Reilly, G., & Meyer, B. (2017). Effectiveness of an individually-tailored computerised CBT programme (Deprexis) for depression: A meta-analysis. Psychiatry Research, 256, 371–377. https://doi.org/10.1016/j.psychres.2017.06.081

Van Straten, A., Hill, J., Richards, D. A., & Cuijpers, P. (2015). Stepped care treatment delivery for depression: a systematic review and meta-analysis. Psychological Medicine, 45, 231–246. https://doi.org/10.1017/S0033291714000701

Yang, M., & Maxwell, S. E. (2014). Treatment effects in randomized longitudinal trials with different types of nonignorable dropout. Psychological Methods, 19, 188–210. https://doi.org/10.1037/a0033804

8. Study III: Outpatient and self-referred participants: Adherence to treatment components and outcome in an internet intervention targeting anxiety disorders

Arndt, A., Rubel, J., Berger, T. & Lutz, W. Outpatient and self-referred participants: Adherence to treatment components and outcome in an internet intervention targeting anxiety disorders. Submitted *Cognitive Behaviour Therapy*, August 09th, 2019.

Author contributions

A. Arndt was responsible for the concept, the data preparation, data analysis and writing of the manuscript. J. Rubel contributed to the study design, contributed to the data analyses, and revised the manuscript. W. Lutz contributed to the study design, the implementation of the study, the data management and revised the manuscript. T. Berger contributed to the study design and was responsible for the development of the intervention as well as the technical support. All authors contributed to and have approved the final manuscript.

8.1. Abstract

While adherence is influencing the effectiveness of internet interventions, few studies report on adherence to treatment components. It remains unclear how adherence and outcome differ between outpatients and self-referred participants. Furthermore, it is important to investigate how outpatients use internet interventions during the waiting time and how they benefit. Adherence to treatment components (relaxation, exposure, cognitive restructuring) and outcome were investigated in participants of an internet intervention targeting anxiety disorders. Outpatients ($N = 50$) were compared to self-referred ($N = 37$) participants and a matched outpatient waitlist sample. Using stepwise regression, adherence to treatment components was investigated as a potential predictor of outcome. Predictors of adherence were investigated using logistic regression analyses. Self-referred participants were more adherent than outpatient participants. Adherence to treatment components was a significant predictor of mean change in anxiety symptoms. Outpatient participants who adhered to relaxation showed greater improvement during the waiting period than outpatients in the outpatient waitlist sample. Self-efficacy, level of education, and motivational incongruence significantly predicted adherence to relaxation in outpatient participants. Adherence to treatment components is associated with outcome and should be fostered, especially if self-efficacy beliefs are low, which may be the case when patients initially seek face-to-face treatment.

Keywords: adherence to treatment components, internet interventions, outpatients, self-referred participants

8.2. Introduction

Internet interventions for anxiety disorders have been found to be effective in numerous studies (see Andrews et al., 2018). Yet not all participants of internet interventions receive the same dosage of treatment as adherence rates vary highly across studies (Beatty & Binnion, 2016). These differences in adherence rates are crucial as higher adherence to internet interventions has been found to be associated with higher outcome (Couper et al., 2010). However, it is not always clear how to best enhance adherence, for example in one study by El Alaoui et al. (2015) more therapist time spent delivering guidance was associated with lower adherence. In addition, most approaches of enhancing adherence focus on providing guidance, which is associated with higher treatment costs and may not always be feasible. Thus, it remains a priority to investigate adherence in internet interventions (Hilvert-Bruce, Rossouw, Wong, Sunderland, & Andrews, 2012).

Also, many studies that have investigated adherence have only used general measures of adherence such as the number of times the website was accessed or the number of sessions completed (e.g. Castro et al., 2018; Couper et al., 2010). However, these general measures may not be the most relevant indicators of adherence, nor do they necessarily imply that the desired outcomes will be achieved (Sieverink, Kelders, & van Gemert-Pijnen, 2017). Instead of this broad definition of adherence, adherence to treatment components should be more closely investigated, as these components are thought to be responsible for treatment change (Domhardt, Geßlein, Rezori, & Baumeister, 2019). Many internet interventions are conceptualized on the basis of CBT components that are deemed to be effective, however most studies do not report adherence to the crucial treatment components. In treatments targeting anxiety disorders, exposure, relaxation, and cognitive restructuring can be considered essential (Borza, 2017). It is likely that some participants experience difficulties conducting these exercises on their own. A more thorough investigation of adherence to these components may indicate how internet interventions could be optimized to increase adherence and, subsequently, outcome.

To date, the identification of consistent predictors of adherence to internet interventions has shown to be difficult (e.g. Castro et al., 2018; El Alaoui et al., 2015; Lutz et al., 2017). A more comprehensive view on adherence to internet interventions could help to better understand what factors affect adherence. In the context of internet interventions, this is especially relevant. Many studies have focused on internet interventions available to the broader public (self-referred participants), but it has been reported that adherence in routine care is less than half of that in research trials (Hilvert-Bruce et al., 2012). Therefore, it remains necessary to investigate how self-referred participants and patients in routine care differ regarding adherence and treatment outcome. In addition, more research is necessary to see, how patients in routine care use internet interventions, and how much they benefit as a result.

In this study, we investigated the impact of participant group (self-referred versus outpatients waiting for face-to face-therapy) on adherence to treatment components and outcome. In addition, we compared outpatient participants to a matched sample of outpatients without access to the internet intervention with regard to change during the waiting period. Finally, we investigated patient variables as potential predictors of adherence to treatment components.

8.3. Methods

Flow of Participants

In this study, an internet intervention was offered to two participant groups: One consisted of outpatients that had registered for a face-to-face therapy in an outpatient clinic and were offered the internet intervention during the waiting period. The second group consisted of interested participants who were recruited by means of advertisements in regional newspapers and the university press. All participants were screened for suicidality via three items: "I have thoughts of ending my life", "During the past seven days, how much were you distressed by thoughts of ending your life", and "In the last week I had thoughts of ending my life". Furthermore, highly depressive symptoms as indicated by a Patient Health Questionnaire-9 (PHQ-9) score of over 21 were an exclusion criterion. As the internet intervention targeted anxiety disorders, only participants who obtained a Generalized Anxiety Disorder Scale-7 (GAD-7) score of 5 or higher were offered the intervention. Outpatients filled out the PHQ-9 and GAD-7 at registration at the clinic, while self-referred participants filled out a screening questionnaire when they registered for the study. All participants that fulfilled the initial inclusion criteria were screened using the Mini International Neuropsychiatric Interview (M.I.N.I, Sheehan et al., 1998). The interviews were conducted by two trained master-level students and seven psychologists in post-graduate clinical training.

In total, 1128 outpatients who registered in the outpatient clinic indicated being interested in taking part in an intervention during the waiting period (see Figure 1). The routinely applied registration questionnaires were used to screen for inclusion and exclusion criteria. After screening for high levels of anxiety (here indicated by a GAD-7 score over 5), excluding outpatients who showed risk of suicidality, and high depressive symptoms (here indicated by a PHQ-9 score over 21) 537 outpatients were contacted and offered information on the study. For 238 outpatients who gave informed consent, an interview appointment was scheduled. The inclusion criteria comprised a diagnosis of panic disorder, social phobia, or generalized anxiety disorder as well as an age of between 18 and 65 years. In addition, outpatients were excluded from the study if acute suicidality or a diagnosis of bipolar disorder or psychosis was reported. After screening for inclusion and exclusion criteria, 86 outpatients with one of the anxiety disorder diagnoses mentioned above were offered the internet intervention.

In total, 104 self-referred participants gave informed consent and filled out the screening questionnaire to participate in the study. After screening level of anxiety (GAD-7 score over 5), suicidality and high level of depressive symptoms (PHQ-9 score over 21), 85 participants were contacted and an appointment for a diagnostic interview was scheduled. Again only participants who fulfilled the criteria of one of the anxiety disorders mentioned above were included. If participants fulfilled reported acute suicidality or fulfilled the criteria of bipolar disorder or psychosis, they were excluded from the study. After screening for inclusion and exclusion criteria, 48 outpatients were allocated to the intervention. After inclusion, participants filled out a pretreatment questionnaire and were then able to access the internet intervention.

Outpatient waitlist sample

151 outpatients did not have access to the internet intervention during the waiting period, fulfilled the described study criteria (PHQ-9 not over 21, GAD-7 over 5), and filled out pre-face-to-face treatment questionnaires. This sample was used as a basis for the matching procedure to select the outpatient waitlist sample.

<u>Intervention</u>

The intervention consisted of eight modules and was primarily based on a cognitive-behavioral approach developed for social anxiety disorder, panic disorder, and generalized anxiety disorder (see also Berger, Boettcher, & Caspar, 2014). The specific content was tailored with regard to the anxiety that was diagnosed with the MINI. The following treatment elements were addressed in the modules: (1) motivational enhancement, (2) psychoeducation and relaxation, (3) cognitive restructuring, (4) self-focused attention and detached mindfulness, (5) exposure and behavioral experiments, (6) summary and repetition, (7) lifestyle modification and problem solving, and (8) repetition and relapse prevention (Berger et al., 2014). The participants were instructed to work with the program for six weeks, with a workload of 1-2 modules per week. If after six weeks the participants wished to continue to use the intervention, they were provided with access to the program for up to another six weeks.

Secure Sockets Layer encryption was used to secure all internet-based communication and participants were identified using anonymous login names and passwords. The study was conducted in compliance with the Declaration of Helsinki and was approved by the local Ethics Committee of the University of Trier.

Participants were informed that they could contact their therapist whenever they wanted to. Once a week, therapists wrote a message with half-standardized supportive feedback to the participants. Three master-level psychology students provided weekly feedback. They received brief training with examples of feedback and were supervised by the first author, a psychologist in post-graduate clinical training. Within the feedback, participants were recognized for making important steps by working with the exercises and motivated to continue treatment. If the participant showed no activity during the past week, participants received a reminder to continue treatment.

Assessments

All participants filled out a questionnaire at registration: Participants that registered at the outpatient clinic received a standardized battery of clinical questionnaires, of which the GAD-7, PHQ-9, and three suicidality items were used for screening. Participants that registered directly for the internet intervention filled out a screening questionnaire (consisting of the PHQ-9, GAD-7, and sociodemographic variables) that was linked to the website containing information on the study and intervention. Following inclusion based on the screening questionnaires and diagnostic interview, all participants filled out a pre-treatment questionnaire consisting of the PHQ-9, the Hopkins Symptom Checklist-11 (HSCL-11), the GAD-7 and the Mini Social Phobia Inventory (Mini-SPIN). Subsequently, participants were asked to fill out in-treatment questionnaires each week during the internet intervention as well as one post-treatment questionnaire after the internet intervention. If they did not fill out the in-treatment or post-treatment questionnaires, they were reminded to do so up to three times.

18 participants did not fill out the pre-treatment questionnaire (NO=14, NSR=4) and a similar number of participants did not log in on the website (NO=15, NSR=3). Eleven participants did not fill out any questionnaires during or after the intervention. Of the remaining 87 participants (NO=50, NSR=37), 58 participants filled out a post-treatment questionnaire.

Diagnostic Instruments

The M.I.N.I is a short structured diagnostic interview (Sheehan et al., 1998). It is based on the Diagnostic and Statistical Manual of Mental Disorders, 4th Edition (DSM-IV) and the International Classification of Diseases, Tenth Edition (ICD-10) and has showed good interrater reliability (Rossi et al., 2004).

Measures

Treatment expectations were assessed before the start of the internet intervention. All symptom measures (GAD-7, Mini-SPIN, HSCL-11, PHQ-9) were assessed before, during, and after the internet intervention. As described above, outpatient participants completed a routine battery of standardized questionnaires at registration at the outpatient clinic. For this study, a subset of these questionnaires was used. Details on these questionnaires are provided below (see Routine measures at the outpatient clinic).

Measures assessed for all participants in the internet intervention

Generalized Anxiety Disorder Scale (GAD-7)

The GAD-7 is an anxiety questionnaire (Löwe et al., 2008) that can be used to screen for generalized anxiety disorder, but can also be used to detect panic disorder or social anxiety disorder. It consists of seven items that reflect the seven core symptoms of generalized anxiety disorder and is rated on a scale from 0 to 3 ("not at all" to "nearly every day"). Good internal consistency have been reported (Cronbachs α = 0.89, Löwe et al., 2008).

