Measurement-Based and Data-Informed Psychological Therapy

Lutz, W., Schwartz, B., & Delgadillo, J. (2021). Measurement-based and data-informed psychological therapy. *Annual Review of Clinical Psychology, 18*. Advance online publication. https://doi.org/10.1146/annurev-clinpsy-071720-014821

Wolfgang Lutz
Department of Psychology University of Trier
E-mail: lutzw@uni-trier.de
ORCiD: 0000-0002-5141-3847

Brian Schwartz
Department of Psychology, University of Trier
E-mail: schwartzb@uni-trier.de
ORCiD: 0000-0003-4695-4953

Jaime Delgadillo
Clinical and Applied Psychology Unit, Department of Psychology, University of Sheffield
E-mail: j.delgadillo@sheffield.ac.uk
ORCiD: 0000-0001-5349-230X

Corresponding author contact information:
Wolfgang Lutz, Ph.D.
Clinical Psychology and Psychotherapy
Department of Psychology
University of Trier
D-54296 Trier, Germany
E-mail: wolfgang.lutz@uni-trier.de
Phone: 0049 651 201 2883

Keywords

Measurement-based psychological therapy, data-informed, routine outcome monitoring, feedback research, prediction, statistical and clinical decision-making, clinical navigation systems

Abstract

Outcome measurement in the field of psychotherapy has developed considerably in the last decade. This paper discusses key issues relating to outcome measurement, modeling, and implementation of data-informed and measurement-based psychological therapy. First, an overview is provided, covering the rationale of outcome measurement by acknowledging some of the limitations to clinical judgement. Second, different models of outcome measurement are discussed, including pre-post, session-by-session, and higher resolution intensive outcome assessments. Third, important concepts related to modeling patterns of change are addressed, including early response, dose–response, and non-linear change. Furthermore, rational and empirical decision tools are discussed as the foundation for measurement-based therapy. Fourth, examples of clinical applications are presented, which show great promise to support the personalization of therapy and to prevent treatment failure. Finally, we build on continuous outcome measurement as the basis for a broader understanding of clinical concepts and data-driven clinical practice in the future.

INTRODUCTION

"A whole class of loosely related errors made in the clinical case conference arises from forgetting (on the part of the psychologist) or never having learned (in the case of the psychiatrist and social worker) certain elementary statistical or psychometric principles" (Meehl 1973, p. 232)

The above quote by Paul Meehl from his essay "Why I do not attend case conferences" is more than 40 years old, however, the topic of measurement in clinical practice is more current than ever. In recent decades, numerous randomized clinical trials (RCTs) have helped to establish psychological therapies as an effective intervention to treat a broad range of psychological disorders (Barkham & Lambert 2021, Cuijpers et al. 2019). As a consequence, the implementation of psychological treatments has become an important asset in most health service systems around the world. However, research demonstrating their average effectiveness does not mean that psychological treatments work for all patients under all circumstances. Furthermore, negative treatment response is usually missed in clinical trials and meta-analytic reviews focusing on average effects. As a consequence, monitoring actual clinical progress in routine care and using such data to improve treatment is a necessary supplement to psychological therapy implementation and has the potential to substantially change the way we think about psychotherapy as a science.

Outcome measurement has developed impressively over the past decade. New technological developments such as the internet or app-based data assessment and feedback tools have made outcome measurement easier to implement than in the past. Recent developments in outcome measurement make use of the information to improve the clinical decision-making process by grounding clinical practice in empirical data. For example, psychometric and demographic data can be used by clinicians to personalize the selection of therapeutic techniques and to monitor a patient's response to therapy in real-time (Lutz et al.

2021a). Therefore, outcome measurement can be seen as an important and integral part of clinical competence, practice, and training. This is comparable to many other areas in the healthcare system, where continuous monitoring of health indicators are common in day-to-day clinical practice (e.g., fever, blood pressure). In this sense, continuous outcome measurement forms the basis of modern measurement-based and data-informed psychological therapy (Delgadillo & Lutz 2020, Lutz et al. 2021b).

Decision-Making in Psychological Therapy

The assessment of treatment response is an issue that therapists have been grappling with since the early history of psychotherapy. Despite their theoretical differences, influential theorists like Sigmund Freud, B.F. Skinner, A.T. Beck, Carl Rogers, John Bowlby and others recognised that psychological disturbances cannot be observed directly, but are rather latent (i.e., hidden) phenomena that can be inferred from behaviour, introspection, and self-report. In many traditions of psychological treatment, latent change is assumed to operate at two levels: at the process level (i.e., maintaining factors or conflicts) and the consequence level (i.e., symptoms, interpersonal and behavioural problems). As such, the practice of psychotherapy implicitly or explicitly involves the monitoring of these changes to determine whether treatment is working as expected and whether it will ultimately benefit the patient. The central theme of this review concerns the continuous monitoring of chang, which is referred to as routine outcome monitoring in contemporary literature, and the monitoring of related processes of change.

Broadly speaking, therapists can monitor treatment response via qualitative (i.e., as part of the therapy dialogue) and quantitative methods (i.e., using validated questionnaires or idiographic measures). Such information must somehow be processed and interpreted to inform treatment decisions. Clinical decision-making is based on one of two general approaches: clinical judgement or actuarial methods. The first approach relies on intuition, which is a function of the therapist's way of processing (qualitative / quantitative) data, their clinical experience, theoretical orientation and consultation with others. This method has been referred to as "informal" (Grove & Meehl 1996), since it is not usually based on a structured set of decision rules or equations, and it is highly variable within (i.e., changes over time) and between therapists. The second approach, referred to as "formal or algorithmic", involves using structured decision rules or equations to combine and interpret data in order to reach a clinical judgement or decision.

In a review of decision-making in the field of clinical psychology, Garb (2005) argued that psychologists' judgments are error-prone due to the influence of heuristics and biases, as defined in Tversky and Kahneman's (1974) seminal work. For example, the representativeness heuristic may partly explain the modest interrater reliability of therapists' diagnostic assessments, as these judgments are influenced by how closely a patient resembles the diagnostic "prototype" that each assessor uses to make a diagnosis (Evans et al. 2002). The influence of selection and confirmation biases in therapists' clinical interviews and interactions is a related issue, as therapists vary widely in the extent to which they gather key information (e.g., symptoms) to determine a diagnosis (Miller et al. 2001). Another bias that has received attention in the field of psychotherapy is clinical over-optimism, as exemplified in the classic study by Hannan et al. (2005) where therapists' prognostic assessments of their patients significantly underestimated the number of patients that eventually made little improvement and those who deteriorated after therapy (as determined using psychometric patient-reported data). In another example, Walfish et al. (2012) showed that therapists estimate about 85% of their patients have improved or recovered, a success rate far higher than those found based on measured outcomes in clinical trials and routine care. Furthermore, in comparison to other therapists, 90% of clinicians rate themselves in the upper quartile of successful therapists and none consider themselves below average. This overly positive self-assessment is known as the *better-than average effect* and can also be found in other areas and professions (e.g., Zell et al. 2020). In part, these clinical intuition biases are a result of therapists having to rely on indirect (i.e., proxy indicators of latent factors) and often subjective information (i.e., the patient's experiences narrated verbally) to make judgements. Experts in the field of decision-making agree that accurate intuition can be developed in situations that provide regular and highly objective feedback, enabling judges to refine their pattern recognition abilities (Kahneman & Klein 2009). Hence, it is not surprising that clinical intuition is not highly reliable in psychotherapy, in the absence of systematic routine outcome monitoring and empirically derived decision-support tools.

Therefore, there is a strong case to advocate the additional use of formal and empirical methods. Of course, formal methods of routine outcome monitoring are also imperfect, but they offer additional empirical support and improvement for clinical decision-making. The aim of this paper is to review the history of formal routine outcome monitoring methods, their strengths and limitations, their implementation and future development. Over the years, several terms have been used to describe this line of research, including practice-oriented research, patient-focused research, practice-based evidence, as well as routine outcome monitoring (Barkham & Lambert 2021; Castonguay et al. 2013, 2021; Lutz et al. 2021a). In this paper, we use the terms patient-focused feedback research, measurement-based as well as data-informed psychological therapy. These terms highlight the clinical decision support function of such endeavours and the new trend to use such data to personalize treatment (Delgadillo & Lutz 2020).

MEASURING OUTCOMES

Supporting clinical decision-making with empirical data can be seen as a major advance in the field of psychotherapy. This process can rely on sparse data (i.e., collected pre and post-therapy) or continuous and intensive data collection (i.e., collected during therapy). In this section we first present a theoretical framework for data-informed psychological therapy. Then we describe the classical pre-post treatment method before moving on to multiple assessments, session-by-session assessments, and intensive longitudinal assessments.

Practice-Based Evidence and Clinical Decision-Making

Figure 1 displays a matrix of assessment modes in clinical practice and their relation to empirically-supported clinical decision-making. The matrix is based on Hammond’s (1978) modes of inquiry in evaluation research and includes five modes and six dimensions. It has been adapted to describe modes and dimensions of measurement-based and data-informed psychological therapy. The six dimensions include (a) the potential for data-informed psychological therapy based on the information provided by the assessment structure, ranging from very low to very high; (b) the extent of empirically-based support to generate practice-based evidence for clinical decision-making (Barkham & Lambert 2021); (c) the challenge of implementing specific data collection modes into clinical routine, ranging from less challenging to very challenging; (d) the intensity of psychometric training, which is necessary to make use of the collected information; (e) the usefulness of the data collection strategy for practice-oriented research (Castonguay et al. 2013), and (f) the mode of cognition, ranging from analytical to intuitive thought.

Mode 1 represents the predominant use of intuitive clinical decision-making, not influenced by outcome measures. This is the traditional psychotherapy mode and likely represents the most common mode to date. Therapists working in this mode adapt their clinical approach based on experience with similar patients, while often simultaneously following guidelines set out by professional organisations, employers or policy makers. This may be a conscious choice (i.e., a preference for intuitive decision-making) or may be due to lack of training, resources or time to support routine outcome monitoring. Nevertheless, this means that clinical decision-making is mainly conducted without routine outcome assessments. Furthermore, in mode 1, empirically-based decision support is non-existent and practice-based evidence is not systematically generated.

Pre-Post and Repeated Assessments

Mode 2 represents the first step of implementing data into clinical practice. Data collection and clinical decision-making is based on a pre-post treatment design (or direct collection of post-treatment information only). This mode allows conclusions to be drawn about the effectiveness of treatments under routine conditions and is relatively easy to implement in clinical practice. However, this form of assessment is often limited by substantial missing data, as patients often drop-out, leave treatment early, or do not fill out the final questionnaire for other reasons. This becomes an even larger issue when longer-term follow-up assessments are added. Mode 3 includes repeated measurements over the course of treatment, for example every 4 or 5 sessions or weeks. Such a repeated measurement setup has greater potential to support clinical decision-making and research. For example, expected treatment response models (covered later on) can be developed and used to monitor and adjust therapy in real-time. These modes require basic training and knowledge of psychometrics in order to make use of the data for clinical decisions regarding individual patients, groups or services.

Session-by-Session Assessments

Most studies on routine outcome monitoring and the effectiveness of psychometric feedback in clinical practice focus on assessments on a weekly or session-by-session basis (see mode 4 in Figure 1). Such an assessment structure provides the context for the effective use of measurement-based treatments, especially in combination with clinical support tools for patients with a high risk for treatment failure (e.g., Lutz et al. 2021b). According to Figure 1, here the potential for data-informed psychological treatment is high. However, more psychometric knowledge and therapist training is needed in order to make use of the information and integrate it into clinical practice. Of course, the implementation of a session-by-session outcome assessment structure is more challenging, however the potential of the generated information for further research is also high.