Hopkins Symptom Checklist-11 (HSCL-11)

The HSCL-11 (Lutz, Tholen, Schürch, & Berking, 2006) is a modified 11-item version of the Symptom Checklist-90-R (Derogatis, 1994). Questions are answered on a four-point Likert scale ranging from "not at all" to "extremely". The questions focus primarily on depressive and anxious symptoms. The HSCL-11 has been found to have adequate psychometric properties (e.g. Cronbach's α = 0.85; Lutz, Tholen, Schürch, & Berking, 2006).

Mini Social Phobia Inventory (Mini-SPIN)

The Mini-SPIN is the short version of the Social Phobia Inventory (Connor et al., 2000), which measures fear, avoidance, and physiological symptoms. The Mini-SPIN consists of three items assessing avoidance and fear of embarrassment experienced in the past week. Answers are provided on a five-point Likert scale (0 – "not at all", 4 – "extremely"). Good internal consistency and good convergent and discriminant validity have been reported (Wiltink et al., 2017).

Patient Health Questionnaire-9 (PHQ-9)

The PHQ-9 (Kroenke, Spitzer, & Williams, 2001) measures depressive symptoms based on the criteria of depression according to the DSM-IV, with higher scores indicating more severe depressive symptoms. Answers are provided on a four-point Likert scale (0 – "not at all" and 3 – "nearly every day"). The test re-test reliability has shown to be good ($r = .84$; (Kroenke et al., 2001).

Treatment expectations

Participants were able to indicate their expectations regarding treatment on three items: how important it was for them to use the internet intervention (1 – "my life is depending on it", 5 – "it is not important at all"), how convinced they were that the interventions could help them (1 – "not convinced at all", 4 – "very convinced"), and how much they believed they could cope in their daily life after the internet intervention (1 – "very poorly, I will not be able to cope at all", 6 – "very well, as I wish").

Adherence measures

The number of logins and the number of completed sessions were documented. Adherence to the treatment components (exposure, relaxation, and cognitive restructuring) was based on the number of reports in the exposure, relaxation, and 'realistic thought' diary, respectively.

Routine measures at the outpatient clinic

Brief Symptom Inventory (BSI)

The BSI measures self-reported psychological symptoms and was developed based on the SCL-90-R (Franke, 2000). It consists of nine scales (somatization, obsessive-compulsive, interpersonal sensitivity, depression, anxiety, anger-hostility, phobic anxiety, paranoid ideation, and psychoticism). The 53 items are answered on a five-point Likert scale that indicates how strong the impact of symptoms was (1 – "not at all", 5 – "very strong"). For the primary symptom dimensions of the BSI, internal consistencies range between .70 and .89 (Geisheim et al., 2002).

Questionnaire for the Evaluation of Psychotherapeutic Progress (FEP-2)

The FEP-2 consists of 40 items and four scales: well-being, symptoms, interpersonal relationships, and incongruence with respect to approach and avoidance goals. It is used to measure therapeutic progress and has been shown to be reliable and change sensitive (Lutz et al., 2009).

Incongruence Questionnaire – short version (INC-S)

The INC-S assesses the degree of satisfaction with approach and avoidance goals that are particularly relevant for psychotherapy (Grosse Holtforth & Grawe, 2003). It consists of 23 items on two subscales: the approach motivational goals (14 items; e.g. “recently, I’ve been independent”) and the avoidance motivational goals (9 items; e.g. “recently, I’ve been criticized”). Answers are provided on a five-point Likert scale (1 – “not enough” to 5 – “entirely sufficient”). A high score for the sum of motivational goals means that both approach and avoidance motivational goals cannot be met. Cronbach’s alpha ranges between .65 and .86 for the approach and avoidance scales and the sum of motivational goals (Grosse Holtforth & Grawe, 2003).

General Self-Efficacy Scale (GSE)

The GSE consists of 10 items that measure the broad and stable sense of personal competence to deal effectively with a variety of stressful situations (Schwarzer, 1999). The response format is a four-point Likert scale (1 – “not at all true”, 4 – “exactly true”). The GSE scale has been used in numerous studies, where it typically yielded internal consistencies between $\alpha = .75$ and .91 (Scholz, Gutiérrez Doña, Sud, & Schwarzer, 2002).

Data Analytic Strategy

In a first step, we compared adherence between self-referred participants and outpatients using χ^2-tests and t-tests. Then Analysis of covariance (ANCOVA) was used to compare pre-post change scores across participant groups while controlling for pre-scores and number of sessions. For participants for whom post scores were missing (N= 29) the last observed value was used. Within-group effect sizes were calculated for each measure by subtracting the symptom score at post-treatment from the symptom score at pre-treatment and dividing the result by the SD of the pre-scores. The reliable change index was used to estimate how many participants showed reliable change (Jacobson & Truax, 1991)-

Next, we used stepwise regression to investigate whether adherence to active treatment components improved the prediction of change in anxiety symptoms (GAD-7, Mini-SPIN) in comparison to baseline variables. Based on the findings from this stepwise analysis, we selected baseline variables to include in the final model.

Next, outpatient participants were compared to outpatients who did not have access to the internet intervention. To control for sample differences, 151 outpatients were selected who did not have access to the internet intervention and fulfilled the inclusion (GAD-7>5), but not the exclusion criteria (PHQ-9>21), that were applied to outpatient participants. Then a tenfold cross-validated LASSO (least absolute shrinkage and selection operator) was used (Tibshirani, 2011) to identify and select the most important predictors of study participation during the waiting period. The VarImp function was used to rank predictors according to their importance. Based on the ten most important predictors, a sample of outpatients (outpatient waitlist sample) was matched to the sample of outpatient participants. Within the software R, The caret package and the matchit package were used to implement LASSO and the matching procedure, respectively.

As findings regarding predictors of adherence have remained largely inconsistent, we again used a tenfold cross-validated LASSO to identify predictors of adherence to active treatment components. Again, the VarImp function was used to rank predictors according to their importance.

<u>Self-referred and outpatient participants</u>

Before the internet intervention began (pre), participants were highly impaired on all measures (see Table 1). On average, they exceeded the GAD-7 score of 15, which is considered to indicate very high anxiety-related impairment (Löwe et al., 2008). On the Mini-SPIN they exceeded the score of 6, making a diagnosis of social phobia probable (Wiltink et al., 2017). PHQ-9 scores were above 15, indicating high impairment in depressive symptoms (Kroenke et al., 2001). General impairment on the HSCL-11 was also high. Outpatient participants showed significantly higher scores on the PHQ-9 (p = .017) and Mini-SPIN (p = .019). There were no significant differences between self-referred and outpatient participants concerning the frequency of M.I.N.I diagnoses of agoraphobia, panic disorder, social phobia, and generalized anxiety disorder. 56% of participants were female and early 64% of participants had a university entrance diploma. Significantly more self-referred participants (p = .006) had a university entrance diploma. On average, participants were approximately 36 years old (*SD* = 12.70). Self-referred and outpatient participants did not differ significantly regarding age or treatment expectations.

<u>Matching procedure and results</u>

A sample of 151 outpatients was used to identify outpatient characteristics that predicted study participation using LASSO: The matching procedure was then based on the ten most important predictors of study participation (FEP-2, PHQ-9, HSCL-11, incongruence, the BSI subscales phobic anxiety and anxiety, age, level of education, using medication as well as self-efficacy; GSE). Via nearest neighbor (NN) matching, an outpatient waitlist sample was selected (N = 40) that was the most similar to the outpatient participants who had accessed the internet intervention during the waiting period and had completed the pre-face-to-face treatment questionnaires (N = 40).

After the application of NN matching, nearly all baseline variables under consideration were sufficiently well balanced: standardized mean difference scores (smd) ranged from .006 for initial impairment on the INC-S to .162 for using medication with a higher standardized mean difference score for sex only (smd = .27). After NN matching, there was still a significant difference in waiting period, but the groups did not differ on any of the baseline variables (all p >.05).

8.4. Results

Adherence of self-referred and outpatient participants

Adherence to the treatment components (exposure, relaxation and cognitive restructuring) varied highly across participants (see Table 2). Participants showed relatively high average adherence to relaxation, but adherence to cognitive restructuring and exposure was low. Self-referred participants completed significantly more sessions than outpatient participants ($t(85) = -2.56$, $p = .012$). On average, self-referred participants did more relaxation exercises, however this difference was not significant ($t(85) = -1.96$, $p = .054$).

As only 11 outpatient participants (22%) and 16 self-referred participants (43%) indicated any exposure in vivo at all, the frequency of reporting the use of exposure (yes/no exposure) was compared between groups. Self-referred participants were significantly more adherent than outpatient participants with regard to exposure in vivo ($\chi^2_1 = 4.48$, $p = .034$). and cognitive restructuring ($t(46.35) = -3.79$, $p < .001$).

Outcome of self-referred and outpatient participants

41 participants showed reliable change on at least one of the measures with most participants showing reliable change on the GAD-7 ($N = 22$) and the Mini-SPIN ($N = 21$). Self-referred and outpatient participants did not differ significantly in the frequency of reliable change on any of the measures (all $p > .05$). Symptoms on all measures were significantly lower at post-treatment compared to pre-treatment ($t(86) = 3.43$-5.80, all $p <.05$). However, only self-referred participants showed improvement on impairment measured by the HSCL-11 ($t(36) = 3.60$, $p = .001$). Within-group effect sizes ranged from small to medium (see Table 3): For outpatients, effect sizes were small with the exception of improvement on symptoms of social anxiety as measured by the Mini-SPIN ($d = 1.03$). Self-referred participants also achieved the most improvement on symptoms of social anxiety ($d = 0.87$), followed by improvement on anxiety symptoms measured by the GAD-7 ($d = 0.66$). When controlling for pre-scores and number of sessions, a significant difference between groups

was found for symptom change (pre-to post-change score) measured with the HSCL-11 ($F(1) = 4.50$, $p = .037$). Self-referred participants improved more (see Table 3).

Impact of adherence on outcome in all participants

Participants who showed higher adherence to treatment components showed a more positive outcome on the GAD-7. Participants with exposure had a significantly higher mean effect size on the GAD-7 ($M = 0.92$, $SD = 1.12$; $t(85) = -3.08$, $p = .003$) than participants without exposure ($M = 0.24$, $SD = 0.87$). Similarly, participants with at least two relaxation exercises reached a higher mean effect size on the GAD-7 ($M = 0.71$, $SD = 1.14$, $t(85) = -2.86$, $p = .005$) than participants who did less relaxation exercises ($M = 0.13$, $SD = 0.76$). Also, the mean effect size on the GAD-7 was significantly higher for participants who used cognitive restructuring ($M = 0.76$, $SD = 1.13$, $t(70.80) = -2.73$, $p = .008$) compared to those who did not ($M = 0.18$, $SD = 0.80$). Regarding adherence to exposure, more adherent participants also showed higher effect sizes on the HSCL-11. This result was not found on any of the other measures (all $p > .05$).

To investigate predictors of mean change in anxiety symptoms, stepwise regression analyses were used. Here mean change on the GAD-7 was used as the dependent variable. In the first step, we included baseline impairment (pre-scores on GAD-7, PHQ-9, Mini-SPIN, and HSCL-11), which resulted in a significant model ($p = .018$). In the second step, we included pre-treatment variables like age, gender, level of education (operationalized as having versus not having a university entrance diploma) as well as treatment expectations. This did not result in a significant improvement of the model ($p = .678$). Finally, in the last step, we added adherence measures (number of sessions, adherence to relaxation and adherence to cognitive restructuring), which resulted in a significant improvement of the model ($p < .001$). Based on the findings from this first stepwise analysis, we implemented a final model (see Table 4).

The final model significantly predicted mean change in anxiety symptoms as measured by the GAD-7 ($F(11) = 6.51$, $p < .001$). Mean initial impairment on the GAD-7 ($b = 0.26$, $p = .043$) and the HSCL-11 ($b = - 0.56$, $p = .002$) significantly predicted mean change on the GAD-7. While higher impairment on the GAD-7 was associated with more mean change on the GAD-7, a higher impairment on the HSCL-11 was associated with less mean change on the GAD-7. In addition to these baseline impairment variables, number of sessions ($b = 0.09$, $p < .001$), number of exposure exercises ($b = 0.06$, $p < .001$) and number of relaxation exercises ($b = -0.02$, $p = .015$) significantly predicted mean change on the GAD-7. However, while a higher number of sessions and more exposure exercises were related to more improvement, a higher number of relaxation exercises was associated with less improvement .

When applying the same procedure to predict mean change on anxiety symptoms as measured by the Mini-SPIN, only the addition of baseline impairment variables in the first step improved the model significantly ($p = .011$).

<u>Outcome of outpatient participants compared to a matched outpatient waitlist sample</u>

After matching we used one-way-ANCOVA to compare effect sizes of outpatients participants and the outpatient waitlist sample while controlling for the duration of the waiting period. Only change in anxiety symptoms measured by the BSI differed significantly between the two groups with outpatient participants showing significantly more improvement ($p =.015$). As adherence to treatment components was low in outpatient participants, in a second step, we compared those outpatient participants who had shown some adherence to relaxation (at least two relaxation exercises ($N = 23$) to the outpatient waitlist sample. Outpatient participants who showed adherence to relaxation showed significantly higher effect sizes on the GAD-7 ($F(58, 1) = 4.68$, $p = .035$) as well as on the BSI subscales anxiety ($F(58, 1) = 5.91$, $p = .018$) and phobic anxiety ($F(59, 1)= 5.59$, $p = .021$). On the GAD-7, outpatient participants reached an effect size of $d = 0.45$ compared to $d = 0.36$ (outpatient waitlist sample). On the BSI anxiety scale, outpatient participants reached an effect size of d = 0.61

compared to $d = 0.59$. The effect size of outpatient participants on the BSI scale phobic anxiety was $d = 0.46$ compared to $d = 0.26$. There were no significant differences in effect sizes on the HSCL-11 or the PHQ-9.

Prediction of adherence to active treatment components

To investigate predictors of adherence to treatment components, first LASSO regression was used to identify the most important baseline variables. We focused on adherence to exposure in vivo (yes/no) and adherence to relaxation, as these variables were predictive of mean change in anxiety symptoms. As participants with at least two relaxation exercises had a better outcome than participants with less relaxation exercises, this indicator of adherence to relaxation was used as the dependent variable. Demographic variables, treatment expectations, measures of baseline impairment, and participant group (self-referred versus outpatient) were entered as potential predictors. For both dependent variables, the first four most important predictors were the same: Sex, level of education, treatment expectations, and participant status (self-referred versus outpatient). In addition, impairment on the PHQ-9 and the GAD-7 were selected as predictors of adherence to exposure and adherence to relaxation, respectively (see supplemental Table 1). Along with the other selected variables, both impairment measures were included in the logistic regression analysis. Only the model predicting adherence to relaxation reached significance ($\chi^2_5 = 14.53$, $p = .034$). Adherence to relaxation was significantly predicted by level of education ($b = 1.21$, $p = .022$) and initial impairment on the GAD-7 ($b = 1.12$, $p = .037$). The probability of reporting relaxation at least twice increased, if the participant had a university entrance diploma and higher initial impairment on the GAD-7.

For outpatient participants, additional baseline variables collected during registration at the outpatient clinic were available . To investigate whether any of these variables predicted adherence to active treatment components for outpatient participants, again, LASSO was used. As too few outpatient participants reported having done exposure, only predictors of adherence to relaxation were investigated. The most important predictors identified using LASSO were level of education, initial impairment on the FEP-2, incongruence as measured by the INC-S, self-efficacy as measured by the GSE, and treatment expectations (see supplemental Table 2). When entered into logistic regression, incongruence (b = -2.56, p = .016), self-efficacy (b = 2.49, p = .016), and level of education (b = 1.53, p = .022) significantly predicted adherence to relaxation (see Supplemental Table 3 for more details).

8.5. Discussion

In this study, we investigated differences in adherence in outcome and to treatment components between self-referred and outpatient participants. In addition, we compared outpatient participants to a matched outpatient waiting sample. Finally we investigated potential predictors of adherence to treatment components.

Adherence varied across treatment components with relatively high adherence to relaxation and low adherence to exposure and cognitive restructuring. Results on the association between adherence and outcome were mixed: While self-referred participants showed higher adherence than outpatient participants, these two groups did not differ significantly in terms of outcome. At the same time, some results did point to an association between adherence to treatment components and outcome: In general, participants showing higher adherence to active treatment components also achieved better outcomes. Compared to outpatients who did not have access to the internet intervention, only outpatient participants who adhered to relaxation showed greater improvement during the waiting period.

In addition to initial impairment, adherence to relaxation and exposure added to the prediction of mean change in anxiety symptoms as measured by the GAD-7. These findings underlie the importance of adherence to treatment components and fits to the findings of other studies, which also reported an association between adherence and outcome (El Alaoui et al., 2015; Lutz et al., 2017). Surprisingly, although outcome was better for participants who had reported at least two relaxation exercises, the number of reported relaxation exercises was negatively associated with mean change in anxiety symptoms. The only other study that has investigated adherence to relaxation in internet interventions, found the number of relaxation exercises to be positively associated with reliable improvement (Alfonsson, Olsson, & Hursti, 2016). However, this study did not focus on participants suffering primarily from anxiety. With regard to anxiety symptoms, it has been noted that in some cases, relaxation can lead to an increase in symptoms (Newman, Lafreniere, & Jacobson, 2018). For example, too much relaxation may be contraindicated when used as a means of avoidance. Thus for some participants, adherence to relaxation should be more closely monitored in internet interventions targeting anxiety.

Besides initial impairment on the GAD-7, level of education was the only other significant predictor of adherence to relaxation in all participants. A higher level of education was associated with higher adherence. However, research findings regarding level of education and adherence are mixed (Beatty & Binnion, 2016), so no clear conclusion can be drawn yet. Remarkably, after including more variables to predict adherence to relaxation for outpatients, self-efficacy was the predictor with the most influence on adherence to relaxation. To the best of our knowledge, only few studies have examined the role of self-efficacy on adherence (e.g. Al-Asadi, Klein, & Meyer, 2014; El Alaoui et al., 2015; Hebert, Vincent, Lewycky, & Walsh, 2010; Wagner et al., 2015). Only one study found a positive association between self-efficacy and adherence (Wagner et al., 2015). This could indicate that participants with low self-efficacy may be especially at risk to show less adherence and achieve poor outcome. Differences in self-efficacy could be one potential explanation of the differences in adherence found between outpatient and self-referred participants. Therefore, future studies should consider self-efficacy when investigating adherence.

In sum, our findings suggest adherence to treatment components has the potential to shed more light on the process of change during internet interventions. While internet interventions have been implemented in various settings, there is not enough research that systematically investigates differences between participant populations regarding adherence and patient characteristics. Only when important patient variables such as self-efficacy are identified can clinical implications regarding optimal allocation of participants be drawn.

8.6. Limitations

There are several limitations, which need to be considered. One limitation is the small sample size limiting the generalizability of the findings as well as power to detect predictor variables and differences between groups. Furthermore, the indicators of adherence used here are only proxies, as it is possible that participants used the treatment components without reporting so in their dairy. Also, the rate of participants showing reliable change was quite low in this study. One explanation could be the low adherence of participants to exposure.

Another limitation is the study design. A dismantling design would have been more appropriate to study the impact of adherence to different treatment components as the use of treatment components was intercorrelated. In addition, no control group with repeated measurements was available, so no causal inference regarding treatment outcome can be drawn. The usage of additional interventions was not controlled, so it is possible that participants also completed other treatments that were available to them in routine care. Also, for economic reasons, not all measures were available for self-referred compared to outpatient participants, so the effect of self-efficacy could only be studied for outpatient participants. Furthermore, no follow-up was conducted, so inferences regarding long-term treatment effects cannot be made. The reported findings are preliminary, future studies should include larger sample sizes and participants from various settings, to investigate differences in adherence between various participant populations. In addition, specific measures of adherence and potentially important predictor variables such as self-efficacy should be investigated in more detail.

8.7. Figures and Tables

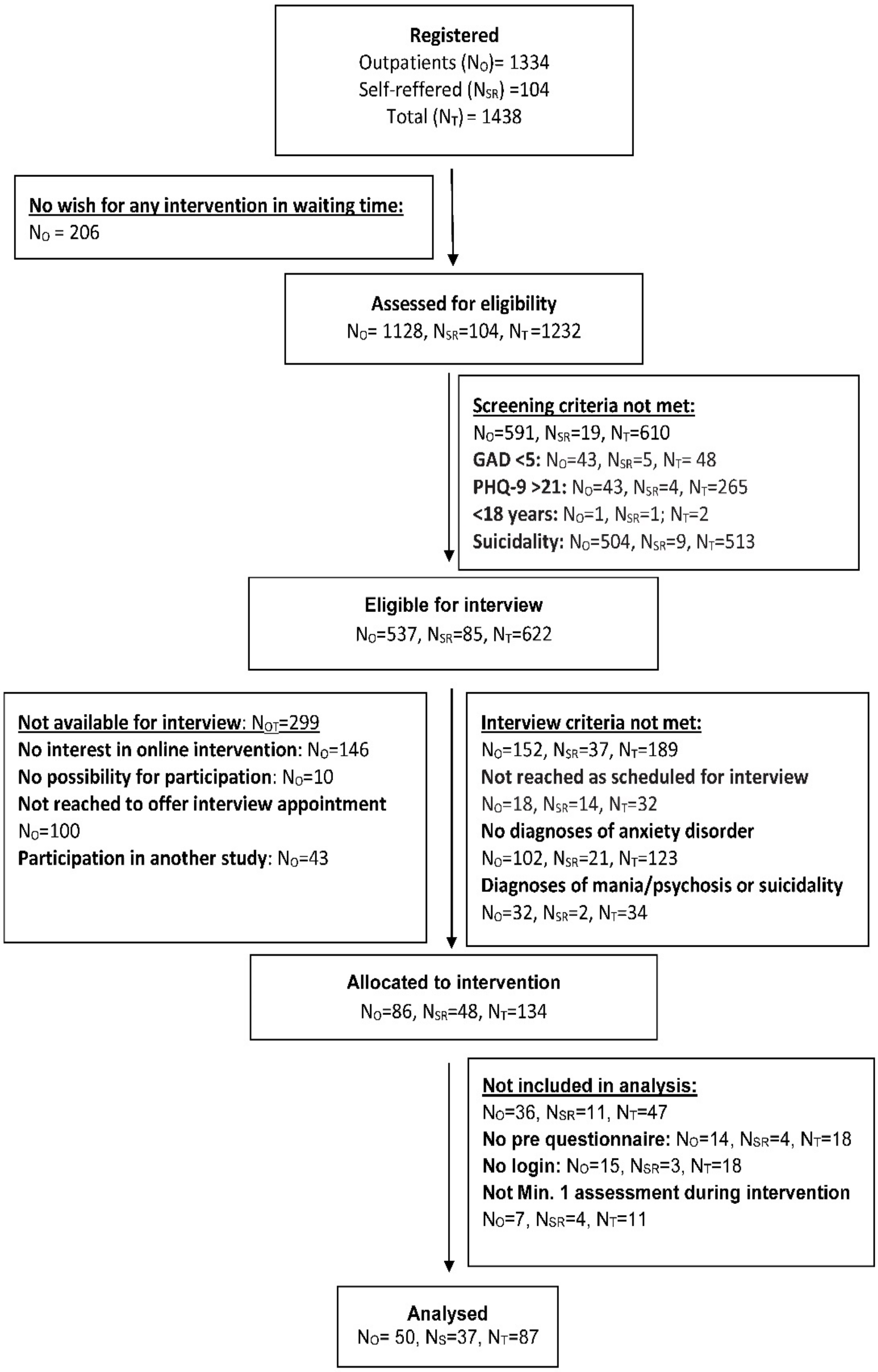

Figure 1. Flowchart of outpatient participants and self-referred participants

Table 1

Baseline variables of participants by group

Baseline variables	Overall (N=87)	Outpatient participants (N_{O}=50)	Self-referred participants (N_{SR}=37)	Test statistic (p)
M.I.N.I. diagnosis agoraphobia *N* (%)	30 (34.5)	20 (40)	10 (27)	$\chi^2_1 = 1.58$ p = .208
M.I.N.I. diagnosis social phobia *N* (%)	43 (49)	25 (50)	18 (48.6)	$\chi^2_1 = 0.02$ p = .901
M.I.N.I. diagnosis GAD[a] *N* (%)	35 (40.2)	24 (48)	11 (29.7)	$\chi^2_1 = 2.95$ p = .086
M.I.N.I. diagnosis panic disorder *N* (%)	30 (34)	16 (32)	14 (37.8)	$\chi^2_1 = 0.59$ p = .443
Sex Female *N* (%)	56 (64.4)	32 (64)	24 (64.9)	$\chi^2_1 = 0.01$ p = .934
University entrance diploma *N* (%)	55 (63.5)	26 (52)	29 (78.4)	$\chi^2_1 = 7.63$ p = .006**
Pre- mean score GAD-7[b] *M* (*SD*)	2.79 (0.59)	2.81 (0.55)	2.75 (0.65)	t(85)= 0.45 p = .652
Pre mean score Mini-SPIN M (*SD*)	3.04 (1.09)	3.28 (1.08)	2.72 (1.04)	t(85)= 2.43 p = .017*
Pre- mean score HSCL-11[c] *M* (*SD*)	2.27 (0.52)	2.32 (0.53)	2.19 (0.49)	t(85)= 1.14 p = .257
Pre- mean score PHQ-9[d] *M* (*SD*)	2.30 (0.59)	2.43 (0.58)	2.13 (0.58)	t(85)= 2.39 p = .019*
Treatment expectations *M* (*SD*)	2.94 (0.48)	2.91 (0.47)	2.97 (0.49)	t(85)= -0.63 p = .532
Age *M* (*SD*)	35.91 (12.70)	34.30 (11.5)	38.08 (14.1)	t(85) = 1.38 p = .171