Furthermore, this mode includes a data-generation strategy, which has laid the foundation for several new developments in clinical research over the last decade. Usually clinical concepts are developed based on theory, or intuitively based on expert clinical experience and then tested in RCTs (or meta-analyses) applied to evaluate the effectiveness of treatment orientations or packages (Cristea et al. 2021). However, there are limitations to thinking about clinical concepts of psychological therapies exclusively based on RCTs and meta-analyses, as was recently pointed out by Baldwin and Imel (2020). The categorization of different variants or orientations of psychological therapy can be difficult (in comparison, for example, to most somatic treatments), and is usually based on arbitrary boundaries and theoretical arguments about the predominance of certain change processes being more or less relevant in specific forms of treatment. Therefore, this line of research requires a complementary approach that can be routinely implemented in clinical practice and which continuously uses research results to improve treatment outcomes in routine clinical practice (e.g., Howard et al. 1996, Lambert et al. 2001). Therapists not only deliver treatment, but simultaneously collect data by assessing patients' progress session-by-session over the course of treatment (Lutz et al. 2019). Large databases generated in this way can then be analysed using new statistical tools (e.g., machine learning algorithms) to develop clinical decision-support tools. This allows therapists not only to track progress (and potential side effects) on an individual level, but also research on the variability of change as a function of patient characteristics, treatment interventions, therapists, and clinics/services (Barkham & Lambert 2021). For example, on an aggregated level, the data can serve to investigate side effects, treatment dosage, therapist effects and service effects (e.g., McKay & Jensen-Doss 2021). The session-by-session assessment structure is accompanied by statistical and methodological developments to analyze the nested structure of longitudinal data (e.g., Raudenbush & Bryk 2002). Such multilevel models allow a better disentanglement of variance components on session, patient, therapist, and service levels. They also allow the disentanglement of between-patient and within-patient variance components to study mechanisms of change (e.g., Hamaker & Wichers 2017).

On a larger scale, such session-by-session information holds the potential to establish a community or national data collection system to which independent research groups can be granted access. Such endeavours have just begun to emerge in the field, as exemplified by the *Improving Access to Psychological Therapies* (IAPT) initiative in England (see Clark 2018). In combination with technological innovations, session-by-session assessments and clinically informative data-driven decision support tools can facilitate the development of large scientist-practitioner infrastructures.

Intensive Longitudinal Assessments

Despite the advances that session-by-session assessments have made possible for data-informed therapy, in recent years, even more fine-grained measurements have been applied in psychotherapy research. Proponents of these intensive longitudinal assessments have argued that higher resolution is necessary to obtain useful information about dynamic changes in patients' everyday experiences.

Data from such high-frequency measurements can be obtained via Ecological Momentary Assessment (EMA), a real-time, within-subject assessment allowing measurements several times a day (mode 5 in Figure 1). EMA has also been referred to as *ambulatory assessment*, or *experience sampling* (Ebner-Priemer & Trull 2009). During the last two decades, technological innovations have drastically facilitated these measurements, allowing laborious paper-and-pencil diaries to be replaced by technological devices such as smartphones, smartwatches, or other mobile hi-tech devices (Miller 2012). EMA circumvents the problem of retrospective bias, is more ecologically valid as it can be conducted within patients' daily lives, has higher resolution and can thus better represent intrapersonal processes (Ebner-Priemer & Trull 2009). It can simultaneously collect psychological, physiological, and behavioral variables, allowing the examination of situation-specific relationships and feedback to the patient in real-time. Furthermore, EMA is face-valid, convenient, and unobtrusive for patients (Miller 2012).

EMA is most often used to collect self-report questionnaires, which patients complete on their own or a study smartphone. The ability to collect psychometric measures multiple times a day supports data-informed therapy and in principle allows for a high level of decision support (Figure 1). In psychotherapy research, intensive longitudinal assessments have been used to improve the prediction of early change (Husen et al. 2016), treatment response (Fisher et al. 2019, Wichers et al. 2012), and treatment dropout (Lutz et al. 2018). At the same time, more therapist training in psychometrics is necessary compared to lower frequency assessment methods, and implementation of the assessments is more challenging, often leading to higher dropout rates during assessments. Furthermore, when implementing EMA, decisions must be made about the duration, frequency, scope (number of items), and the timing of assessments. To date, intensive longitudinal assessments have been implemented very heterogeneously and a consensus on EMA survey design standards is lacking. However, some authors have already offered suggestions to improve transparency and user acceptability (e.g., Fisher et al. 2021). An important limitation of intensive longitudinal assessments concerns patient burden (completion of questionnaires), where longer questionnaires, rather than higher assessment frequency seems to increase the burden and affect data quantity and quality (Eisele et al. 2020).

In addition to the high-frequency collection of psychometric measures, EMA can also be used to collect biological and physiological measures (e.g., heart rate, electrodermal activity, motor activity, sleep, etc.). For example, wristbands that collect physiological parameters are already better than chance at predicting epileptic seizures (Meisel et al. 2020). Smartwatches and associated mobile sensors enable the passive and continuous collection of huge amounts of data within patients' daily lives with minimal patient burden (e.g., Jacobson et al. 2019, Wright & Woods 2020). This has been referred to as *digital phenotyping* and has recently been applied in psychotherapy research to identify diagnostic groups, symptom severity, patterns of change, as well as to inform treatment selection (e.g., Hehlmann et al. 2021, Jacobson et al. 2019, Webb et al. 2021).

Intensive longitudinal data can also help to optimise the assessment of therapy processes during treatment. This allows intra-individual changes to be assessed with a higher resolution and standardized change measurements to be supplemented by idiographic approaches. Intensive longitudinal assessments are also necessary for the continuous evaluation of video and audio recordings of a therapy session. For example, physical movements of the patient and therapist (Ramseyer & Tschacher 2011), speech content and prosodic features (Imel et al. 2014, Paz et al. 2021) as well as emotions based on gestures and facial expressions (Baur et al. 2020) can be recorded automatically. This allows a patient, a patient-therapist dyad and the entire therapeutic process to be captured not only unidimensionally based on known pre-therapy self-report items, but multidimensionally based on high-frequency data from different modalities (e.g., audio, video, self-report, or physiology).

Recently, data-driven tools have been developed to provide therapists with personalized feedback by dynamic visualizations of intensive longitudinal data (e.g., Bringmann et al. 2020, Hehlmann et al. 2021). However, these approaches still have to contend with a number of problems in addition to the challenging implementation and large training effort. These data require the use of sophisticated statistical analysis methods, which are currently still being developed and refined (Bringmann 2021). In addition, there are many analytical options and decisions that need to be made during data preparation and evaluation, whereby small changes can lead to significant differences in the results. Therefore, these approaches are currently more likely to be found in the context of pilot and proof-of-concept studies.

Summary

Practice-based evidence and clinical decision-making are based on data assessed before, during, and after treatment. In measurement-based and data-informed psychological therapy, outcomes are measured and observed at varying frequencies: (1) Pre-post-assessments are easy to implement, but are limited by missing data and can only represent simple changes; (2) Repeated measures throughout treatment can help to model patterns of change, to inform treatment decisions, and to enable psychometric feedback; (3) Session-by-session assessments allow progress to be tracked on an individual level to develop clinical support tools and examine variability in change as a function of patient, therapist or treatment characteristics; and (4) Intensive longitudinal assessments can improve the representation of high-frequency intrapersonal processes in patients' daily lives. Currently, however, these high-frequency measurements still face several problems, which is why session-by-session assessment remains the best suited method to assess treatment progress and to inform decisions in routine care.

MODELING PATTERNS OF CHANGE

Interest in how patients change over time has been documented for several decades in the field of psychotherapy, although earlier investigations were hampered by unreliable methods of assessing longitudinal patterns of change. In the last 35 years, however, a large body of empirical research on this topic has accumulated. To a great extent, the development of brief psychometric measures that could be regularly collected during psychotherapy led to important methodological and conceptual advances. Patients are not sufficiently described by their diagnosis, as they are quite different to each other across several other demographic, clinical and interpersonal characteristics that are related to treatment response (e.g., Lutz et al. 2021a). These individual differences influence the longitudinal trends that characterise symptomatic change over time, as well as symptom fluctuations that can be used to develop empirically-supported clinical decision rules. These concepts and related findings are discussed below.

Dose–Response

Research on trajectories of change can be traced back to Ken Howard's seminal examination of the dose–response effect (Howard et al. 1986). Alluding to the pharmacological notion of a relationship between the dose of medication and its expected effects, Howard et al. (1986) aggregated routine outcomes data from N = 2,431 psychotherapy patients across 15 studies and modelled the statistical relationship between the number of treatment sessions (dose) and symptomatic improvements. This relationship was nonlinear, characterised by a negatively accelerating (log-linear) trend where most of the improvements were observed in the earlier sessions, showing diminishing improvements thereafter. They proposed that an *optimal dose of therapy* could be defined as an interval, between the point at which at least 50% of cases respond to treatment, and the point after which response rates plateau, which in their study was around 8 to 26 sessions. Conceptually, the dose–response notion is based on several assumptions: therapy sessions cause change, they have a cumulative effect (more sessions are better), but the potency of this effect diminishes over time. This study motivated a surge of investigations over the following decades, in clinical samples with various diagnoses and treatment modalities (e.g., Barkham et al. 1996, Baldwin et al. 2009). Despite the diversity of statistical methods applied throughout these studies, a systematic review of over 20 dose–response studies (Robinson et al. 2020a) concluded that a curvilinear relationship between treatment duration and outcomes has been extensively replicated, although the “optimal dose” apparently varies according to clinical heterogeneity (e.g., impairment level), treatment intensities (e.g., guided self-help vs. psychotherapy), and treatment settings (e.g., student counselling, outpatient or inpatient care). Consistent with these assumptions, more recent investigations (Robinson et al. 2020b) in highly standardized treatments and more diagnostically homogenous samples report differential dose–response patterns according to diagnosis (e.g., PTSD requires lengthier treatment than generalized anxiety disorder) and treatment intensity (e.g., the dose–effect of guided self-help plateaus sooner than that of CBT).

Additional research has revealed that not all patients follow a uniform dose-effect relation and subgroups of patients follow different latent trends of change (e.g., Lutz et al. 2014, Owen et al. 2015). Such findings motivated the proposal of an alternative perspective referred to as the “good-enough level” (GEL) model. Barkham et al. (1996) argued that the classic log-linear model could be a statistical artefact that results from the aggregation of data from cases that in fact have heterogeneous treatment response patterns (i.e., early responders, gradual responders, non-responders with lengthy treatments, and those that drop out). Rather than assuming uniformity across all patients, the GEL model assumes that the number of sessions needed to attain improvement varies from case to case, resulting in a process of “responsive regulation” of treatment duration (Barkham et al. 2006, Stiles et al. 1998). Several studies have provided support for this model, although studies comparing the goodness-of-fit between the dose-response and GEL models have tended to yield mixed findings (e.g., Stulz et al. 2013). A systematic review of fifteen studies in this area found empirical support for some of the assumptions of the GEL model, such as the observation that higher intake severity tends to require lengthier treatments, and the observation that some subgroups of cases show curvilinear change while others show linear change (Bone et al. 2021a).

Overall, the dose-effect and GEL models count with empirical support and together indicate that (1) subgroups of patients change in similar ways; (2) change is most often non-linear; (3) intake severity influences the linearity and duration of treatment; (4) although some patients respond early and others more gradually, the net benefit of therapy tends to occur within a predictable window of time; (5) patients who remain in treatment beyond a typical optimal dose without showing reliable change are more likely to be non-responders. The precise operationalisation of the latter two points, however, remains a matter of controversy and debate in the field (e.g., Nordmo et al. 2020, Robinson et al. 2020a). Consequently, predicting and monitoring which patients are likely to require shorter or lengthier interventions is a goal that is empirically justified and clinically important.

Early Response

Early response refers to symptomatic improvement during the initial phase of treatment (usually the first month[1]). Over the last three decades, numerous studies have investigated early response and its prognostic value in different patient conditions such as major depressive disorder, panic disorder, generalized anxiety disorder, eating disorders, and in samples of patients with heterogeneous presenting problems (e.g., Chang et al. 2021, Delgadillo et al. 2014, Lutz et al. 2014, Moggia et al. 2020). A systematic review aggregated available data from 15 studies in a random effects meta-analysis and reported a large pooled effect size (g = 0.87; Beard & Delgadillo 2019). Early responders had significantly better post-treatment outcomes compared to patients without early response. The detected effect seems to be larger in anxiety measures (g = 1.37) compared to depression measures (g =

[1] Early response is often conceptualized by reliable and/or clinically significant change (Jacobson & Truax 1991). However, several studies have also used growth mixture models (GMM) to identify patient early change patterns (e.g., Lutz et al. 2014, Moggia et al. 2020).