[a]GAD: Generalized Anxiety Disorder.

[b]GAD-7: Generalized Anxiety Disorder Screener-7.

[c]HSCL Hopkins Symptom Checklist-11.

[d]PHQ-9: Patient Health Questionnaire-9.

Table 2

Adherence to intervention by participant group

Variable	Participant group	*M* (*SD*)/*N*	Test statistic	*p*
No. of sessions	Outpatient	4.48 (2.73)	t(85) = -2.56	.012*
	Self-referred	5.95 (5.95)		
	all	5.10 (2.72)		
Adherence to relaxation (No. of exercises)	Outpatient	4.38 (7.82)	t(85) = -1.96	.054
	Self-referred	8.24 (9.89)		
	all	6.02 (8.01)		
Adherence to exposure (yes)	Outpatient	11 (22.0%)	$\chi^2_1 = 4.48$	.034*
	Self-referred	16 (59.3%)		
	all	27 (31%)		
Adherence to cognitive restructuring (yes)	Outpatient	16 (32%)	$\chi^2_1 = 10.80$	.001*
	Self-referred	25 (50%)		
	all	41 (42%)		

Note. Outpatients: N_o= 50. Self-referred: N_{SR}= 37. All: N= 87.

Table 2

Outcome measures at pre and post and effect sizes by participant group

Measure	Participant group	Pre Treatment *M* (*SD*)	Post Treatment *M* (*SD*)	Effect-size pre to post Cohens *d*	Difference in effect size for Outpatient vs. Self-referred
GAD-7[a]	Outpatient	2.81 (0.55)	2.62 (0.79)	[illegible]	-0.34
	Self-referred	2.75 (0.65)	2,36 (0.82)	0.66	
	all	2.79 (0.59)	2.51 (0.81)	0.46	
Mini-SPIN[b]	Outpatient	3.28 (1.08)	2.85 (1.11)	1.03	0.16
	Self-referred	2.72 (1.04)	2.18 (0.72)	0.87	
	all	3.04 (1.09)	2.57 (1.02)	0.95	
HSCL-11[c]	Outpatient	2.32 (0.53)	2.25 (0.68)	0.13	-0.48*
	Self-referred	2.19 (0.49)	1.89 (0.57)	0.61	
	all	2.27 (0.52)	2.10 (0.66)	0.34	
PHQ-9[d]	Outpatient	2.43 (0.58)	2.31 (0.67)	0.20	-0.21
	Self-referred	2.13 (0.58)	1.89 (0.59)	0.41	
	all	2.30 (0.59)	2.13 (0.67)	0.29	

Note. Outpatients: N_O= 50. Self-referred: N_{SR}= 37. All: N= 87.

[a]GAD-7: Generalized Anxiety Disorder Screener-7.

[b]Mini-SPIN: Mini Social Phobia Inventory.

[c]HSCL-11: Hopkins Symptom Checklist-11.

[d]PHQ-9: Patient Health Questionnaire-9.

Table 4

Baseline impairment and adherence measures as predictors of change on the GAD-7 estimated with stepwise regression analysis (N=87)

Steps	Predictors	R^2	Beta (*SE*)	*t*	*p*
1		.133			.018
	GAD-7[a]pre		.26 (0.14)	1.85	.069
	Mini-SPIN[b]pre		-.18 (-0.07)	-1.46	.149
	HSCL-11[c]pre		-.42 (0.20)	-2.45	.016
	PHQ-9[d]pre		.18 (0.17)	1.05	.299
2		.267			<.001
	GAD-7pre	.	.25 (0.12)	2.06	.043
	Mini-SPINpre		-.16 (0.06)	-1.52	.134
	HSCL-11pre		-.48 (0.17)	-3.24	.002
	PHQ-9pre		.24 (0.15)	1.65	.104
	Number of sessions		.38 (0.02)	3.65	<.001
	Adherence to relaxation (No. of exercises)		-.30 (0.01)	-2.48	.015
	Adherence to cognitive restructuring (No. of exercises)		.02 (0.01)	0.19	.854
	Adherence to exposition (No. of exercises)		.35 (0.02)	3.15	.002
Total R^2		.400			
N		87			

[a]GAD-7: Generalized Anxiety Disorder Screener-7.

[b]Mini-SPIN: Mini Social Phobia Inventory.

[c]HSCL-11: Hopkins Symptom Checklist-11.

[d]PHQ-9: Patient Health Questionnaire-9.

8.8. References

Al-Asadi, A. M., Klein, B., & Meyer, D. (2014). Pretreatment attrition and formal withdrawal during treatment and their predictors: An exploratory study of the anxiety online data. Journal of Medical Internet Research, 16(6), e152. https://doi.org/10.2196/jmir.2989

Alfonsson, S., Olsson, E., & Hursti, T. (2016). Motivation and Treatment Credibility Predicts Dropout, Treatment Adherence, and Clinical Outcomes in an Internet-Based Cognitive Behavioral Relaxation Program: A Randomized Controlled Trial. Journal of Medical Internet Research, 18(3), e52. https://doi.org/10.2196/jmir.5352

Andrews, G., Basu, A., Cuijpers, P., Craske, M. G., McEvoy, P., English, C. L., & Newby, J. M. (2018). Computer therapy for the anxiety and depression disorders is effective, acceptable and practical health care: An updated meta-analysis. Journal of Anxiety Disorders, 55, 70–78. https://doi.org/10.1016/j.janxdis.2018.01.001

Beatty, L., & Binnion, C. (2016). A Systematic Review of Predictors of, and Reasons for, Adherence to Online Psychological Interventions. International Journal of Behavioral Medicine, 23(6), 776–794. https://doi.org/10.1007/s12529-016-9556-9

Berger, T., Boettcher, J., & Caspar, F. (2014). Internet-based guided self-help for several anxiety disorders: A randomized controlled trial comparing a tailored with a standardized disorder-specific approach. Psychotherapy (Chicago, Ill.), 51(2), 207–219. https://doi.org/10.1037/a0032527

Borza, L. (2017). Cognitive-behavioral therapy for generalized anxiety. Dialogues in Clinical Neuroscience, 19(2), 203–208.

Castro, A., López-Del-Hoyo, Y., Peake, C., Mayoral, F., Botella, C., García-Campayo, J., . . . Gili, M. (2018). Adherence predictors in an Internet-based Intervention program for depression. Cognitive Behaviour Therapy, 47(3), 246–261. https://doi.org/10.1080/16506073.2017.1366546

Connor, K. M., Davidson, J. R. T., Churchill, L. E., Sherwood, A., Weisler, R. H., & Foa, E. (2000). Psychometric properties of the Social Phobia Inventory (SPIN). British Journal of Psychiatry, 176(04), 379–386. https://doi.org/10.1192/bjp.176.4.379

Couper, M. P., Alexander, G. L., Zhang, N., Little, R. J. A., Maddy, N., Nowak, M. A., . . . Cole Johnson, C. (2010). Engagement and retention: Measuring breadth and depth of participant use of an online intervention. Journal of Medical Internet Research, 12(4), e52. https://doi.org/10.2196/jmir.1430

Domhardt, M., Geßlein, H., Rezori, R. E. von, & Baumeister, H. (2019). Internet- and mobile-based interventions for anxiety disorders: A meta-analytic review of intervention components. Depression and Anxiety, 36(3), 213–224. https://doi.org/10.1002/da.22860

El Alaoui, S., Ljótsson, B., Hedman, E., Kaldo, V., Andersson, E., Rück, C., . . . Lindefors, N. (2015). Predictors of Symptomatic Change and Adherence in Internet-Based Cognitive Behaviour Therapy for Social Anxiety Disorder in Routine Psychiatric Care. PloS One, 10(4), e0124258. https://doi.org/10.1371/journal.pone.0124258

Franke, G. H. (2000). Brief symptom inventory von LR Derogatis; BSI;(kurzform der SCL-90-R); deutsche Version: Beltz Test.

Geisheim, C., Hahlweg, K., Fiegenbaum, W., Frank, M., Schröder, B., & Witzleben, I. von. (2002). Das Brief Symptom Inventory (BSI) als Instrument zur Qualitätssicherung in der Psychotherapie. Diagnostica, 48(1), 28–36. https://doi.org/10.1026//0012-1924.48.1.28

Grosse Holtforth, M., & Grawe, K. (2003). Der Inkongruenzfragebogen (INK). Zeitschrift für Klinische Psychologie und Psychotherapie, 32(4), 315–323. https://doi.org/10.1026//1616-3443.32.4.315

Hebert, E. A., Vincent, N., Lewycky, S., & Walsh, K. (2010). Attrition and adherence in the online treatment of chronic insomnia. Behavioral Sleep Medicine, 8(3), 141–150. https://doi.org/10.1080/15402002.2010.487457

Hilvert-Bruce, Z., Rossouw, P. J., Wong, N., Sunderland, M., & Andrews, G. (2012). Adherence as a determinant of effectiveness of internet cognitive behavioural therapy for anxiety and depressive disorders. Behaviour Research and Therapy, 50(7-8), 463–468. https://doi.org/10.1016/j.brat.2012.04.001

Jacobson, N. S., & Truax, P. (1991). Clinical significance: A statistical approach to defining meaningful change in psychotherapy research. Journal of Consulting and Clinical Psychology, 59(1), 12–19. https://doi.org/10.1037/0022-006X.59.1.12

Kroenke, K., Spitzer, R. L., & Williams, J. B. (2001). The PHQ-9: Validity of a brief depression severity measure. Journal of General Internal Medicine, 16(9), 606–613.