0.76). Furthermore, early response seems to be a phenomenon that exists in a variety of treatment models, such as CBT, cognitive-behavioral analysis system of psychotherapy (CBASP), behavioral activation (BA), interpersonal psychotherapy (IPT), psychodynamic therapy, guided self-help, internet-based CBT, and group therapy (Beard & Delgadillo 2019, Moggia et al. 2020).

Expected Treatment Response

The development of expected treatment response (ETR) models was key to the rise of measurement-based care. The ETR concept was first proposed by Lutz et al. (1999), who used archival data from a large routine outcome monitoring database and generated an expected recovery curve for individual patients based on growth curve modelling and seven predictors (well-being, symptoms, life functioning, past use of therapy, problem duration, treatment expectations, global assessment of functioning). Finch et al. (2001) applied a similar ETR model including 80% tolerance intervals around the predicted ETR curve in order to identify the 10% of deteriorating cases (outside the upper tolerance interval, also referred to as a *failure boundary*). This model was applied as an empirical decision rule in several progress feedback studies to identify patients who are on track (OT) - meaning their symptom change is within the expected course of treatment, or not on track (NOT) - meaning their symptom change is not sufficient and crosses the upper failure boundary (e.g., Lambert 2017).

After these original applications, further ETR approaches were developed to model change and support clinical decision-making. For example, Lutz et al. (2005) modelled expected treatment response for patients using the machine learning algorithm nearest neighbors (NN). The NN approach allows the generation of an ETR model for each patient based on a similar reference group (nearest neighbors) with similar intake characteristics. The predictive accuracy of this advanced ETR model increased further when information on early response was included and ETR curves were dynamically adapted over the course of treatment (Bone et al. 2021b, Lutz et al. 2019). Furthermore, methods modeling latent growth trajectories for outcome and suicidal ideation (e.g., Growth Mixture Models, GMM) have also been developed to generate ETRs based on change patterns of similar subgroups (e.g., Kyron et al. 2019, Lutz et al. 2014, Saunders et al. 2016, Stulz et al. 2013). Other variations on the ETR concept include systems that model expected change patterns for type of modality and treatment. For example, Lutz et al. (2006) generated individual nearest neighbor ETR models for patients receiving CBT and patients receiving an integrative CBT and IPT treatment in order to predict the optimal modality before patients began treatment. Further ETR models have been developed to monitor and provide feedback on treatment drop-out, common factors and therapeutic processes such as the alliance, outcome expectations, empathy, and substance use (e.g., Crits-Christoph et al. 2012, McClintock et al. 2017, Miller et al. 2005).

Non-linear Change, Sudden Gains and Losses

As described above, individual patients' treatment response can substantially deviate from aggregated or predicted group trends and dismissing this phenomenon as error variance leads to a loss of information (e.g., Krause 2018). The investigation of sudden gains and losses is an area of research dealing with nonlinear change patterns (e.g., Tang & DeRubeis 1999). A reliable sudden gain or loss is defined as a significant symptomatic positive (gain) or negative (loss) change between two consecutive sessions during the course of treatment. In their seminal work on the topic, Tang & DeRubeis (1999) identified 79% of patients with (at least one) sudden gain in the group of patients with a positive outcome in cognitive behavioral therapy for depression. In most cases, these gains occur early in treatment, relating the phenomenon to early response. However, gains also occurred between later sessions and some therapists also seem to be better at facilitating and initiating such gains (Deisenhofer et al. 2021). Aderka et al. (2012) conducted a meta-analysis on 16 studies identifying sudden gains not only in major depressive disorder, but also in social anxiety disorder, post-traumatic stress disorder, and other anxiety problems as well as different treatment approaches (e.g., CBT, IPT and supportive expressive therapy). Overall, the analysis showed that sudden gains had an effect on primary outcome measures at post-treatment ($g = 0.62$) as well as follow-up ($g = 0.56$). Although only a few studies have been conducted on sudden losses, findings are still useful for the development of clinical decision rules (e.g., Krüger et al. 2014, Lutz et al. 2013). Patients with sudden losses seem to benefit less from treatment than patients with sudden gains or patients with no significant shifts. Furthermore, sudden losses are less frequent and more equally distributed over the course of treatment than sudden gains.

Monitoring Clinical Processes and Mechanisms

In most applications of measurement-based care, outcome parameters are tracked and fed back to therapists. Therefore, mainly predictors and moderators are relevant, and they are applied in the development of decision rules. While predictors and moderators provide information about for whom treatment works, mediators inform us about how treatment works (e.g., Kraemer et al. 2001). However, in recent years, several investigations using repeated assessments of process and outcome variables over the course of therapy (e.g., session-by-session) and multilevel modeling allowed advancements in the establishment of mechanisms and outcome relations by separating within- and between-patient variance components of the process-outcome relation (e.g., Falkenström et al. 2016). Therefore, several within- and between-patient associations between processes (e.g., coping skills, therapeutic alliance/ruptures, emotional involvement, competence and skills) and symptom change in psychological therapy have been successfully investigated (e.g., Crits-Christoph & Connolly Gibbons 2021; Gómez Penedo et al. 2021; Lorenzo-Luaces & DeRubeis 2018; Rubel et al. 2017, 2020; Strunk et al. 2012; Zilcha-Mano 2017).

Furthermore, measures on change mechanisms have also been added to applications of measurement-based care in order to enrich the clinical usefulness of such systems (e.g., McAleavey et al. 2021, Miller et al. 2005, Lutz et al. 2019). Several authors have recently argued that psychotherapy practice and research should focus on therapy processes instead of treatment schools or the classical illness model of psychopathology, and to move forward towards a process-based and transdiagnostic approach with a focus on monitoring and targeting central evidence-based therapeutic procedures and processes (e.g., Goldfried 2019, Hofmann & Hayes 2019, Lutz & Schwartz 2021).

Complex Models and Critical Instability

Critical transitions represent another line of research on non-linear patterns of change (Scheffer et al. 2009). These transitions refer to sudden changes in the state of a complex system, which thereby tips from one state (e.g., symptoms meeting diagnostic criteria for a mental disorder) to another distinct state (e.g., a full remission of symptoms). This phenomenon has already been observed in ecosystems, the climate, the financial market, and neurological disorders (e.g., Lenton et al. 2008, May et al. 2008).

Conceptually comparable, critical instability can be seen as a universal early warning signal (EWS) that indicates the system's increasing instability and thus the increasing probability of a change of state (Scheffer et al. 2009). These warning signals show the critical slowing down of a system (i.e., decelerated recovery from small perturbations) which occurs when a system is approaching a transition. For example, a flurry of obsessional thoughts and related emotions may increase temporarily, immediately preceding the transition towards a state of remission of symptoms of obsessive compulsive disorder (Heinzel et al. 2014). This can be measured in patient time-series data via, for example by detecting increasing autocorrelation and regression at low lags, increasing variance, changes in spectral properties, and increasing skewness and kurtosis (Dakos et al. 2012). However, only few studies have investigated intraindividual change in EWS (Wichers et al. 2016). For example, Olthof et al. (2020) found early warning signals in a large sample to be associated with sudden gains and losses. In a feasibility study on seven cases, time-varying change point autoregressive models (TVCP-AR) detected gradual and abrupt changes in continuously measured physiological stress levels that might improve outcome prediction (Hehlmann et al. 2021). However, actual intraindividual changes of EWS and their predictive power need to be investigated in future studies (Bringmann 2021).

Network models of psychopathology are also considered complex models. The network approach assumes that symptoms (e.g., affect, thoughts, or behaviour) influence each other rather than being triggered by an underlying latent disease factor (e.g., Borsboom 2017, Hofmann & Hayes 2019, Wright & Woods 2020). Network analysis models the connections (edges) between the symptoms (nodes) and allows the calculation of so-called centrality measures, which indicate the centrality of a variable within the network and thus its potential influence on other nodes. This approach was applied to identify relevant psychopathological variables via their centrality in a network and the dynamics among symptoms (Contreras et al. 2019). Beyond psychopathology, it was used in process research to identify bridges between intersession processes and symptom stress (Kaiser & Laireiter 2018) as well as in outcome research to predict treatment dropout (Lutz et al. 2018) and recovery from major depressive disorder (van Borkulo et al. 2015).

Recently, however, the fit of network models and centrality measures to psychological data has been questioned. Especially the idea of betweenness of a node, but also the assumption of symptoms as distinct nodes has been questioned (Bringmann 2021). Furthermore, networks of symptoms were found to be unstable, questioning their reliability and validity, and the methodological heterogeneity challenges comparability and interpretability of the results (Bringmann 2021). Therefore, more sophisticated models need to be developed and the acceptability of new applications needs to be evaluated under real-world clinical conditions (Epskamp 2020). For example, in a pilot study on twelve patient-therapist dyads, Frumkin et al. (2021) found that patients were willing to participate in these assessments, were interested in the results and tended to evaluate the data and models as useful, while therapists were less open to this method and not yet convinced of the added value of the results.

Summary

Four findings on patterns of change are especially important for outcome monitoring in clinical practice: (1) Change, outcome, and dropout are influenced by patient intake characteristics; (2) Change is influenced by patients' early response to treatment; (3) Change is influenced by nonlinear phenomena such as sudden gains and sudden losses; and (4) Probably most importantly, individual patient change can vary widely, deviating from general trends or symptom trajectories derived from clinical samples. This information can be used to identify patients at risk for treatment failure.

CLINICAL DECISION TOOLS

Identifying empirically-derived decision rules to support clinical decision-making is one of the central goals of clinical research (e.g., Meehl 1973). In this section, we provide examples of some of the key developments arising from routinely-collected data, and which enable data-informed psychological therapy and decision-making.

Rational versus Empirical Decision Rules

Two distinct approaches to developing decision rules based on empirical data have been applied in studies to monitor progress and tailor treatments to patients' needs (cf. Lambert et al. 2002). Rationally-derived models are based on clinical assessments and the predefined classification of patients. A widely used model is based on the concepts of reliable change and clinically significant improvement (Jacobson & Truax 1991). Typically, routine outcome monitoring is applied to enable a prognostic assessment (e.g., likelihood of response to treatment) based on symptomatic change since the initial intake assessment. Thereby, reliable change criteria are used as a classification rule. Studies applying these methods show that these classification rules reliably predict treatment response in diverse treatment settings and clinical populations (e.g., Delgadillo et al. 2014, Flood et al. 2019, Lutz et al. 2006). Another example of a rationally-derived model is the routine outcome monitoring method applied in IAPT services in England. In these services, all patients complete depression and anxiety questionnaires on a session-to-session basis and therapists apply conventional cut-off scores to assess treatment progress. Symptom reduction below the cut-off on at least one of the two measures indicates a response, while a reduction below both measures' cut-offs is considered a full recovery (Clark 2018).

In contrast, empirically-derived methods are based on prediction models and the above-described expected treatment response (ETR) concept and associated developments. Some newer applications also include prediction models and critical failure boundaries (beyond which the probability of negative treatment outcome is greater than positive outcome) for individual patients by dynamically re-calculating failure boundaries based on current progress data, i.e., assessments from previous treatment sessions (e.g., Bone et al. 2021b, Lutz et al. 2019). For example, Bone et al. (2021b) used session-by-session self-report depression and anxiety measures from n = 42,992 patients in IAPT services to train a dynamic prediction model (adapting failure boundaries over the course of treatment) using iterative logistic regression analysis. Subsequently, the model was evaluated on an external test sample of n=30,026 patients. This dynamic prediction model improved the accuracy of empirically-derived decision rules to identify patients at risk for treatment failure.

Prognostic Indices and Treatment Selection

Prognostic models can complement the therapist's clinical impression and thus support decision-making. Data-informed prognostic indices (PI) and prediction models based on patient information have recently been developed using machine learning approaches (see review by Chekroud et al. 2021). These prognostic indices and prediction algorithms can be used as criteria for treatment selection to identify the optimal treatment package or treatment strategy that suits the individual best. For example, Lorenzo-Luaces et al. (2017) found that patients with a worse prognosis, i.e. a higher PI, benefited more from CBT than from treatment as usual or brief therapy, while the treatment alternative made no difference if the overall prognosis was good.

Another option for treatment selection is to estimate the therapy outcome for each treatment alternative, and to recommend the treatment with the most favourable prognosis, i.e., with the best outcome prediction. For this purpose, Lutz et al. (2005) used a nearest neighbour model (NN) to predict patient-specific differential response to treatments. Another concept, the Personalised Advantage Index (PAI), was introduced by DeRubeis et al. (2014), which represents the difference of two predictions and quantifies the estimated superiority of one treatment model over another one (a PAI of zero means that both treatments should be equally effective for this patient). Patients who received their recommended treatment have been shown to have better outcomes in studies evaluating a wide range of treatment packages, e.g., cognitive, interpersonal, person-centered, psychodynamic therapy, antidepressant medication, eye movement desensitization and reprocessing (Cohen et al. 2021, Delgadillo et al. 2020, Deisenhofer et al. 2018, Huibers et al. 2015, Webb et al. 2019). Another line of research does not focus on predictions of different treatment packages, but rather on expected change in clinical modules, strategies and mechanisms (e.g., such as problem-solving and motivation-oriented interventions) in order to tailor clinical interventions within a treatment package (e.g., Gómez Penedo et al. 2021; Lutz et al. 2019; Ng et al. 2021; Rubel et al. 2018, 2020).

In addition, recent studies have begun to cross-validate treatment selection models using independent holdout data (e.g., Delgadillo & Gonzalez Salas Duhne 2020, Schwartz et al. 2021) and to data from another study (cross-trial validation; van Bronswijk et al. 2021), or to apply these models prospectively (Delgadillo et al. 2021, Lutz et al. 2021b).

In the future, predictions and recommendations based on pre-therapy assessments collected at one point in time could be complemented by EMA data (Fisher et al. 2019, Webb et al. 2021; see mode 5 on intensive longitudinal assessments above). Furthermore, variables from different data sources can be integrated multimodally to improve predictions and recommendations (e.g., clinical and video data; Atzil Slonim et al. 2021).

Therapist- Patient Matching

Variability in treatment outcomes between therapists (also known as therapist effects) has been extensively documented in the psychotherapy literature (Baldwin & Imel 2013). Even if they apply the same treatment protocol (i.e., CBT) for the same target problem (i.e., depression), some therapists attain impressive clinical outcomes compared to their peers, while others are less effective than average. Studies indicate that therapist effects may be partly related to patients' features, such that highly effective therapists are especially helpful for patients with severe levels of distress and risks of self-harm (Saxon & Barkham 2012). Furthermore, therapist effects also appear to be related to therapists' interpersonal skills (Heinonen & Nissen-Lie 2020) and particularly their ability to work effectively in highly challenging cases (Anderson et al. 2016). It logically follows that matching some patients to specific therapists could be a potentially effective method of treatment allocation. Recent research shows that, in fact, patients with certain combinations of demographic (e.g., employment status) and clinical features (e.g., comorbidity of depression and anxiety, symptom severity) respond better to some therapists compared to others (Delgadillo et al. 2020). On this basis, using archival clinical data, it is possible to predict the likely outcome for new patients assigned to specific therapists (i.e., % probability of recovery), potentially enabling an evidence-based approach to patient-therapist matching (Constantino et al. 2021).

Clinical Support Tools

In the context of routine outcome monitoring, clinical support tools (CST) have been developed. They implement several findings from psychotherapy process research to support clinical decision-making. As monitoring symptom change alone does not provide any information on how to adjust the treatment strategy, the goal of such clinical problem solving tools is to alert therapists to potential obstacles to treatment progress and to provide suggestions for interventions to improve treatment for patients at risk of treatment failure (Boswell et al. 2015). A number of CST models have been developed and applied in the field of psychotherapy.

One such model is Lambert's assessment of signal cases (e.g., Lambert 2017), which guides therapists to routinely monitor four domains using a validated patient-reported questionnaire covering: therapeutic alliance, social support, motivation, and negative life events. If a patient is not on track according to feedback from routine outcome measures, the therapist is prompted to assess problems across one or more of these domains and to apply clinical skills that target the perspective problem (e.g., rupture-repair for alliance deficits, or a decisional balance exercise for motivational deficits). A meta-analysis of controlled trials that supplemented progress feedback with this clinical problem-solving approach concluded that the addition of CST significantly improves clinical outcomes (Shimokawa et al. 2010). Subsequent secondary analyses of data from these trials indicated that not on track signals were significantly associated with elevated scores on the domains of social support and adverse life events (Probst et al. 2020, White et al. 2015).

A unidimensional approach to CST is featured in the Partners for Change Outcome Management System (PCOMS; e.g., Miller et al. 2005), where clinicians routinely monitor a measure of the therapeutic alliance (SRS) in addition to treatment outcomes (ORS). This is based on the well-known association between alliance and treatment outcomes (e.g., see Flückiger et al. 2020), and follows the hypothesis that poor progress is likely to be related to alliance problems and potentially rectified by close attention to the alliance. However, a meta-analysis of clinical trials applying the PCOMS system found that subgroup analyses excluding studies that did not use the SRS (i.e., did not systematically monitor the alliance) did not influence the overall meta-analytic finding that simple outcome monitoring (using the ORS) improves treatment outcomes (Østergård et al. 2020).

Another multi-domain model (Lutz et al. 2019) involves the routine monitoring of five domains: (1) risk/suicidality; (2) motivation/therapy goals; (3) therapeutic alliance; (4) social support and critical life events; (5) emotion regulation/self-regulation. Each domain is assessed at every fifth session using a battery of validated patient-reported questionnaires. Like in previous CST systems, this model primes the therapist to assess potential problems across these domains in cases where clinical outcomes are classed as not on track using computerized feedback. In an analysis of archival data from over 400 patients who were monitored using this battery of measures, Schilling et al. (2021) found that patients classed as not on track had significantly elevated scores on the domains of risk/suicidality, social support, and life events.

Summary

The collection and systematic investigation of outcome measures in routine care has generated several new options to make use of continuous data assessments in clinical practice. Research findings allow the generation of clinically meaningful decision-rules for the selection of treatment options as well as the monitoring of treatment and the early detection of negative developments. In particular, the development of clinical support tools to supplement outcome monitoring seems to constitute a step forward to improve psychological therapy for those patients who are at greatest risk of poor treatment response.

IMPLEMENTING MEASUREMENT-BASED AND DATA INFORMED PSYCHOLOGICAL THERAPY

Continuous assessment of change during treatment enables clinicians to adapt their decisions about the best available treatment option, but also allows the immediate application of research findings into clinical practice. A particular strength of data-rich health research and medicine is the ability to develop individualized diagnosis and treatment options. This has been also demonstrated in other public health areas, ranging for example from tailor-made immune cells and tumor therapies to the search for genes that can influence the risk of heart disease, the development of life-threatening pulmonary hypertension or therapy options for Parkinson's disease (e.g., Han et al. 2020, Stoker & Barker 2020).

Feedback-Informed Treatment

Important developments such as brief psychometrically validated self-report measures, their repeated administration in routine care, and the advancement of methods such ETR models discussed above, have enabled the development of a new paradigm of measurement-based care. This approach makes best use of psychometric and statistical methods to support the delivery of effective care, in a way that enables therapists to make timely decisions that are supported by evidence, rather than clinical judgment alone. Furthermore, since most commonly used measurement-based care approaches use patient-reported measures (of symptoms, alliance, processes, goals, etc.), this approach places the patient’s perspective at the centre of the therapeutic process.

Feedback-informed treatment (FIT) is one of the most well established examples of MBC; also known as outcome feedback or patient-focused feedback research (Howard et al.

1996, Lambert et al. 2001). This involves regularly monitoring a patient's response to treatment by comparing their response to predicted trajectories of improvement (clinical norms). These clinical norms are derived from ETR models (e.g., Finch et al. 2001) or other advanced statistical approaches such as nearest neighbors analysis (Lutz et al. 2006) or dynamic prediction models that are recalibrated from session-to-session (Bone et al. 2021b). Typically, patients are asked to complete questionnaires about their symptoms before every therapy session, and the results are entered into a computerized FIT system. The system compares each patient's symptoms to those observed in hundreds of similar patients, and then provides a prognosis: either the patient is "on track" (likely to recover) or "not on track" (unlikely to recover). This feedback is provided to therapists every week, which prompts them to quickly identify and resolve obstacles to improvement in cases classified as "not on track".

There are now over 50 clinical trials and several meta-analyses indicating that FIT systems can help to improve treatment outcomes, to prevent dropout and to improve the efficiency of psychological treatment (for comprehensive narrative and meta-analytic reviews, see de Jong et al. 2021, Lambert et al. 2018). In particular, the evidence is stronger for studies that offered evidence-based psychological treatments for depression and/or anxiety problems, supplemented by a FIT system, and it appears to generalise to different countries and healthcare systems. Furthermore, supplementing FIT technology with clinical support tools (CST) described earlier has been shown to enhance the effects of FIT (Lambert et al. 2018).

Moderator analyses of feedback-informed treatment trials reveal that the effects of feedback are enhanced by using ETR-based models to track treatment response and supplementing these systems with training for therapists and with clinical support tools (de Jong et al. 2021). Providing feedback to therapists and patients, rather than only to therapists, also seems to enhance its effectiveness (de Jong et al. 2014). Importantly, the extent to which individual clinicians adhere to the use of feedback systems determines the extent to which measurement-based therapy will be effective (see Lutz et al. 2021a). As such, there is already compelling evidence demonstrating that feedback-informed treatment is more effective than usual psychological care (purely guided by clinical judgement) in terms of symptomatic improvement and dropout prevention (de Jong et al. 2021), but adequate adherence by therapists varies considerably even in clinical trials. Thus, the current challenges and research questions in this field concern how to optimize the implementation of feedback systems in routine practice. Numerous obstacles to implementation have been identified in prior studies; such as organizational, technological, practical and attitudinal barriers (for an overview see Lutz et al. 2021a). Several examples of successful implementation and strategies to optimize adoption by therapists have also been documented and we refer interested readers to a previously published collection of implementation-focused case studies (de Jong 2016).

Precision Mental Health Care

The debate about the development and implementation of "personalized" or "precision" medicine has also influenced psychotherapy (research) in recent years. Several new developments have emerged within this tradition of outcome prediction and monitoring, which can be summarized as precision mental health research. This branch of research is intertwined with the traditions and improvements in continuous outcome measurement (e.g., Delgadillo & Lutz 2020, DeRubeis et al. 2014, Huibers et al. 2015). The overarching goal is to use evidence-based strategies to support decision-making in clinical practice (e.g., Bickman 2020, Chekroud et al. 2021, Cohen et al. 2021, Page et al. 2019, Zilcha-Mano 2019).

One example of a comprehensive treatment selection and tracking system, which is also augmented by tools from e-mental health, is the Trier Treatment Navigator (TTN). This system includes a personalized treatment recommendation at the beginning as well as adjustments based on continuous outcome assessments during the course of treatment (Lutz et al. 2019, 2021b). For each patient, the TTN generates individual predictions based on a large, previously treated patient sample for several important indicators, e.g., dropout risk and the optimal treatment strategy to start the treatment (see Figure 2a). In order to predict the optimal treatment strategy, the most similar patients (based on the nearest neighbour algorithm) who have already been treated are identified for an individual patient out of a large archival patient sample (N = 1,234). These nearest neighbor groups of patients are generated for two treatment strategies or a combination of both (problem-solving-oriented, relationship- and motivation-oriented, or mixed approach in the first ten sessions) and effect sizes for each group achieved in the first ten sessions are generated. Effect sizes of the three treatment strategies and confidence intervals are then mapped and the therapist can directly evaluate whether a particular treatment strategy is predicted to have a clear advantage for this patient (the problem-solving approach for the patient example in Figure 2a).