Löwe, B., Decker, O., Müller, S., Brähler, E., Schellberg, D., Herzog, W., & Herzberg, P. Y. (2008). Validation and standardization of the Generalized Anxiety Disorder Screener (GAD-7) in the general population. Medical Care, 46(3), 266–274. https://doi.org/10.1097/MLR.0b013e318160d093

Lutz, W., Arndt, A., Rubel, J., Berger, T., Schröder, J., Späth, C., . . . Moritz, S. (2017). Defining and Predicting Patterns of Early Response in a Web-Based Intervention for Depression. Journal of Medical Internet Research, 19(6), e206. https://doi.org/10.2196/jmir.7367

Lutz, W., Schürch, E., Stulz, N., Böhnke, J. R., Schöttke, H., Rogner, J., & Wiedl, K. H. (2009). Entwicklung und psychometrische Kennwerte des Fragebogens zur Evaluation von Psychotherapieverläufen (FEP). Diagnostica, 55(2), 106–116. https://doi.org/10.1026/0012-1924.55.2.106

Lutz, W., Tholen, S., Schürch, E., & Berking, M. (2006). Reliabilität von Kurzformen gängiger psychometrischer Instrumente zur Evaluation des therapeutischen Fortschritts in Psychotherapie und Psychiatrie. Diagnostica, 52(1), 11–25. https://doi.org/10.1026/0012-1924.52.1.11

Newman, M. G., Lafreniere, L. S., & Jacobson, N. C. (2018). Relaxation-induced anxiety: Effects of peak and trajectories of change on treatment outcome for generalized anxiety disorder. Psychotherapy Research : Journal of the Society for Psychotherapy Research, 28(4), 616–629. https://doi.org/10.1080/10503307.2016.1253891

Rossi, A., Alberio, R., Porta, A., Sandri, M., Tansella, M., & Amaddeo, F. (2004). The Reliability of the Mini-International Neuropsychiatric Interview-Italian Version. Journal of Clinical Psychopharmacology, 24(5), 561–563. https://doi.org/10.1097/01.jcp.0000139758.03834.ad

Scholz, U., Gutiérrez Doña, B., Sud, S., & Schwarzer, R. (2002). Is General Self-Efficacy a Universal Construct?1. European Journal of Psychological Assessment, 18(3), 242–251. https://doi.org/10.1027//1015-5759.18.3.242

Schwarzer, R. (Ed.). (1999). Skalen zur Erfassung von Lehrer- und Schülermerkmalen: Dokumentation der psychometrischen Verfahren im Rahmen der wissenschaftlichen Begleitung des Modellversuchs Selbstwirksame Schulen. Berlin: R. Schwarzer.

Sheehan, D. V., Lecrubier, Y., Sheehan, K. H., Amorim, P., Janavs, J., Weiller, E., . . . Dunbar, G. C. (1998). The Mini-International Neuropsychiatric Interview (M.I.N.I.): The development and validation of a structured diagnostic psychiatric interview for DSM-IV and ICD-10. The Journal of Clinical Psychiatry, 59 Suppl 20, 22-33;quiz 34-57.

Sieverink, F., Kelders, S. M., & van Gemert-Pijnen, J. E. (2017). Clarifying the Concept of Adherence to eHealth Technology: Systematic Review on When Usage Becomes Adherence. Journal of Medical Internet Research, 19(12), e402. https://doi.org/10.2196/jmir.8578

Tibshirani, R. (2011). Regression shrinkage and selection via the lasso: a retrospective. Journal of the Royal Statistical Society: Series B (Statistical Methodology), 73(3), 273–282. https://doi.org/10.1111/j.1467-9868.2011.00771.x

Wagner, G., Penelo, E., Nobis, G., Mayrhofer, A., Wanner, C., Schau, J., . . . Karwautz, A. (2015). Predictors for good therapeutic outcome and drop-out in technology assisted guided self-help in the treatment of bulimia nervosa and bulimia like phenotype. European Eating Disorders Review : the Journal of the Eating Disorders Association, 23(2), 163–169. https://doi.org/10.1002/erv.2336

Wiltink, J., Kliem, S., Michal, M., Subic-Wrana, C., Reiner, I., Beutel, M. E., . . . Zwerenz, R. (2017). Mini - social phobia inventory (mini-SPIN): Psychometric properties and population based norms of the German version. BMC Psychiatry, 17(1), 377. https://doi.org/10.1186/s12888-017-1545-2

9. General discussion

The studies included in the current dissertation represent three important contributions to the field of research on change processes in internet interventions.

Study I concentrated on the identification of early change patterns in participants of an internet intervention targeting depression. Study II expanded on Study I's findings and investigated change as well as dropout by applying the Muthen-Roy model to identify change-dropout patterns.

Study III investigated how adherence and outcome differed in outpatient and self-selected participants of an internet intervention targeting anxiety. In all participants, the impact of adherence to treatment components on outcome as well as possible predictors of adherence were investigated.

Studies I and II used statistical approaches, which use latent variables to identify different change patterns. In Study III, LASSO regression was used to identify important predictor variables. In Studies I and II, data stemmed from participants of an internet intervention targeting depressive symptoms that was implemented by a cooperation of different study sites. In Study III, data stemmed from a project that was realized at the outpatient clinic at the University of Trier, where the internet intervention targeting anxiety was offered to outpatients as well as interested persons suffering from anxiety problems.

Study I transfers research findings on change processes in face-to-face therapies (e.g. Lutz et al., 2014) to internet interventions. It is one of the few studies to investigate early change patterns in participants undergoing an internet intervention (e.g. Schibbye et al., 2014) and the only one to apply GMM in this context. Results suggested that participants of internet interventions show varying early change patterns, which are associated with final treatment outcome and, to some degree, adherence.

Study II expands on Study I's findings by applying the Muthen-Roy model, which models change as well as dropout to identify change-dropout patterns (Muthén et al., 2011).

This study remains one of few studies to investigate dropout mechanisms in a psychological treatment using a NMAR model. Results suggested that dropping out of in-treatment questionnaires and change were interrelated, with participants showing four different change-dropout patterns: One group of participants showed improvement and low dropout risk, another group showed deterioration and high dropout risk. Unexpectedly some participants also showed improvement, but had a high dropout risk, while others experienced no change and had a low drop out risk. Change-dropout patterns were related to adherence and long-term treatment outcome.

Study III broadens our knowledge of adherence to internet interventions in different participant groups, which had often been addressed rather superficially in the past (e.g. Berger, Boettcher, & Caspar, 2014; Castro et al., 2018; Klein et al., 2016). Adherence to treatment components (relaxation, cognitive restructuring, and exposure) as well as outcome were compared between outpatient and self-referred participants of an internet intervention targeting anxiety. In addition, it was investigated whether outpatient participants who had access to the internet intervention during the waiting period for face-to-face treatment improved more during waiting period than outpatients without access to the intervention. Results suggested that participant referral impacts adherence to treatment components and, in part, also treatment outcome. On average, more adherent participants showed better outcomes, with only adherent outpatient participants showing additional benefit compared to outpatients without access to the intervention. Furthermore, specific patient characteristics such as self-efficacy predicted adherence to treatment components.

The following sections draw general conclusions from the findings based on the three studies. Future research directions are summarized and limitations are discussed.

9.1. General conclusions and future research

Some general conclusions and future research directions can be derived from considering the results of all three studies simultaneously. The main finding of all studies seems to be that even in rather standardized internet interventions, the change process matters. As outlined above, studying the change process in internet interventions has attracted less attention than investigating treatment efficacy using RCTs (Ebert et al., 2018). Despite many results pointing to the efficacy of internet interventions, Andrews & Williams (2014) estimate that nearly half of the participants in internet interventions do not improve. Therefore, it seems rather important to put more focus on the process of change in internet interventions to gain more insight regarding the following questions: (1) Which participants are most likely to benefit from internet interventions? (2) Which participants are at risk of treatment failure and how can treatments be adjusted to increase adherence and treatment outcome? With a better understanding of important patient characteristics and change processes, it may be possible to develop informed clinical guidelines for the application of internet interventions, which are currently still lacking on an international level (Ebert et al., 2018).

In Study I, almost half of participants (45%) showed low initial impairment and symptom improvement already during the waiting period for the internet intervention. In line with the finding that these participants' impairment was lower at the beginning of treatment (indicating less need of treatment), they were less adherent than participants who improved only after treatment began. In contrast, participants with high impairment showed either early response (39%) or early deterioration (16%), with more participants showing early response. Participants whose physical health was more impaired showed a greater risk of deterioration. In general, positive attitudes towards internet interventions were associated with lower impairment at post, but not with greater adherence.

While an early response rate of nearly 50% seems quite high, these results match Delgadillo et al.'s (2014) findings. This group reported that 50% of participants who had completed at least four sessions in low intensity interventions fulfilled criteria for reliable improvement. However, compared to the early response rate reported by Lutz et al. (2014) regarding early change in psychotherapy, the early response rate identified in Study I was much higher (39% vs. 20%). However differences between patient samples regarding impairment and diagnosis need to be considered. Furthermore, one explanation for the different early response rates found in our study and the study of Lutz et al. (2014) may be that while most participants take part in internet interventions voluntarily (Crisp & Griffiths, 2014), there are situations in which patients feel obliged or pressured into pursuing face-to-face treatment (e.g. by relatives or by especially difficult life circumstances such as work disability). In addition, while the early response rate was higher in Study I than in the study by Lutz et al. (2014), the same is true of the early deterioration rate (16% vs. 5%). This could indicate that the probability of early deterioration is lower in face-to-face treatment. Factors unique to face-to-face therapy such as the initiation of a positive therapeutic bond, a good fit between therapist and patient, positive effects of the working alliance as well as therapist-specific factors such as experience may provide one explanation. Some of these factors may be especially important in association with specific patient characteristics such as chronicity of illness, negative treatment expectations, ambivalent motivation, and experience of stigma or degradation by peers and relatives (see also Delgadillo et al., 2017). This would imply that it is important to identify relevant participant characteristics to identify specific treatment needs. Study I's findings indicate that one such important participant characteristic is physical health.

To conclude, identifying and considering specific participants' needs regarding treatment could improve adherence and lower risk of treatment failure (Ebert et al., 2018). To achieve this, it is important to develop clearer clinical guidelines regarding treatment allocation and adaptation. Interesting treatment options that are currently discussed include the use of stepped-care approaches and personalized treatments (Ebert et al., 2018).

Investigating these treatment options for different participants will take time and effort. Currently, clinicians should be informed of the impact of early change on later treatment outcome in order to be able to identify cases at risk of deterioration. In contrast to participants at risk of no change or deterioration, especially participants who already improve during the waiting period may need less intensive treatment than other participants who enter treatment with a high level of impairment.

Study II indicated two potential mechanisms of dropout: On the one hand, participants who were only mildly impaired and showed very rapid improvement had a higher risk of dropout. On the other hand, participants who deteriorated also had a higher risk of dropout. In contrast, participants who improved after the beginning of treatment had a low dropout risk and a high probability of adhering to treatment. There was also a group of participants, who showed no change and still had a low dropout risk.

In line with Study I's findings, both highly impaired participant groups (the one deteriorating and the one showing no change) were more likely to show lower levels of physical health than improving participants. Both participants groups who did not improve showed poorer treatment outcome at 12-month follow-up than participants, who improved and did not drop out. This is in line with reports that, overall, patients who drop out of treatment have poorer treatment effects (e.g. Delgadillo et al., 2014; Lutz et al., 2014). Thus, Study II's findings also indicate that especially for deteriorating participants at risk of dropout, but also for non-improving participants, treatment adaption is important.

Regarding the risk of deterioration and dropout, it remains difficult to clearly identify the reasons for deterioration. To date, there are no clear findings regarding causes of treatment failure in internet interventions (Andersson et al., 2019). Fernandez, Salem, Swift, & Ramtahal (2015) reported that dropout was significantly associated with depression, which matches our findings that participants showing deterioration and a high risk of dropout, more often fulfilled diagnostic criteria of depression than other participants. Likely, more severely depressed participants feel overwhelmed with the effort necessary to complete measurements or are frustrated with the lack of progress. Although these participants probably would have benefited more from another treatment, it is also possible that they are less likely to improve compared to other patients independent of the treatment offered (DeRubeis et al., 2014). It is unclear, what the best procedure regarding these participants may be.

By experiencing treatment failure, participants of internet interventions may experience negative developments that aggravate the situation such as low treatment expectations, negative attitudes towards psychological treatment, and negative self-efficacy. Primarily, an attempt should likely be made to prevent such negative developments in non-improving participants. For example, Schibbye et al. (2014) suggest that clinical supervision and engagement efforts should particularly focus on those participants, who have not shown pre-treatment symptom improvement or early reliable change. In addition, they underline the importance of routine monitoring of symptom change in internet interventions. As the dose-response effect appears to decline after session 6, they suggest using this timepoint to decide whether a patient should be stepped-up.

Possible adaptations of treatment in the context of a personalized treatment approach could include the treatment format, the intensity and modality of contact, as well as the level of therapist directiveness and patient autonomy (Forsell et al., 2019). While it could be important to examine participants' attitudes towards internet interventions early and, if possible, foster positive attitudes, an approach to reduce the risk of treatment failure must also take patients' treatment preferences into account.