After the start of treatment, the navigation and monitoring system includes personalized treatment adjustment based on a dynamic risk index and clinical support tools. The system alerts therapists to the risk of an adverse treatment outcome. Figure 2b shows the course of treatment (session by session) and the visual feedback for a sample patient. The feedback graph is supplemented by the 30 most similar neighbors' average course of treatment and the above described dynamic risk index. This risk index is also based on the expected treatment response of the most similar previously treated patients and is adaptively recalculated after each session. A patient progress score above the risk index indicates a significantly increased risk of a negative outcome. In such cases, therapists receive a warning signal as in the example in Figure 2b: "(!) General symptom course is not as expected". In such cases, therapists also receive feedback on the potential problem area(s) (Risk /

Suicidality, Motivation / Therapy Goals, Therapeutic Relationship, Social Support / Critical Life Events, and Emotion Regulation / Self-Regulation) in which the patient has indicated high scores. Furthermore, the system additionally supports the implementation of helpful alternative clinical interventions by a number of support tools (e.g., videos on how to perform specific techniques, worksheets, and audio files for download). In the example in Figure 2b, the patient shows high scores in the areas of motivation / therapy goals and emotion regulation / self-regulation) and the feedback system suggests the therapist click on the support tools in these two areas.

Figure 2 shows the sample treatment of a 35 year old, single male patient with major depressive disorder (single episode) and sub-threshold personality disorder. At the beginning of therapy, the patient showed great ambivalence toward treatment and only started because of his friends' and family's suggestions. He attributed all his problems to his environment and, for example, in session 7 he was getting extremely upset about a friend who had canceled a vacation. This event led him to express strong doubts whether therapy could help him at all. The navigation system proposed several techniques concerning doubts about treatment (clinical support tool motivation/therapy goals) and the therapist used these techniques to question the patient's perspective and attributions concerning the cause of his problems. Together, they developed alternative explanations for the origin of his problems and were able to strengthen his motivation for treatment by including additional motivational interviewing techniques.

In a randomized-controlled trial (RCT) with 538 patients, the TTN was prospectively evaluated. Patients showed an increased effect size of about .3, when therapists followed the recommended treatment strategy in the first ten sessions. Moreover, the linear mixed models revealed therapist symptom awareness and therapist attitude and confidence as significant predictors of outcome as well as therapist-rated usefulness of feedback as a significant moderator of the feedback-outcome and the not on track-outcome associations. These results demonstrate the importance of prospective studies and the necessity of a high quality implementation of digital decision support tools in clinical practice.

Implementation

From the beginning, the integration of psychometric measures into clinical practice has been associated with barriers and implementation issues. Historically, psychotherapists have often been hesitant or even critical of implementing measures on a routine basis (e.g., Boswell et al. 2015, de Jong 2016, Douglas et al. 2016). This phenomenon has often been described as the scientist-practitioner gap in the psychological therapies (e.g., Lilienfeld et al. 2015). This contrasts with the patient perspective, where outcome evaluation is usually well received. For example, when patients are asked whether they find it important to monitor the results of psychotherapeutic treatments, for example using questionnaires, more than 90% seem to agree or partially agree (e.g., Lutz et al. 2021a).

Furthermore, therapist's behavior and attitudes, perceived usefulness and commitment to outcome measurement have some impact on its effectiveness in clinical practice (e.g., de Jong et al. 2021, Lutz et al. 2021b). Several factors contribute to the hesitant reception of outcome measurement in clinical practice. For example, the technical equipment, financial support or necessary time might be not available in daily practice. Furthermore, therapists' general aversion to the use of technology in their practice has often been mentioned as a reason in the literature. In addition, clinicians might not trust the ecological validity of measures, thinking that empirical findings do not reflect their everyday practice (e.g., Boswell et al. 2015, Gilbody et al. 2002). Sometimes, measures can also be perceived as controlling or concerns arise with regard to data security (e.g., Mütze et al. 2021). All these factors culminate in outcome monitoring not yet often enough implemented in clinical practice and clinicians lack the training and support necessary to make good use of the information (e.g., Boswell et al. 2015). As described throughout this chapter, in the future the cultural shift to a measurement-based and data-informed psychological therapy including outcome measures as well as feedback and navigation systems is one of the most important steps to improve clinical services in mental health.

Summary

Measurement-based and data-informed psychological therapy evolved from efforts to extend the external validity of outcome research, which is traditionally based on efficacy research and RCTs. However, after a treatment effect has been established, observational, non-interventional designs and the monitoring of real world applications are necessary to investigate implementation issues. The main focus of these Phase IV surveillance investigations is the transportability of a treatment into real life conditions, side-effects, subgroup differences and treatment failures (e.g., Suvarna 2010). The implementation of outcome monitoring supplemented by feedback and navigation systems takes this investigation a step further by not only monitoring real-world effects, but also improving effects in real time.

CONCLUSIONS

Common to all endeavours described in this paper is the idea of providing clinicians with personalized recommendations for their everyday clinical decision-making based on continuous data collection. The generation of large practice-based datasets allows the development of new data disaggregation strategies to more precisely define relevant reference groups for each patient, thereby improving the applicability of information to the individual patient and supporting therapists in their clinical practice with an expert decision support system.

So far, clinical judgement has been largely based on theory and intuition. However, we now have adequate assessment tools that are psychometrically reliable, brief, clinically useful, and sensitive to change (Lutz et al. 2021a). Furthermore, feedback and decision support tools are available, which have shown that, when used, lead to more evidence-based and effective treatments. Clinicians can make use of feedback to identify patients with a high likelihood of a negative treatment development, and to obtain empirically-supported recommendations about treatment strategies that could improve outcomes. In the future, further research efforts in this area should focus on implementation. Furthermore, the improvement of freely available and easy-to-apply measures remains a priority as well as efforts to replicate results under large-scale routine care conditions. Of course, further improvements are also necessary in terms of item overlap and standardization of measures (e.g., Fried 2017). Furthermore, important future research perspectives include (1) new statistical methods (e.g., machine learning) to analyze large cross-sectional as well as intensive longitudinal datasets; (2) improved research on processes and mechanisms of change; (3) better dissemination and cross-cultural adaptation (Kazdin 2018), and (4) better implementation and testing of clinical decision-support systems to identify and treat patients at risk for treatment failure.

The vision for the future is to support clinical work not only with one algorithm, but with multiple algorithms integrated into expert decision-support systems, which include algorithms and problem solving modules to address several difficulties encountered in clinical practice such as: how to personalize the choice of treatment modality or techniques for the individual; how to predict and prevent dropout, symptomatic deterioration or side effects; how to predict and prevent relapse. The aim is to support human decision-making, not to replace it with automatized interventions. Well-trained clinicians should receive empirically-based support to solve the complex problems that are commonly encountered in the field of psychological therapy.

SUMMARY POINTS

1. Measurement-based and data-informed psychological therapy uses algorithmic decision tools to overcome some of the limitations of clinical judgement and intuition. It is enabled by the accumulation of large-scale routine clinical datasets.
2. Studies using longitudinal data show that the course of psychotherapy is influenced by patients' baseline characteristics as well as early symptomatic changes, and it often follows a non-linear pattern. Furthermore, individual trajectories can deviate significantly from aggregated clinical population trends.
3. Intensive (i.e., daily) longitudinal assessments can improve the representation of high-frequency intrapersonal processes in patients' daily lives. However, due to their complexity and existing data analysis problems, less frequent session-by-session assessments are easier to implement and to use in clinical practice.
4. Data-informed decision tools can guide treatment selection, monitoring of treatment progress, early detection of not-on-track patients, and adaptive recommendation of clinical strategies.
5. The effectiveness of measurement-based psychological therapy depends on therapists' behaviors and attitudes toward outcome measurement and feedback, as well as their use of technology. Therefore, to realize the full potential of data-informed therapy, future generations of psychotherapists need to be trained to understand psychometrics and to embrace technology as a means of making smarter and more effective decisions.

ACKNOWLEDGEMENT

The authors are most thankful for the tremendously helpful proofreading by Kaitlyn Poster (University of Trier) and Paulina Gonzalez Salas Duhne (University of Sheffield). We also thank Björn Bennemann and Jana Wasserheß (University of Trier) for their helpful comments, supporting figure design, and compiling the reference list.

LITERATURE CITED

Aderka IM, Nickerson A, Bøe HJ, Hofmann SG. 2012. Sudden gains during psychological treatments of anxiety and depression: A meta-analysis. *J. Consult. Clin. Psych.* 80(1):93–101

Anderson T, McClintock AS, Himawan L, Song X, Patterson CL. 2016. A prospective study of therapist facilitative interpersonal skills as a predictor of treatment outcome. *J. Consult. Clin. Psych.* 84(1):57–66

Atzil-Slonim D, Juravski D, Bar-Kalifa E, Gilboa-Schechtman E, Tuval-Mashiach R, et al. 2021. Using topic models to identify clients' functioning levels and alliance ruptures in psychotherapy. *Psychotherapy.* https://doi.org/10.1037/pst0000362

Baldwin SA, Berkeljon A, Atkins DC, Olsen JA, Nielsen SL. 2009. Rates of change in naturalistic psychotherapy: Contrasting dose-effect and good-enough level models of change. *J. Consult. Clin. Psych.* 77(2):203–11

Baldwin SA, Imel ZE. 2013. Therapist effects: Findings and methods. In *Bergin and Garfield's handbook of psychotherapy and behavior change*, 6th ed., ed. MJ Lambert, pp. 258–297. New York: Wiley.

Baldwin SA, Imel ZE. 2020. Studying specificity in psychotherapy with meta-analysis is hard. *Psychother. Res.* 30(3):294–96

Barkham M, Connell J, Stiles WB, Miles JNV, Margison F, et al. 2006. Dose-effect relations and responsive regulation of treatment duration: The good enough level. *J. Consult. Clin. Psych.* 74(1):160–67

Barkham M, Lambert MJ. 2021. The efficacy and effectiveness of psychological therapies. In *Bergin and Garfield's handbook of psychotherapy and behavior change*, 7th ed., ed. M Barkham, W Lutz, LG Castonguay, pp. 129–86. New York: Wiley

Barkham M, Rees A, Stiles WB, Shapiro DA, Hardy GE, Reynolds S. 1996. Dose–effect relations in time-limited psychotherapy for depression. *J. Consult. Clin. Psych.* 64(5):927–35

Baur T, Clausen S, Heimerl A, Lingenfelser F, Lutz W, André E. 2020. NOVA: A tool for explanatory multimodal behavior analysis and its application to psychotherapy. In *Lecture Notes in Computer Science: Vol. 11962. MultiMedia modeling: 26th international conference, MMM 2020, Daejeon, South Korea, January 5-8, 2020: proceedings*, Vol. 11962, ed. W-H Cheng, pp. 577–88. Cham: Springer

Beard JIL, Delgadillo J. 2019. Early response to psychological therapy as a predictor of depression and anxiety treatment outcomes: A systematic review and meta-analysis. *Depress. Anxiety* 36(9):866–78

Bickman L. 2020. Improving mental health services: A 50-year journey from randomized experiments to artificial intelligence and precision mental health. *Adm. Policy Ment. Hlth* 47:795–843

Bone C, Delgadillo J, Barkham M. 2021a. A systematic review and meta-analysis of the good-enough level (GEL) literature. *J. Couns. Psychol.* 68(2):219–31

Bone C, Simmonds-Buckley M, Thwaites R, Sandford D, Merzhvynska M, et al. 2021b. Dynamic prediction of psychological treatment outcomes: Development and validation of a prediction model using routinely collected symptom data. *Lancet Digital Health* 3(4):e231-e240

Borsboom D. 2017. A network theory of mental disorders. *World Psychiatry* 16(1):5–13

Boswell JF, Kraus DR, Miller SD, Lambert MJ. 2015. Implementing routine outcome monitoring in clinical practice: Benefits, challenges, and solutions. *Psychother. Res.* 25(1):6–19

Bringmann LF. 2021. Person-specific networks in psychopathology: Past, present, and future. *Curr. Opin. Psychol.* 41:59–64

Bringmann LF, van der Veen DC, Wichers M, Riese H, Stulp G. 2020. Esmvis: A tool for visualizing individual Experience Sampling Method (ESM) data. *Qual. Life Res.* https://doi.org/10.1007/s11136-020-02701-4

Castonguay LG, Barkham M, Lutz W, McAleavey A. 2013. Practice-oriented research: Approaches and applications. In *Bergin and Garfield's handbook of psychotherapy and behavior change*, 6th ed., ed. MJ Lambert, pp. 85–133. New York: Wiley

Castonguay LG, Barkham M, Youn SJ, Page AC. 2021. Practice-based evidence: Findings from routine clinical settings. In *Bergin and Garfield's handbook of psychotherapy and behavior change,* 7th ed., ed. M Barkham, W Lutz, LG Castonguay, pp. 187–220. New York: Wiley.