It has been reported that some 20% of participants in internet interventions dropout, because they have received sufficient benefit (Andrews & Williams, 2014). This is also in line with the good enough level model proposed by Barkham et al. (2006). Among others, this phenomenon may explain the rather high dropout rates reported in studies on internet interventions (Fernandez et al., 2015). Other important variables related to dropout include self-efficacy, treatment credibility, intrinsic motivation for treatment, and focus on immediate consequences (Alfonsson, Johansson, Uddling, & Hursti, 2017). Thus, participants that drop out of internet interventions may perceive them as either not relevant to their problem, unlikely to help or too difficult (see Matsuzawa et al., 2019). In addition, it is important to consider that people who are interested in internet interventions may represent a specific group with characteristics such as fear of stigma, aversion of disclosure or preferring to solve problems on their own. In addition, because of the high availability of internet interventions, it is possible that people consider treatment that otherwise would not consider it, e.g. because of rather low impairment, time or cost issues. Concerning this group the questions remain who of them really needs treatment, what do they want from treatment and how can they be supported best.

In addition to these interesting points, one might deduce further research questions when combining findings from both Studies I and II.

Are those participants, who show early response after screening the same participants who are mildly impaired and have a high dropout probability? And is the improvement shown by these participants stable in the long term? This could have important implications regarding clinicians' reactions to dropout, e.g. participants who show improvement before treatment, could be offered to "come back", should symptoms reoccur.

Further questions include: What about participants who experience improvement only after the beginning of treatment? Are they the same as the participants who show improvement and a low dropout probability? If yes, why don't they leave treatment after having received a good enough dosage of treatment as the other participants do?

A possible explanation could be that in contrast to participants, who improve during the waiting period, these participants attribute their improvement to treatment and possibly to their own efforts.

In Study III, the results indicated that self-referred participants were more adherent to treatment components than outpatient participants. While adherence to exposure was positively associated with outcome, results regarding adherence to relaxation were mixed: Participants who completed at least two relaxation exercises showed better treatment outcome, but the total number of relaxation exercises was associated with less positive treatment outcome. This could indicate that there is an optimal dosage of relaxation in the treatment of anxiety disorders. Furthermore, only outpatient participants who had adhered to relaxation had an additional benefit compared to outpatients without access to the internet intervention. Outpatient participants with higher self-efficacy at the beginning of the waiting period were more likely to adhere to relaxation.

The finding that adherence differed between self-referred and outpatient participants could have important implications for treatment selection and provision of guidance, but first, further studies with larger sample sizes are necessary to systematically investigate differences between participant groups regarding adherence and treatment outcome. To date, some studies have examined referral context, however they have mostly focused on differences in pre-treatment patient characteristics (Crisp & Griffiths, 2014; Lindner, Nyström, Hassmén, Andersson, & Carlbring, 2015), even when these characteristics have not consistently been found to be associated with treatment outcome.

The difference between self-referred and outpatient participants regarding adherence could provide an explanation for the lower effects that are often reported in studies implementing internet interventions in primary care (see Romijn et al., 2019). In line with this finding, Mohr et al. (2010) reported that patients in a primary care setting were less interested in internet interventions than in face-to-face treatment. In addition, Alfonsson et al. (2016) reported that participants who felt pressured to engage in treatment showed higher stress levels at post-measurement, underlining the importance of patient motivation and treatment preference. More research is necessary to establish how participant groups differ in order to gain a better understanding of how these differences can be acknowledged when considering internet interventions as treatment options for specific patients. As higher self-efficacy predicted higher adherence to relaxation in outpatients, self-efficacy may be an important patient characteristic impacting adherence and possibly outcome in internet interventions.

In addition to these important implications, in future research, it should be taken into consideration that the use of general measures of adherence may be insufficient to understand the change process in internet interventions. The vague definition of treatment dosage as the number of sessions has already been criticized in psychotherapy research (e.g. Cuijpers, Reijnders, & Huibers, 2019). Adherence to treatment components could be especially crucial in internet interventions, where there is no therapist to ensure that specific exercises are completed, at least during sessions. Furthermore, some participants may experience specific treatment components to be more difficult to complete than others or they may have preferences for specific components (see Alfonsson et al., 2016). Further research is necessary to investigate how much participants adhere to specific treatment components, why they do so and how much they benefit subsequently. This could explain some of the variance regarding the outcome of internet interventions across different samples and also have important implications for the provision of feedback to participants, e.g. the timing of feedback and content.

When combining the findings of all three studies, some interesting suggestions for research and clinical practice as well as interesting research questions can be deduced. One interesting question focuses on the mechanism that leads to the occurrence of early response in some patients. As this response occurs early in treatment, before important treatment components have been introduced, Haas, Hill, Lambert, & Morrell (2002) suggested that patient characteristics and common factors such as initiation of hope may play a role in the occurrence of early response. Interestingly, in our study, some participants showing early response achieved improvement even before treatment onset, with other participants showing early response only after treatment began. It is possible that those participants, who showed early response after treatment began, experienced hope and an early rise in self-efficacy, because they gained a better understanding of their condition and potentially successfully applied some initial therapeutic techniques.

Despite the potential to shed some light on the processes that contribute to early change, to my knowledge, self-efficacy remains to be investigated in this context. In addition, many studies do not specify the timing of early response and thus do not allow regression to the mean to be distinguished from early response that may be triggered by factors such as hope or self-efficacy.

Interestingly, Ebert et al. (2018) have noted that especially depressed participants benefit from guided internet interventions. Possibly, a combination of specific characteristics could make participants more prone to react positively to therapeutic contact, including a lack of positive interaction in daily life, low self-efficacy and frequent trouble structuring tasks. Supporting the idea that self-efficacy and adherence are associated, Zarski et al. (2018) reported that planning predicted adherence and better planning was associated with higher levels of self-efficacy maintenance. This could indicate how to provide guidance in internet interventions for participants at risk of deterioration, e.g. by helping participants to plan the next treatment steps and by fostering self-efficacy. Still, it remains unclear, which dose of guidance may be optimal for whom at which stage of treatment (Ebert et al., 2018).

In general, self-efficacy has seldom been investigated in internet interventions (Beatty & Binnion, 2016). This is an important shortcoming, because especially in the treatment context of internet interventions, a high level of self-efficacy may be crucial to adherence and outcome. There are findings that suggest that internet interventions are equally effective as face-to-face interventions (Andrews et al., 2018), however this may only be true for a selected group of participants. Further studies should investigate how both treatments work for different participants and should also consider important patient characteristics, such as self-efficacy.

It would be interesting to see how much and how fast self-efficacy changes in participants of internet interventions compared to patients in face-to-face psychological treatment. While face-to-face therapies may work better than internet interventions for patients with low self-efficacy, internet interventions may promote change in self-efficacy more quickly than face-to-face psychological treatments, as they promote self-empowerment (see Ebert et al., 2018).

Similarly, it would be interesting to investigate, whether participants who improve rapidly during internet interventions and dropout also have high self-efficacy. This could potentially explain why they drop out, e.g. when they have confidence that they can cope on their own. In contrast, some of the non-improving participants may not dropout, e.g. because they feel like they are unable to cope on their own or even find a treatment that works better for them.

In summary, more research on treatment change, dropout and adherence in participants of internet interventions is necessary. As internet interventions use a standardized approach, the risk may be higher that specific participant needs remain unmet, leading to poor outcome. In addition, negative effects beyond symptom deterioration may also occur in internet interventions and should also be investigated (Andersson et al., 2019).

To conclude, the findings of this dissertation suggest that internet interventions can be seen less as a substitute for face-to-face interventions, but rather as a useful addition to treatment options (see Ebert et al., 2018). However, this only applies, when actions are taken to optimize treatment quality for all patients. Specifically, several points might be important to improve treatment availability as well as treatment outcomes for participants of internet interventions:

1. Applying quality control of internet interventions in a form that also considers ethical issues such as data protection.
2. Increasing awareness of potential benefits and limitations of internet interventions in important stakeholders such as providers of interventions, guiding clinicians as well as people in need of psychological treatment.
3. Applying specific treatment options to improve outcome in internet interventions (e.g. empirically informed stepped-care or personalized treatment).
4. Applying outcome monitoring and providing feedback to clinicians and guidance to participants of internet interventions.

Regarding point four, in face-to-face treatments, outcome monitoring and feedback tools exist (see Lambert, 2007) that allow the identification of risk cases and provide clinical suggestions. They have already been successfully implemented in several settings and their effects are currently being investigated, e.g. in an outpatient clinic in Germany (Lutz et al., 2019). Therefore, research on internet interventions should continue to investigate the change process in internet interventions, so that such important developments can be transferred to internet interventions. This would increase their potential as useful treatment options for psychological disorders, which is important, considering the low availability of empirically-based psychological treatments for patients in need.

Regardless of the strengths of the studies discussed above, they are not free of limitations. These general limitations are discussed in the subsequent section.

9.2. General limitations

As described above, Studies I and II were based on data from the same participants. It would have been interesting to compare the results of the GMM and the Muthen-Roy model across different samples. Furthermore, as participants of Deprexis were recruited in varying settings including the media, it can be assumed that most participants were self-referred. Therefore, findings most likely cannot be easily generalized to primary care settings. In both studies, only symptom change was investigated, ignoring other potentially interesting change processes.

Study III attempted to overcome the limitation regarding the specific referral context present in Studies I and II by comparing self-referred participants to outpatient participants. However, no randomization occurred, there was no control group available for self-referred participants, and the sample size was quite small. Thus, it is necessary to replicate the findings in studies investigating larger randomized samples. In addition, participation in the study was voluntary for all outpatients, so even when they were referred to the intervention by a clinician from the outpatient clinic, self-selection also occurred in outpatient participants.

In all studies, participants were allowed to make use of treatment as usual, including medication, so it remains unclear, whether the observed symptom change was due to the internet intervention.

Study III tried to control for patient characteristics and effects of medication by comparing outpatient participants to a matched outpatient waitlist sample, however there were no multiple measurements available to monitor change processes in the matched waitlist sample. Furthermore, it remains unclear, whether outpatients in the matched waitlist sample made use of any other treatments during the waiting period for face-to-face treatment.

Another limitation to Study III is that many specific measures were only available for outpatient participants, not for self-referred participants. Also, in contrast to Studies I and II, general attitudes towards internet interventions were not measured in Study III. This is a shortcoming, as it could have provided additional information regarding differences in attitudes between self-referred participants and outpatients.

Another difference between the studies is that in Studies I and II, only general measures of adherence were used, limiting what can be said about possible mechanisms of change. In Study III, more specific adherence measures were used, however the findings cannot be generalized to other internet interventions that target different disorders.

In all studies, it would have been interesting to have multiple measures of self-efficacy as well as adherence to better understand the underlying process of change.

Based on Study I and II's findings, it was suggested that some participants at risk of poor treatment outcome should be stepped-up. However, this is a preliminary conclusion, as no control group was available in which the stepped-care approach was applied.

In addition, several limitations must be mentioned regarding the methodological approaches that were applied. In Study I, GMM was applied and results should be interpreted with caution. First, the identified patterns are a mere simplification of a much more complex reality. Second, while several criteria for model selection exist, researchers should still bear in mind that different model specifications may lead to very different results. Similarly, the Muthen-Roy model that was applied in Study II is a very complex model and thus may be especially vulnerable to potential model specification errors. In both cases, estimating the applied models in a control sample with multiple measurements may have provided more clarity regarding the identified patterns (early change patterns and change-dropout patterns). In general, when more complex models such as GMM and Muthen-Roy are applied, large data sets with many repeated measurements are often beneficial.

Regarding feature selection conducted using LASSO in Study III, it should be noted that while this approach protects its estimates from biases such as overfitting, it is not considered stable, as results are based on multiple subsets of features and on different subsamples. This is a shortcoming, if the results cannot be tested in another dataset with the same features.

9.3. Concluding remarks

The aim of this dissertation was to investigate the change process in internet interventions by investigating patterns of change and dropout as well as adherence to treatment components and their impact on treatment outcome. Despite several limitations and the caution necessary, when interpreting the results, it can be concluded that focusing on the process of change in internet interventions might help to further improve treatment. This could be relevant in the context of different participant groups with individual needs as well as a varying risk of poor adherence, poor treatment outcome or dropout. Only with a better understanding of how improvement or deterioration occur and why, can decisions regarding treatments options be made. These options include stepped-care approaches and treatment personalization (e.g. regarding level of guidance, treatment form, and treatment content). Empirically-based clinical guidelines could help to reduce the number of participants who do not respond to internet interventions. In this context, clinicians should keep an eye on important patient variables that may lower the probability of success, even in future treatment (e.g. negative treatment expectations and low self-efficacy).

All three studies are more or less starting points in their specific fields. More research is needed to investigate whether findings can be replicated in larger samples and in internet interventions targeting different disorders. Investigating differences in treatment change can be seen as an opportunity to identify individual mechanisms of change and further improve treatments according to patients' specific needs.

9.4. References

References

Alfonsson, S., Johansson, K., Uddling, J., & Hursti, T. (2017). Differences in motivation and adherence to a prescribed assignment after face-to-face and online psychoeducation: An experimental study. *BMC Psychology*, *5*(1), 3. https://doi.org/10.1186/s40359-017-0172-5

Alfonsson, S., Olsson, E., & Hursti, T. (2016). Motivation and Treatment Credibility Predicts Dropout, Treatment Adherence, and Clinical Outcomes in an Internet-Based Cognitive Behavioral Relaxation Program: A Randomized Controlled Trial. *Journal of Medical Internet Research*, *18*(3), e52. https://doi.org/10.2196/jmir.5352

Andersson, G., Rozental, A., Rück, C., & Carlbring, P. (2015). Guided Internet-delivered CBT: Can it really be as good as seeing a therapist? *The Behavior Therapist*, *38*, 123–126.

Andersson, G., Titov, N., Dear, B. F., Rozental, A., & Carlbring, P. (2019). Internet-delivered psychological treatments: From innovation to implementation. *World Psychiatry : Official Journal of the World Psychiatric Association (WPA)*, *18*(1), 20–28. https://doi.org/10.1002/wps.20610

Andrews, G., Basu, A., Cuijpers, P., Craske, M. G., McEvoy, P., English, C. L., & Newby, J. M. (2018). Computer therapy for the anxiety and depression disorders is effective, acceptable and practical health care: An updated meta-analysis. *Journal of Anxiety Disorders*, *55*, 70–78. https://doi.org/10.1016/j.janxdis.2018.01.001

Andrews, G., & Williams, A. D. (2014). Internet psychotherapy and the future of personalized treatment. *Depression and Anxiety*, *31*(11), 912–915. https://doi.org/10.1002/da.22302

Barkham, M., Connell, J., Stiles, W. B., Miles, J. N. V., Margison, F., Evans, C., & Mellor-Clark, J. (2006). Dose-effect relations and responsive regulation of treatment duration: The good enough level. *Journal of Consulting and Clinical Psychology*, *74*(1), 160–167. https://doi.org/10.1037/0022-006X.74.1.160

Beatty, L., & Binnion, C. (2016). A Systematic Review of Predictors of, and Reasons for, Adherence to Online Psychological Interventions. *International Journal of Behavioral Medicine*, *23*(6), 776–794. https://doi.org/10.1007/s12529-016-9556-9

Berger, T., Boettcher, J., & Caspar, F. (2014). Internet-based guided self-help for several anxiety disorders: A randomized controlled trial comparing a tailored with a standardized disorder-specific approach. *Psychotherapy (Chicago, Ill.)*, *51*(2), 207–219. https://doi.org/10.1037/a0032527

Berlin, K. S., Parra, G. R., & Williams, N. A. (2014). An introduction to latent variable mixture modeling (part 2): Longitudinal latent class growth analysis and growth mixture models. *Journal of Pediatric Psychology*, *39*(2), 188–203. https://doi.org/10.1093/jpepsy/jst085

Carlbring, P., Andersson, G., Cuijpers, P., Riper, H., & Hedman-Lagerlöf, E. (2018). Internet-based vs. Face-to-face cognitive behavior therapy for psychiatric and somatic disorders: An updated systematic review and meta-analysis. *Cognitive Behaviour Therapy*, *47*(1), 1–18. https://doi.org/10.1080/16506073.2017.1401115

Castro, A., López-Del-Hoyo, Y., Peake, C., Mayoral, F., Botella, C., García-Campayo, J., . . . Gili, M. (2018). Adherence predictors in an Internet-based Intervention program for depression. *Cognitive Behaviour Therapy*, *47*(3), 246–261. https://doi.org/10.1080/16506073.2017.1366546

Clark, D. M. (2018). Realizing the Mass Public Benefit of Evidence-Based Psychological Therapies: The IAPT Program. *Annual Review of Clinical Psychology*, *14*(1), 159–183. https://doi.org/10.1146/annurev-clinpsy-050817-084833

Coertjens, L., Donche, V., Maeyer, S. de, Vanthournout, G., & van Petegem, P. (2017). To what degree does the missing-data technique influence the estimated growth in learning strategies over time? A tutorial example of sensitivity analysis for longitudinal data. *PloS One*, *12*(9), e0182615. https://doi.org/10.1371/JOURNAL.PONE.0182615

Crisp, D. A., & Griffiths, K. M. (2014). Participating in online mental health interventions: Who is most likely to sign up and why? *Depression Research and Treatment*, *2014*, 790457. https://doi.org/10.1155/2014/790457

Cuijpers, P., Reijnders, M., & Huibers, M. J. H. (2019). The Role of Common Factors in Psychotherapy Outcomes. *Annual Review of Clinical Psychology*, *15*, 207–231. https://doi.org/10.1146/annurev-clinpsy-050718-095424

Delgadillo, J., Gellatly, J., & Stephenson-Bellwood, S. (2015). Decision making in stepped care: How do therapists decide whether to prolong treatment or not? *Behavioural and Cognitive Psychotherapy*, *43*(3), 328–341. https://doi.org/10.1017/S135246581300091X

Delgadillo, J., Huey, D., Bennett, H., & McMillan, D. (2017). Case complexity as a guide for psychological treatment selection. *Journal of Consulting and Clinical Psychology*, *85*(9), 835–853. https://doi.org/10.1037/ccp0000231

Delgadillo, J., McMillan, D., Lucock, M., Leach, C., Ali, S., & Gilbody, S. (2014). Early changes, attrition, and dose-response in low intensity psychological interventions. *The British Journal of Clinical Psychology*, *53*(1), 114–130. https://doi.org/10.1111/bjc.12031

Demirtas, H., & Schafer, J. L. (2003). On the performance of random-coefficient pattern-mixture models for non-ignorable drop-out. *Statistics in Medicine*, *22*(16), 2553–2575. https://doi.org/10.1002/sim.1475

DeRubeis, R. J., Cohen, Z. D., Forand, N. R., Fournier, J. C., Gelfand, L. A., & Lorenzo-Luaces, L. (2014). The Personalized Advantage Index: Translating research on prediction into individualized treatment recommendations. A demonstration. *PloS One*, *9*(1), e83875. https://doi.org/10.1371/journal.pone.0083875

Ebert, D. D., & Baumeister, H. (2017). Internet-Based Self-help Interventions for Depression in Routine Care. *JAMA Psychiatry*, *74*(8), 852–853. https://doi.org/10.1001/jamapsychiatry.2017.1394

Ebert, D. D., van Daele, T., Nordgreen, T., Karekla, M., Compare, A., Zarbo, C., . . . Baumeister, H. (2018). Internet- and Mobile-Based Psychological Interventions: Applications, Efficacy, and Potential for Improving Mental Health. *European Psychologist*, *23*(2), 167–187. https://doi.org/10.1027/1016-9040/a000318

El Alaoui, S., Ljótsson, B., Hedman, E., Kaldo, V., Andersson, E., Rück, C., . . . Lindefors, N. (2015). Predictors of Symptomatic Change and Adherence in Internet-Based Cognitive Behaviour Therapy for Social Anxiety Disorder in Routine Psychiatric Care. *PloS One*, *10*(4), e0124258. https://doi.org/10.1371/journal.pone.0124258

Enders, C. K. (2011). Missing not at random models for latent growth curve analyses. *Psychological Methods*, *16*(1), 1–16. https://doi.org/10.1037/a0022640

Fernandez, E., Salem, D., Swift, J. K., & Ramtahal, N. (2015). Meta-analysis of dropout from cognitive behavioral therapy: Magnitude, timing, and moderators. *Journal of Consulting and Clinical Psychology*, *83*(6), 1108–1122. https://doi.org/10.1037/ccp0000044

Fisher, A. J. (2015). Toward a dynamic model of psychological assessment: Implications for personalized care. *Journal of Consulting and Clinical Psychology*, *83*(4), 825–836. https://doi.org/10.1037/ccp0000026

Forsell, E., Jernelöv, S., Blom, K., Kraepelien, M., Svanborg, C., Andersson, G., . . . Kaldo, V. (2019). Proof of Concept for an Adaptive Treatment Strategy to Prevent Failures in Internet-Delivered CBT: A Single-Blind Randomized Clinical Trial With Insomnia Patients. *The American Journal of Psychiatry*, *176*(4), 315–323. https://doi.org/10.1176/appi.ajp.2018.18060699

Grünzig, S.-D., Bengel, J., Göhner, W., & Krämer, L. V. (2019). Niedrigintensive Interventionen zur Reduktion depressiver Symptome vor Beginn einer ambulanten Psychotherapie [Low-Intensity Interventions to Reduce Depressive Symptoms before Outpatient Psychotherapy - A Systematic Literature Review]. *Psychotherapie,* *Psychosomatik, medizinische Psychologie*, *69*(6), 212–223. https://doi.org/10.1055/a-0630-2397

Gunlicks-Stoessel, M., & Mufson, L. (2011). Early patterns of symptom change signal remission with interpersonal psychotherapy for depressed adolescents. *Depression and Anxiety*, *28*(7), 525–531. https://doi.org/10.1002/da.20849

Haas, E., Hill, R. D., Lambert, M. J., & Morrell, B. (2002). Do early responders to psychotherapy maintain treatment gains? *Journal of Clinical Psychology*, *58*(9), 1157–1172. https://doi.org/10.1002/jclp.10044

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

Hilvert-Bruce, Z., Rossouw, P. J., Wong, N., Sunderland, M., & Andrews, G. (2012). Adherence as a determinant of effectiveness of internet cognitive behavioural therapy for anxiety and depressive disorders. *Behaviour Research and Therapy*, *50*(7-8), 463–468. https://doi.org/10.1016/j.brat.2012.04.001

Howard, K. I., Kopta, S. M., Krause, M. S., & Orlinsky, D. E. (1986). The dose–effect relationship in psychotherapy. *American Psychologist*, *41*(2), 159–164. https://doi.org/10.1037/0003-066X.41.2.159

Hunter, A. M., Muthén, B. O., Cook, I. A., & Leuchter, A. F. (2010). Antidepressant response trajectories and quantitative electroencephalography (QEEG) biomarkers in major depressive disorder. *Journal of Psychiatric Research*, *44*(2), 90–98. https://doi.org/10.1016/j.jpsychires.2009.06.006

Johnson, W. (2015). Analytical strategies in human growth research. *American Journal of Human Biology : the Official Journal of the Human Biology Council*, *27*(1), 69–83. https://doi.org/10.1002/ajhb.22589

Jung, T., & Wickrama, K. A. S. (2008). An Introduction to Latent Class Growth Analysis and Growth Mixture Modeling. *Social and Personality Psychology Compass*, *2*(1), 302–317. https://doi.org/10.1111/j.1751-9004.2007.00054.x

Klein, J. P., Berger, T., Schröder, J., Späth, C., Meyer, B., Caspar, F., . . . Hohagen, F. (2016). Effects of a Psychological Internet Intervention in the Treatment of Mild to Moderate Depressive Symptoms: Results of the EVIDENT Study, a Randomized Controlled Trial. *Psychotherapy and Psychosomatics*, *85*(4), 218–228. https://doi.org/10.1159/000445355

Lambert, M. J. (2007). Presidential address: What we have learned from a decade of research aimed at improving psychotherapy outcome in routine care. *Psychotherapy Research*, *17*(1), 1–14. https://doi.org/10.1080/10503300601032506