Chang PGRY, Delgadillo J, Waller G. 2021. Early response to psychological treatment for eating disorders: A systematic review and meta-analysis. *Clin. Psychol. Rev.* 86(1):102032

Chekroud AM, Bondar J, Delgadillo J, Doherty G, Wasil A, et al. 2021. The promise of machine learning in predicting treatment outcomes in psychiatry. *World Psychiatry* 20(2):154–70

Clark DM. 2018. Realizing the mass public benefit of evidence-based psychological therapies: The IAPT program. *Annu. Rev. Clin. Psycho.* 14:159–83

Cohen ZD, Delgadillo J, DeRubeis RJ. 2021. Personalized treatment approaches. In *Bergin and Garfield's handbook of psychotherapy and behavior change*, 7th ed., ed. M Barkham, W Lutz, LG Castonguay, pp. 667–700. New York: Wiley.

Constantino MJ, Boswell JF, Coyne AE, Swales TP, Kraus DR. 2021. Effect of matching therapists to patients vs assignment as usual on adult psychotherapy outcomes: A randomized clinical trial. *JAMA Psychiat.* https://doi.org/10.1001/jamapsychiatry.2021.1221

Contreras A, Nieto I, Valiente C, Espinosa R, Vazquez C. 2019. The study of psychopathology from the network analysis perspective: A systematic review. *Psychother. Psychosom.* 88(2):71–83

Cristea IA, Vecchi T, Cuijpers P. 2021. Top-down and bottom-up pathways to developing psychological interventions. *JAMA Psychiat.* 78(6):593–94

Crits-Christoph P, Connolly Gibbons MB. 2021. Psychotherapy process-outcome research: Advances in understanding causal connections. In *Bergin and Garfield's handbook of psychotherapy and behavior change*, 7th ed., ed. M Barkham, W Lutz, LG Castonguay, pp. 259–92. New York: Wiley.

Crits-Christoph P, Ring-Kurtz S, Hamilton JL, Lambert MJ, Gallop R, et al. 2012. A preliminary study of the effects of individual patient-level feedback in outpatient substance abuse treatment programs. *J. Subst. Abuse Treat.* 42(3):301–9

Cuijpers P, Reijnders M, Huibers MJH. 2019. The role of common factors in psychotherapy outcomes. *Annu. Rev. Clin. Psycho.* 15:207–31

Dakos V, Carpenter SR, Brock WA, Ellison AM, Guttal V, et al. (2012). Methods for detecting early warnings of critical transitions in time series illustrated using simulated ecological data. *PLOS ONE* 7(7):e41010

de Jong K. 2016. Challenges in the implementation of measurement feedback systems. *Adm. Policy Ment. Hlth* 43(3):467–70

de Jong K, Conijn JM, Gallagher RAV, Reshetnikova AS, Heij M, Lutz MC. 2021. Using progress feedback to improve outcomes and reduce drop-out, treatment duration, and deterioration: A multilevel meta-analysis. *Clin. Psychol. Rev.* 85:102002 de Jong K, Timman R, Hakkaart-Van Roijen L, Vermeulen P, Kooiman K, et al. 2014. The effect of outcome monitoring feedback to clinicians and patients in short and long-term psychotherapy: A randomized controlled trial. *Psychother. Res.* 24(6):629–39

Deisenhofer A-K, Delgadillo J, Rubel JA, Böhnke JR, Zimmermann D, et al. 2018. Individual treatment selection for patients with posttraumatic stress disorder. *Depress. Anxiety* 35(6):541–50

Deisenhofer A-K, Rubel JA, Bennemann B, Aderka IM, Lutz W. 2021. Are some therapists better at facilitating and consolidating sudden gains than others?. *Psychother. Res*. https://doi.org/10.1080/10503307.2021.1921302

Delgadillo J, Ali S, Fleck K, Agnew C, Southgate A, et al. 2021. Stratcare: A pragmatic, multi-site, single-blind, cluster randomised controlled trial of stratified care for depression. Submitted.

Delgadillo J, Gonzalez Salas Duhne P. 2020. Targeted prescription of cognitive-behavioral therapy versus person-centered counseling for depression using a machine learning approach. *J. Consult. Clin. Psych*. 88(1):14–24

Delgadillo J, Lutz W. 2020. A development pathway towards precision mental health care. *JAMA Psychiat*. *77*(9):889–90

Delgadillo J, McMillan D, Lucock M, Leach C, Ali S, Gilbody S. 2014. Early changes, attrition, and dose-response in low intensity psychological interventions. *Brit. J. Clin. Psychol.* 53(1):114–30

Delgadillo J, Rubel J, Barkham M. 2020. Towards personalized allocation of patients to therapists. *J. Consult. Clin. Psych*. 88(9):799–808

DeRubeis RJ, Cohen ZD, Forand NR, Fournier JC, Gelfand LA, Lorenzo-Luaces L. 2014. The personalized advantage index: Translating research on prediction into individualized treatment recommendations. A demonstration. *PLOS ONE* 9(1):e83875

Douglas S, Button S, Casey SE. 2016. Implementing for sustainability: Promoting use of a measurement feedback system for innovation and quality improvement. *Adm. Policy Ment. Hlth* 43(3):286–91

Ebner-Priemer UW, Trull TJ. 2009. Ecological momentary assessment of mood disorders and mood dysregulation. *Psychol. Assessment* 21(4):463–75

Eisele G, Vachon H, Lafit G, Kuppens P, Houben M, et al. 2020. The effects of sampling frequency and questionnaire length on perceived burden, compliance, and careless responding in experience sampling data in a student population. *Assessment.* https://doi.org/10.1177/1073191120957102

Epskamp S. 2020. Psychometric network models from time-series and panel data. *Psychometrika* 85(1):206–31

Evans DL, Herbert JD, Nelson-Gray RO, Gaudiano BA. 2002. Determinants of diagnostic prototypicality judgments of the personality disorders. *J. Pers. Disord.* 16(1):95–106

Falkenström F, Ekeblad A, Holmqvist R. 2016. Improvement of the working alliance in one treatment session predicts improvement of depressive symptoms by the next session. *J. Consult. Clin. Psych*. 84(8):738–51

Finch AE, Lambert MJ, Schaalje BG. 2001. Psychotherapy quality control: The statistical generation of expected recovery curves for integration into an early warning system. *Clin. Psychol. Psychot.* 8(4):231–42

Fisher AJ, Bosley HG, Fernandez KC, Reeves JW, Soyster PD, et al. 2019. Open trial of a personalized modular treatment for mood and anxiety. *Behav. Res. Ther.* 116:69–79

Fisher AJ, Soyster P, Ashlock L. 2021. Machine learning algorithms for generating early warning signals in real time. *Biol. Psychiat.* 89(9):S58–S59

Flood N, Page A, Hooke G. 2019. A comparison between the clinical significance and growth mixture modelling early change methods at predicting negative outcomes. *Psychother. Res.* 29(7):947–58

Flückiger C, Rubel J, Del Re AC, Horvath AO, Wampold BE, et al. 2020. The reciprocal relationship between alliance and early treatment symptoms: A two-stage individual participant data meta-analysis. *J. Consult. Clin. Psych*. 88(9):829–43

Fried EI. 2017. The 52 symptoms of major depression: Lack of content overlap among seven common depression scales. *J. Affect. Disorders* 208:191–97

Frumkin MR, Piccirillo ML, Beck ED, Grossman JT, Rodebaugh TL. 2021. Feasibility and utility of idiographic models in the clinic: A pilot study. *Psychother. Res.* 31(4):520–34

Garb HN. 2005. Clinical judgment and decision making. *Annu. Rev. Clin. Psycho.* 1:67–89

Gilbody SM, House AO, Sheldon TA. 2002. Psychiatrists in the UK do not use outcomes measures: National survey. *Brit. J. Psychiat.* 180(2):101–3

Goldfried MR. 2019. Obtaining consensus in psychotherapy: What holds us back? *Am. Psychol.* 74(4):484–96

Gómez Penedo JM, Schwartz B, Giesemann J, Rubel JA, Deisenhofer A-K, Lutz W. 2021. For whom should psychotherapy focus on problem coping? A machine learning algorithm for treatment personalization. *Psychother. Res*. https://doi.org/10.1080/10503307.2021.1930242

Grove WM, Meehl PE. 1996. Comparative efficiency of informal (subjective, impressionistic) and formal (mechanical, algorithmic) prediction procedures: The clinical–statistical controversy. *Psychol. Public Pol. L.* 2(2):293–323

Hamaker EL, Wichers M. 2017. No time like the present: Discovering the hidden dynamics in intensive longitudinal data. *Curr. Dir. Psychol. Sci.* 26(1):10–15

Hammond KR. 1978. Toward increasing competence of thought in public policy formation. In *Judgement and decision in public policy formation,* ed. KR Hammond, pp. 11–32. New York: Routledge

Han S, Shuen WH, Wang W-W, Nazim E, Toh HC. 2020. Tailoring precision immunotherapy: Coming to a clinic soon?. *ESMO Open* 5 Suppl 1:e000631

Hannan C, Lambert MJ, Harmon C, Nielsen SL, Smart DW, et al. 2005. A lab test and algorithms for identifying clients at risk for treatment failure. *J. Clin. Psychol.* 61(2):155–63

Hehlmann MI, Schwartz B, Lutz T, Gómez Penedo JM, Rubel JA, Lutz W. 2021. The use of digitally assessed stress levels to model change processes in CBT - a feasibility study on seven case examples. *Front. Psychiatry* 12:613085

Heinonen E, Nissen-Lie HA. 2020. The professional and personal characteristics of effective psychotherapists: A systematic review. *Psychother. Res.* 30(4):417–32

Heinzel S, Tominschek I, Schiepek G. 2014. Dynamic patterns in psychotherapy-discontinuous changes and critical instabilities during the treatment of obsessive compulsive disorder. *Nonlinear Dyn. Psychol. Life Sci* 18(2):155–76

Hofmann SG, Hayes SC. 2019. The future of intervention science: Process-Based Therapy. *Clin. Psychol. Sci.* 7(1):37–50

Howard KI, Kopta SM, Krause MS, Orlinsky DE. 1986. The dose–effect relationship in psychotherapy. *Am. Psychol.* 41(2):159–64

Howard KI, Moras K, Brill PL, Martinovich Z, Lutz W. 1996. Evaluation of psychotherapy: Efficacy, effectiveness, and patient progress. *Am. Psychol.* 51(10):1059–64

Huibers MJH, Cohen ZD, Lemmens LH, Arntz A, Peeters FP, et al. 2015. Predicting optimal outcomes in cognitive therapy or interpersonal psychotherapy for depressed individuals using the personalized advantage index approach. *PLOS ONE* 10(11):e0140771

Husen K, Rafaeli E, Rubel JA, Bar-Kalifa E, Lutz W. 2016. Daily affect dynamics predict early response in CBT: Feasibility and predictive validity of EMA for outpatient psychotherapy. *J. Affect. Disorders* 206:305–14

Imel ZE, Barco JS, Brown HJ, Baucom BR, Baer JS, et al. 2014. The association of therapist empathy and synchrony in vocally encoded arousal. *J. Couns. Psychol.* 61(1):146–53

Jacobson NS, Truax P. 1991. Clinical significance: A statistical approach to defining meaningful change in psychotherapy research. . *J. Consult. Clin. Psych*. 59(1):12–19

Jacobson NC, Weingarden H, Wilhelm S. 2019. Digital biomarkers of mood disorders and symptom change. *NPJ Digit. Med.* 2:3

Kahneman D, Klein G. 2009. Conditions for intuitive expertise: A failure to disagree. *Am. Psychol.* 64(6):515–26

Kaiser T, Laireiter A-R. 2018. Process-symptom-bridges in psychotherapy: An idiographic network approach. *J. Pers.-Oriented Res.* 4(2):49–62

Kazdin AE. 2018. *Innovations in psychosocial interventions and their delivery: Leveraging cutting-edge science to improve the world's mental health*. New York: Oxford University Press.