Laurenceau, J.-P., Hayes, A. M., & Feldman, G. C. (2007). Some methodological and statistical issues in the study of change processes in psychotherapy. *Clinical Psychology Review*, *27*(6), 682–695. https://doi.org/10.1016/j.cpr.2007.01.007

Lindner, P., Nyström, M. B.T., Hassmén, P., Andersson, G., & Carlbring, P. (2015). Who seeks ICBT for depression and how do they get there? Effects of recruitment source on patient demographics and clinical characteristics. *Internet Interventions*, *2*(2), 221–225. https://doi.org/10.1016/j.invent.2015.04.002

Lutz, W., Clausen, S. A., Bennemann, B., Zimmermann, D., Prinz, J., Rubel, J., & Deisenhofer, A.-K. (2019). Chancen von E-Mental-Health und eProzessdiagnostik in der ambulanten Psychotherapie: Der Trierer Therapie Navigator. *Verhaltenstherapie*, 1–10. https://doi.org/10.1159/000501026

Lutz, W., Hofmann, S. G., Rubel, J., Boswell, J. F., Shear, M. K., Gorman, J. M., . . . Barlow, D. H. (2014). Patterns of early change and their relationship to outcome and early treatment termination in patients with panic disorder. *Journal of Consulting and Clinical Psychology*, *82*(2), 287–297. https://doi.org/10.1037/a0035535

Lutz, W., Stulz, N., & Köck, K. (2009). Patterns of early change and their relationship to outcome and follow-up among patients with major depressive disorders. *Journal of Affective Disorders*, *118*(1-3), 60–68. https://doi.org/10.1016/j.jad.2009.01.019

Matsuzawa, Y., Lee, Y. S. C., Fraser, F., Langenbahn, D., Shallcross, A., Powers, S., . . . Minen, M. (2019). Barriers to Behavioral Treatment Adherence for Headache: An Examination of Attitudes, Beliefs, and Psychiatric Factors. *Headache*, *59*(1), 19–31. https://doi.org/10.1111/head.13429

McNeish, D. M. (2015). Using Lasso for Predictor Selection and to Assuage Overfitting: A Method Long Overlooked in Behavioral Sciences. *Multivariate Behavioral Research*, *50*(5), 471–484. https://doi.org/10.1080/00273171.2015.1036965

Mohr, D. C., Siddique, J., Ho, J., Duffecy, J., Jin, L., & Fokuo, J. K. (2010). Interest in behavioral and psychological treatments delivered face-to-face, by telephone, and by internet. *Annals of Behavioral Medicine : a Publication of the Society of Behavioral Medicine*, *40*(1), 89–98. https://doi.org/10.1007/s12160-010-9203-7

Muthén, B., Asparouhov, T., Hunter, A. M., & Leuchter, A. F. (2011). Growth modeling with nonignorable dropout: Alternative analyses of the STAR*D antidepressant trial. *Psychological Methods*, *16*(1), 17–33. https://doi.org/10.1037/a0022634

Muthukrishnan, R. (2016). Lasso: A Feature Selection Technique In Predictive Modeling For Machine Learning. In R. Rajeswari (Ed.), *2016 IEEE International Conference on Advances in Computer Applications: 2016 IEEE ICACA : 24th October 2016* (pp. 18–20). Piscataway, NJ: IEEE.

National Collaborating Centre for Mental Health. (2011). *Common Mental Health Disorders: Identification and Pathways to Care: NICE Clinical guidlines, No. 123*. Leicester (UK): British Psychological Society.

Nordberg, S. S., Castonguay, L. G., Fisher, A. J., Boswell, J. F., & Kraus, D. (2014). Validating the rapid responder construct within a practice research network. *Journal of Clinical Psychology*, *70*(9), 886–903. https://doi.org/10.1002/jclp.22077

Nylund, K. L., Asparouhov, T., & Muthén, B. O. (2007). Deciding on the Number of Classes in Latent Class Analysis and Growth Mixture Modeling: A Monte Carlo Simulation Study. *Structural Equation Modeling: A Multidisciplinary Journal*, *14*(4), 535–569. https://doi.org/10.1080/10705510701575396

Ram, N., & Grimm, K. J. (2009). Growth Mixture Modeling: A Method for Identifying Differences in Longitudinal Change Among Unobserved Groups. *International Journal of Behavioral Development*, *33*(6), 565–576. https://doi.org/10.1177/0165025409343765

Robinson, L., Delgadillo, J., & Kellett, S. (2019). The dose-response effect in routinely delivered psychological therapies: A systematic review. *Psychotherapy Research : Journal of the Society for Psychotherapy Research*, 1–18. https://doi.org/10.1080/10503307.2019.1566676

Romijn, G., Batelaan, N., Kok, R., Koning, J., van Balkom, A., Titov, N., & Riper, H. (2019). Internet-Delivered Cognitive Behavioral Therapy for Anxiety Disorders in Open Community Versus Clinical Service Recruitment: Meta-Analysis. *Journal of Medical Internet Research*, *21*(4), e11706. https://doi.org/10.2196/11706

Rubel, J., Lutz, W., Kopta, S. M., Köck, K., Minami, T., Zimmermann, D., & Saunders, S. M. (2015). Defining early positive response to psychotherapy: An empirical comparison between clinically significant change criteria and growth mixture modeling. *Psychological Assessment*, *27*(2), 478–488. https://doi.org/10.1037/pas0000060

Schibbye, P., Ghaderi, A., Ljótsson, B., Hedman, E., Lindefors, N., Rück, C., & Kaldo, V. (2014). Using early change to predict outcome in cognitive behaviour therapy: Exploring timeframe, calculation method, and differences of disorder-specific versus general measures. *PloS One*, *9*(6), e100614. https://doi.org/10.1371/journal.pone.0100614

Schröder, J., Berger, T., Meyer, B., Lutz, W., Hautzinger, M., Späth, C., . . . Moritz, S. (2017). Attitudes Towards Internet Interventions Among Psychotherapists and Individuals with Mild to Moderate Depression Symptoms. *Cognitive Therapy and Research*, *41*(5), 745–756. https://doi.org/10.1007/s10608-017-9850-0

Schuster, R., Berger, T., & Laireiter, A.-R. (2018). Computer und Psychotherapie – geht das zusammen? *Psychotherapeut*, *63*(4), 271–282. https://doi.org/10.1007/s00278-017-0214-8

Schwarz, G. (1978). Estimating the Dimension of a Model. *The Annals of Statistics*, *6*(2), 461–464. https://doi.org/10.1214/aos/1176344136

Stulz, N., Gallop, R., Lutz, W., Wrenn, G. L., & Crits-Christoph, P. (2010). Examining differential effects of psychosocial treatments for cocaine dependence: An application of latent trajectory analyses. *Drug and Alcohol Dependence*, *106*(2-3), 164–172. https://doi.org/10.1016/j.drugalcdep.2009.08.009

Stulz, N., Lutz, W., Kopta, S. M., Minami, T., & Saunders, S. M. (2013). Dose-effect relationship in routine outpatient psychotherapy: Does treatment duration matter? *Journal of Counseling Psychology*, *60*(4), 593–600. https://doi.org/10.1037/a0033589

Thibodeau, M. A., Quilty, L. C., Fruyt, F. de, Bolle, M. de, Rouillon, F., & Bagby, R. M. (2015). Latent classes of nonresponders, rapid responders, and gradual responders in depressed outpatients receiving antidepressant medication and psychotherapy. *Depression and Anxiety*, *32*(3), 213–220. https://doi.org/10.1002/da.22293

Tibshirani, R. (1996). Regression Shrinkage and Selection Via the Lasso. *Journal of the Royal Statistical Society: Series B (Methodological)*, *58*(1), 267–288. https://doi.org/10.1111/j.2517-6161.1996.tb02080.x

Titov, N., Dear, B. F., Johnston, L., Lorian, C., Zou, J., Wootton, B., . . . Rapee, R. M. (2013). Improving adherence and clinical outcomes in self-guided internet treatment for anxiety and depression: Randomised controlled trial. *PloS One*, *8*(7), e62873. https://doi.org/10.1371/journal.pone.0062873

Wampold, B. E. (2015). How important are the common factors in psychotherapy? An update. *World Psychiatry : Official Journal of the World Psychiatric Association (WPA)*, *14*(3), 270–277. https://doi.org/10.1002/wps.20238

Yang, M., & Maxwell, S. E. (2014). Treatment effects in randomized longitudinal trials with different types of nonignorable dropout. *Psychological Methods*, *19*(2), 188–210. https://doi.org/10.1037/a0033804

Zarski, A.-C., Berking, M., Reis, D., Lehr, D., Buntrock, C., Schwarzer, R., & Ebert, D. D. (2018). Turning Good Intentions Into Actions by Using the Health Action Process Approach to Predict Adherence to Internet-Based Depression Prevention: Secondary Analysis of a Randomized Controlled Trial. *Journal of Medical Internet Research*, *20*(1), e9. https://doi.org/10.2196/jmir.8814

ERKLÄRUNG

Ich erkläre hiermit, dass ich die vorliegende Arbeit ohne unzulässige Hilfe Dritter und ohne Benutzung anderer als der angegebenen Hilfsmittel angefertigt habe. Die aus anderen Quellen direkt oder indirekt übernommenen Daten und Konzepte sind unter Angabe der Quelle gekennzeichnet.
Bei der Auswahl und Auswertung der in dieser Dissertation dargestellten Methoden und Experimente haben mir die nachstehend aufgeführten Personen in der jeweils beschriebenen Weise entgeltlich/unentgeltlich geholfen:

1. W. Lutz war verantwortlich für das Konzept und das Datenmanagement und hat das finale Manuskript kritisch durchgesehen. A. Arndt hat die Auswertung durchgeführt und das Skript geschrieben. J. Rubel plante die Datenauswertung und korrigierte das Manuskript. J. P. Klein plante und koordinierte die Studie. J. Schröder war ebenfalls an der Koordination beteiligt und entwickelte den APOI- Fragebogen. B. Meyer war verantwortlich für die Kooperation mit GAIA AG, die die Intervention für Patienten zugänglich machte. C. Späth und T. Berger waren ebenfalls an Planung und Koordination der Studie beteiligt. Alle anderen Autoren beteiligten sich an der Rekrutierung und der Erhebung der Daten.

2. A. Arndt was verantwortlich für das Konzept, die Datenauswertung und das Verfassen den Manuskripts. W. Lutz war verantwortlich für das Datenmanagement und hat das finale Manuskript kritisch durchgesehen. J. Rubel hat zur Auswertung beigetragen, und hat das finale Manuskript kritisch durchgesehen. J. P. Klein plante und koordinierte die Studie. J. Schröder war ebenfalls an der Koordination beteiligt und entwickelte den APOI-Fragebogen. B. Meyer war verantwortlich für die Kooperation mit GAIA AG, die die Intervention für Patienten zugänglich machte. C. Späth und T. Berger waren ebenfalls an Planung und Koordination der Studie beteiligt.Alle anderen Autoren beteiligten sich an der Rekrutierung und der Erhebung der Daten.

3. A. Arndt was verantwortlich für das Studien-Design, die Datenauswertung und das Verfassen den Manuskripts. J. Rubel trug bei zum Studiendesign und zur Auswertung und hat das finale Manusript kritisch durchgesehen. W. Lutz war mitverantwortlich für das Studien Design und die Implementierung und hat das finale Manusript kritisch durchgesehen. T. Berger war verantwortlich für das Konzept und die Entwicklung der Intervention und hat das finale Manusript kritisch durchgesehen.

Weitere Personen waren an der inhaltlich-materiellen Erstellung der vorliegenden Arbeit nicht beteiligt. Insbesondere habe ich hierfür nicht die entgeltliche Hilfe von Vermittlungs- bzw. Beratungsdiensten (Promotionsberater oder andere Personen) in Anspruch genommen. Niemand hat von mir unmittelbar oder mittelbar geldwerte Leistungen für Arbeit erhalten, die im Zusammenhang mit dem Inhalt der vorliegenden Dissertation stehen.

Die Arbeit wurde bisher weder im In- noch im Ausland in gleicher oder ähnlicher Form einer anderen Prüfungsbehörde vorgelegt.

Ich versichere die Richtigkeit der vorangegangenen Erklärung und bin mir der strafrechtlichen Folgen einer Falschaussage bewusst.

....

Ort, Datum Unterschrift