Kraemer HC, Stice E, Kazdin A, Offord D, Kupfer D. 2001. How do risk factors work together? Mediators, moderators, and independent, overlapping, and proxy risk factors. *Am. J. Psychiat.* 158(6):848–56

Krause MS. 2018. Associational versus correlational research study design and data analysis. *Qual. Quant.* 52(6):2691–707

Krüger A, Ehring T, Priebe K, Dyer AS, Steil R, Bohus M. 2014. Sudden losses and sudden gains during a DBT-PTSD treatment for posttraumatic stress disorder following childhood sexual abuse. *Eur. J. Psychotraumato.* 5(1):24470

Kyron MJ, Hooke GR, Page AC. 2019. Assessing interpersonal and mood factors to predict trajectories of suicidal ideation within an inpatient setting. *J. Affect. Disorders* 252:315–24

Lambert MJ. 2017. Maximizing psychotherapy outcome beyond evidence-based medicine. *Psychother. Psychosom.* 86(2):80–89

Lambert MJ, Hansen NB, Finch AE. 2001. Patient-focused research: Using patient outcome data to enhance treatment effects. *J. Consult. Clin. Psych.* 69(2):159–72

Lambert MJ, Whipple JL, Bishop MJ, Vermeersch DA, Gray GV, Finch AE. 2002. Comparison of empirically-derived and rationally-derived methods for identifying patients at risk for treatment failure. *Clin. Psychol. Psychot.* 9(3):149–64

Lambert MJ, Whipple JL, Kleinstäuber M. 2018. Collecting and delivering progress feedback: A meta-analysis of routine outcome monitoring. *Psychotherapy* 55(4):520–37

Lenton TM, Held H, Kriegler E, Hall JW, Lucht W, et al. 2008. Tipping elements in the earth's climate system. *P. Natl. Acad. Sci. USA* 105(6):1786–93

Lilienfeld SO, Ritschel LA, Lynn SJ, Cautin RL, Latzman RD. 2015. Science-practice gap. In *The encyclopedia of clinical psychology*, ed. RL Cautin, SO Lilienfeld, pp. 2548–55. Hoboken: Wiley.

Lorenzo-Luaces L, DeRubeis RJ. 2018. Miles to go before we sleep: Advancing the understanding of psychotherapy by modeling complex processes. *Cognitive Ther. Res.* 42(2):212–17

Lorenzo-Luaces L, DeRubeis RJ, 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. *J. Affect. Disorders* 213:78–85

Lutz W, Deisenhofer A-K, Rubel J, Bennemann B, Giesemann J, et al. 2021b. Prospective evaluation of a clinical decision support system in psychological therapy. *J. Consult. Clin. Psych*. https://doi.org/10.1037/ccp0000642

Lutz W, Ehrlich T, Rubel J, Hallwachs N, Röttger M-A, et al. 2013. The ups and downs of psychotherapy: Sudden gains and sudden losses identified with session reports. *Psychother. Res.* 23(1):14–24

Lutz W, Hofmann SG, Rubel J, Boswell JF, Shear MK, et al. 2014. Patterns of early change and their relationship to outcome and early treatment termination in patients with panic disorder. *J. Consult. Clin. Psych*. 82(2):287–97

Lutz W, de Jong K, Rubel JA, Delgadillo J. 2021a. Measuring, predicting and tracking change in psychotherapy. In *Bergin and Garfield's handbook of psychotherapy and behavior change*, 7th ed., ed. M Barkham, W Lutz, LG Castonguay, pp. 89–133. New York: Wiley

Lutz W, Leach C, Barkham M, Lucock M, Stiles WB, et al. 2005. Predicting change for individual psychotherapy clients on the basis of their nearest neighbors. *J. Consult. Clin. Psych*. 73(5):904–13

Lutz W, Martinovich Z, Howard KI. 1999. Patient profiling: An application of random coefficient regression models to depicting the response of a patient to outpatient psychotherapy. *J. Consult. Clin. Psych*. 67(4):571–77

Lutz W, Rubel JA, Schwartz B, Schilling V, Deisenhofer A-K. 2019. Towards integrating personalized feedback research into clinical practice: Development of the Trier Treatment Navigator (TTN). *Behav. Res. Ther.* 120:103438

Lutz W, Saunders SM, Leon SC, Martinovich Z, Kosfelder J, et al. 2006. Empirically and clinically useful decision making in psychotherapy: Differential predictions with treatment response models. *Psychol. Assessment* 18(2):133–41

Lutz W, Schwartz B. 2021. Trans-theoretical clinical models and the implementation of precision mental health care. *World Psychiatry* 20(3). https://doi.org/10.1002/wps.20888

Lutz W, Schwartz B, Hofmann SG, Fisher AJ, Husen K, Rubel JA. 2018. Using network analysis for the prediction of treatment dropout in patients with mood and anxiety disorders: A methodological proof-of-concept study. *Sci. Rep.* 8(1):7819

May RM, Levin SA, Sugihara G. 2008. Complex systems: Ecology for bankers. *Nature* 451(7181):893–95

McAleavey AA, Nordberg SS, Moltu C. 2021. Initial quantitative development of the Norse Feedback system: A novel clinical feedback system for routine mental healthcare. *Qual. Life Res.* https://doi.org/10.1007/s11136-021-02825-1

McClintock AS, Perlman MR, McCarrick SM, Anderson T, Himawan L. 2017. Enhancing psychotherapy process with common factors feedback: A randomized, clinical trial. *J. Couns. Psychol.* 64(3):247–60

McKay D, Jensen-Doss A. 2021. Harmful treatments in psychotherapy. *Clin. Psychol.–Sci. Pr.* 28(1):2–4

Meehl PE. 1973. Why I do not attend case conferences. In *Psychodiagnosis: Selected papers*, ed. PE Meehl, pp. 225–302. Minneapolis: University of Minnesota Press

Meisel C, El Atrache R, Jackson M, Schubach S, Ufongene C, Loddenkemper T. 2020. Machine learning from wristband sensor data for wearable, noninvasive seizure forecasting. *Epilepsia* 61(12):2653–66

Miller G. 2012. The smartphone psychology manifesto. *Perspect. Psychol. Sci.* 7(3):221–37

Miller PR, Dasher R, Collins R, Griffiths P, Brown F. 2001. Inpatient diagnostic assessments: 1. Accuracy of structured vs. unstructured interviews. *Psychiat. Res.* 105(3):255–64

Miller SD, Duncan BL, Sorrell R, Brown GS. 2005. The partners for change outcome management system. *J. Clin. Psychol.* 61(2):199–208

Moggia D, Lutz W, Arndt A, Feixas G. 2020. Patterns of change and their relationship to outcome and follow-up in group and individual psychotherapy for depression. . *J. Consult. Clin. Psych*. 88(8):757–73

Mütze K, Witthöft M, Lutz W, Bräscher A-K. 2021. Matching research and practice: Prediction of individual patient progress and dropout risk for basic routine outcome monitoring. *Psychother. Res.* 31(1). https://doi.org/10.1080/10503307.2021.1930244

Ng MY, Schleider JL, Horn RL, Weisz JR. 2021. Psychotherapy for children and adolescents: From efficacy to effectiveness, scaling, and personalizing. In *Bergin and Garfield's handbook of psychotherapy and behavior change*, 7th ed., ed. M Barkham, W Lutz, LG Castonguay, pp. 621–666. New York: Wiley.

Nordmo M, Monsen JT, Høglend PA, Solbakken OA. 2020. Investigating the dose-response effect in open-ended psychotherapy. *Psychother. Res.* https://doi.org/10.1080/10503307.2020.1861359

Olthof M, Hasselman F, Strunk G, van Rooij M, Aas B, et al. (2020). Critical fluctuations as an early-warning signal for sudden gains and losses in patients receiving psychotherapy for mood disorders. *Clin. Psychol. Sci.* 8(1):25–35

Østergård OK, Randa H, Hougaard E. 2020. The effect of using the Partners for Change Outcome Management System as feedback tool in psychotherapy-a systematic review and meta-analysis. *Psychother. Res.* 30(2):195–212

Owen J, Adelson J, Budge S, Wampold B, Kopta M, et al. 2015. Trajectories of change in psychotherapy. *J. Clin. Psychol.* 71(9):817–27

Page AC, Camacho KS, Page JT. 2019. Delivering cognitive behaviour therapy informed by a contemporary framework of psychotherapy treatment selection and adaptation. *Psychother. Res.* 29(8):971–73

Paz A, Rafaeli E, Bar-Kalifa E, Gilboa-Schectman E, Gannot S, et al. 2021. Intrapersonal and interpersonal vocal affect dynamics during psychotherapy. *J. Consult. Clin. Psych*. 89(3):227–39

Probst T, Kleinstäuber M, Lambert MJ, Tritt K, Pieh C, et al. 2020. Why are some cases not on track? An item analysis of the assessment for signal cases during inpatient psychotherapy. *Clin. Psychol. Psychot.* 27(4):559–66

Ramseyer F, Tschacher W. 2011. Nonverbal synchrony in psychotherapy: Coordinated body movement reflects relationship quality and outcome. *J. Consult. Clin. Psych*. 79(3):284–95

Raudenbush SW, Bryk AS. 2002. *Hierarchical linear models: Applications and data analysis methods*. Thousand Oaks: Sage

Robinson L, Delgadillo J, Kellett S. 2020a. The dose-response effect in routinely delivered psychological therapies: A systematic review. *Psychother. Res.* 30(1):79–96

Robinson L, Kellett S, Delgadillo J. 2020. Dose-response patterns in low and high intensity cognitive behavioral therapy for common mental health problems. *Depress. Anxiety* 37(3):285–94

Rubel JA, Fisher AJ, Husen K, Lutz W. 2018. Translating person-specific network models into personalized treatments: Development and demonstration of the dynamic assessment treatment algorithm for individual networks (DATA-IN). *Psychother. Psychosom.* 87(4):249–51

Rubel JA, Rosenbaum D, Lutz W. 2017. Patients' in-session experiences and symptom change: Session-to-session effects on a within- and between-patient level. *Behav. Res. Ther.* 90:58–66

Rubel JA, Zilcha-Mano S, Giesemann J, Prinz J, Lutz W. 2020. Predicting personalized process-outcome associations in psychotherapy using machine learning approaches—A demonstration. *Psychother. Res.* 30(3):300–9

Saunders R, Cape J, Fearon P, Pilling S. 2016. Predicting treatment outcome in psychological treatment services by identifying latent profiles of patients. *J. Affect. Disorders* 197:107–15

Saxon D, Barkham M. 2012. Patterns of therapist variability: Therapist effects and the contribution of patient severity and risk. *J. Consult. Clin. Psych*. 80(4):535–46

Scheffer M, Bascompte J, Brock WA, Brovkin V, Carpenter SR, et al. 2009. Early-warning signals for critical transitions. *Nature* 461(7260):53–59

Schilling VNLS, Zimmermann D, Rubel JA, Boyle KS, Lutz W. 2021. Why do patients go off track? Examining potential influencing factors for being at risk of psychotherapy treatment failure. *Qual. Life Res.* https://doi.org/10.1007/s11136-020-02664-6

Schwartz B, Cohen ZD, Rubel JA, Zimmermann D, Wittmann WW, Lutz W. 2021. Personalized treatment selection in routine care: Integrating machine learning and statistical algorithms to recommend cognitive behavioral or psychodynamic therapy. *Psychother. Res.* 31(1):33–51

Shimokawa K, Lambert MJ, Smart DW. 2010. Enhancing treatment outcome of patients at risk of treatment failure: Meta-analytic and mega-analytic review of a psychotherapy quality assurance system. *J. Consult. Clin. Psych*. 78(3):298–311

Stiles WB, Honos-Webb L, Surko M. 1998. Responsiveness in psychotherapy. *Clin. Psychol.–Sci. Pr.* 5(4):439–58

Stoker TB, Barker RA. 2020. Recent developments in the treatment of Parkinson’s Disease. *F1000Research* 9:F1000 Faculty Rev-862

Strunk DR, Cooper AA, Ryan ET, DeRubeis RJ, Hollon SD. 2012. The process of change in cognitive therapy for depression when combined with antidepressant medication: Predictors of early intersession symptom gains. *J. Consult. Clin. Psych*. 80(5):730–38

Stulz N, Lutz W, Kopta SM, Minami T, Saunders SM. 2013. Dose-effect relationship in routine outpatient psychotherapy: Does treatment duration matter?. *J. Couns. Psychol.* 60(4):593–600

Suvarna V. 2010. Phase IV of drug development. *Perspect. Clin. Res.* 1(2):57–60

Tang TZ, DeRubeis RJ. 1999. Sudden gains and critical sessions in cognitive-behavioral therapy for depression. *J. Consult. Clin. Psych*. 67(6):894–904

Tversky A, Kahneman D. 1974. Judgment under uncertainty: Heuristics and biases. *Science* 185(4157):1124–31

van Borkulo C, Boschloo L, Borsboom D, Penninx BWJH, Waldorp LJ, Schoevers RA. 2015. Association of symptom network structure with the course of corrected depression. *JAMA Psychiat.* 72(12):1219–26

van Bronswijk SC, Bruijniks SJE, Lorenzo-Luaces L, DeRubeis RJ, Lemmens LHJM, et al. 2021. Cross-trial prediction in psychotherapy: External validation of the Personalized Advantage Index using machine learning in two Dutch randomized trials comparing CBT versus IPT for depression. *Psychother. Res.* 31(1):78–91

Walfish S, McAlister B, O'Donnell P, Lambert MJ. 2012. An investigation of self-assessment bias in mental health providers. *Psychol. Rep.* 110(2):639–44

Webb CA, Forgeard M, Israel ES, Lovell-Smith N, Beard C, Björgvinsson T. 2021. Personalized prescriptions of therapeutic skills from patient characteristics: An ecological momentary assessment approach. *J. Consult. Clin. Psych*. https://doi.org/10.1037/ccp0000555

Webb CA, Trivedi MH, Cohen ZD, Dillon DG, Fournier JC, et al. 2019. Personalized prediction of antidepressant v. placebo response: Evidence from the EMBARC study. *Psychol. Med.* 49(7):1118–27

White MM, Lambert MJ, Ogles BM, Mclaughlin SB, Bailey RJ, Tingey KM. 2015. Using the Assessment for Signal Clients as a feedback tool for reducing treatment failure. *Psychother. Res.* 25(6):724–34

Wichers M, Groot PC, Psychosystems, ESM Group, EWS Group. 2016. Critical slowing down as a personalized early warning signal for depression. *Psychother. Psychosom.* 85(2):114–16

Wichers M, Peeters F, Rutten BPF, Jacobs N, Derom C, et al. 2012. A time-lagged momentary assessment study on daily life physical activity and affect. *Health Psychol.* 31(2):135–44

Wright AGC, Woods WC. 2020. Personalized models of psychopathology. *Annu. Rev. Clin. Psycho.* 16:49–74

Zell E, Strickhouser JE, Sedikides C, Alicke MD. 2020. The better-than-average effect in comparative self-evaluation: A comprehensive review and meta-analysis. *Psychol. Bull.* 146(2):118–49

Zilcha-Mano S. 2017. Is the alliance really therapeutic? Revisiting this question in light of recent methodological advances. *Am. Psychol.* 72(4):311–25

Zilcha-Mano S. 2019. Major developments in methods addressing for whom psychotherapy may work and why. *Psychother. Res.* 29(6):693–708

FIGURE CAPTIONS

Figure 1

Matrix of assessment modes in measurement-based and data-informed psychological therapy.

Figure 2

Example of a patient-specific pre-therapy strategy and an adaptive personalized recommendation during treatment as they are displayed in the clinical decision support system. A: Patient scores at the beginning of each session; B: Expected recovery curve; C: Failure boundary; D: As soon as the patient’s score exceeds the failure boundary on the HSCL-11, the therapist receives this warning signal, which is defined in more detail in the clinical problem-solving tools (CPST) below; E: CPST are divided into five domains. The exclamation mark indicates the domains in which the patient has specific problems. The therapist is able to click on these icons to gain access to the activated tools. The check mark indicates that the patient has few or no problems in this area.

FIGURES

Figure 1

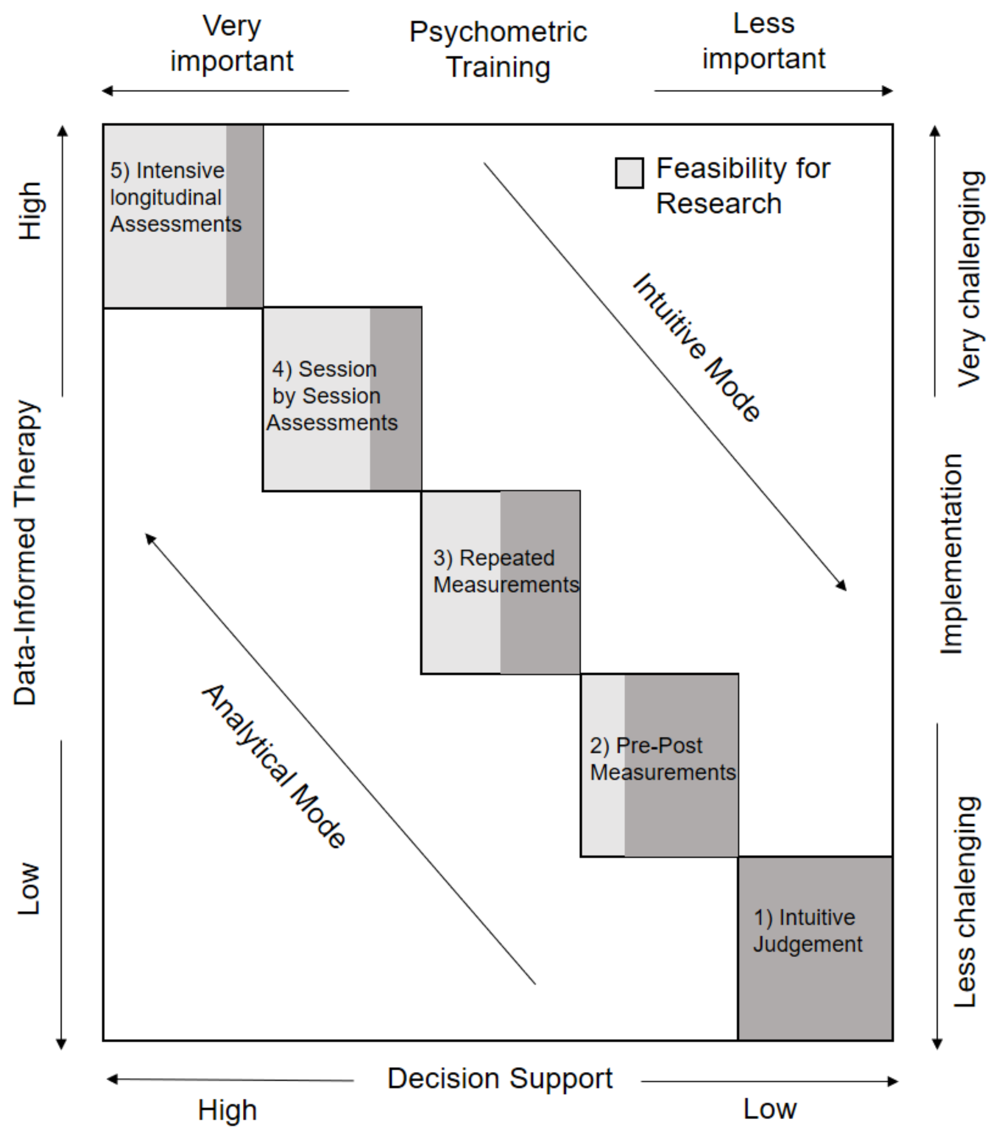
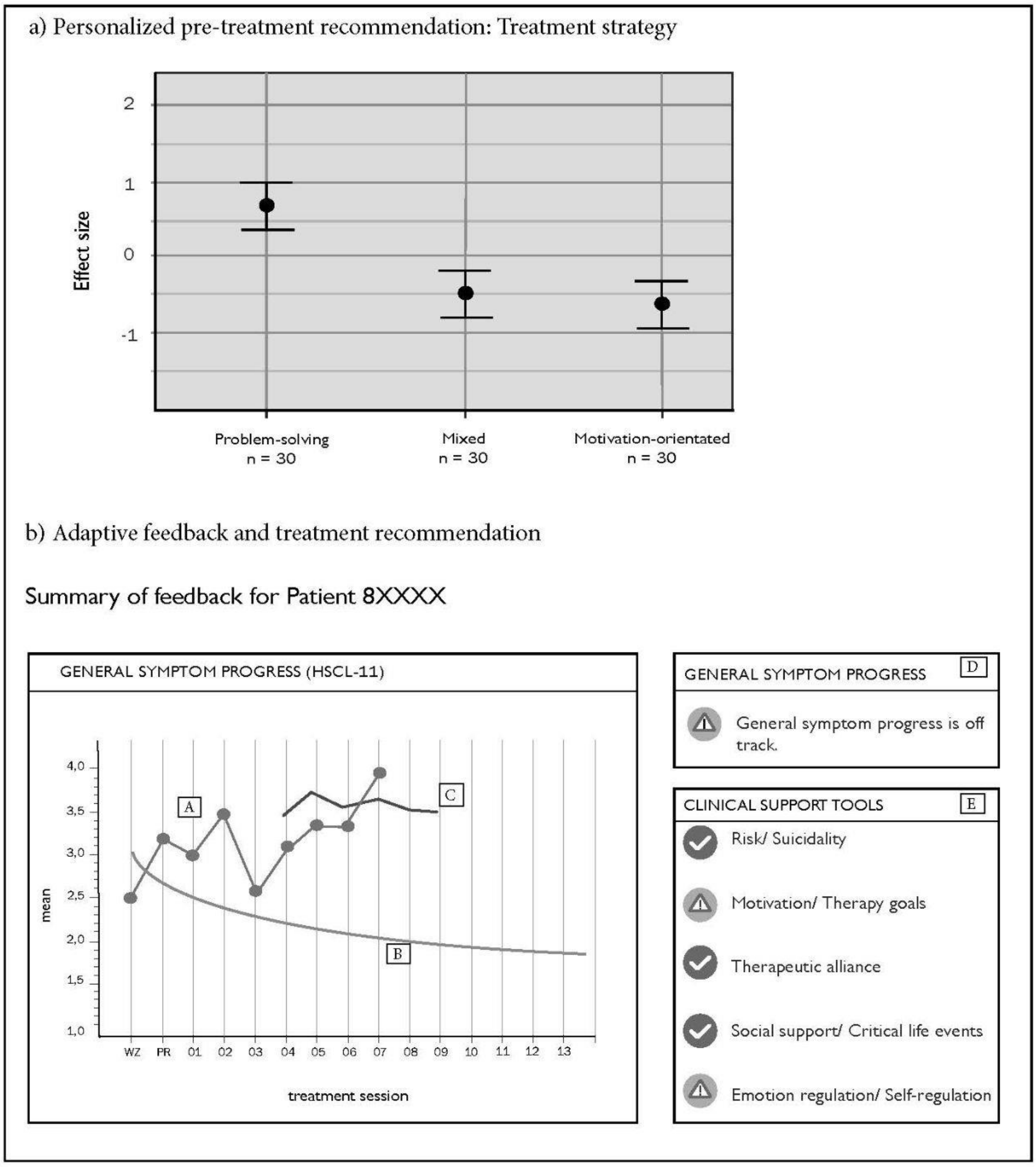


Figure 2


a) Personalized pre-treatment recommendation: Treatment strategy
Effect size
2
1
0
-1
Problem-solving
n = 30
Mixed
n = 30
Motivation-orientated
n = 30
b) Adaptive feedback and treatment recommendation
Summary of feedback for Patient 8XXXX
GENERAL SYMPTOM PROGRESS (HSCL-11)
mean
4,0
3,5
3,0
2,5
2,0
1,5
1,0
A
B
C
WZ PR 01 02 03 04 05 06 07 08 09 10 11 12 13
treatment session
GENERAL SYMPTOM PROGRESS
D
General symptom progress is off track.
CLINICAL SUPPORT TOOLS
E
Risk/ Suicidality
Motivation/ Therapy goals
Therapeutic alliance
Social support/ Critical life events
Emotion regulation/ Self-regulation