Routine outcome monitoring

Barkham, M.; Jong, K. de; Delgadillo, J.; Lutz, W.; Hill, ; C.E, ; ... ; J.C.,

15

Routine Outcome Monitoring

Michael Barkham, Kim de Jong, Jaime Delgadillo, and Wolfgang Lutz

In this chapter, we summarize the research evidence on conducting routine outcome monitoring (ROM) in individual adult psychological therapies. A previous account of this topic (Lambert et al., 2019) focused on two specific ROM systems: the OQ-45 measure and feedback system (Lambert et al., 2013) and the Partners in Change Outcome System (PCOMS; Duncan & Reese, 2015). By contrast, we adopt a structural perspective, focusing on the process of ROM, recognizing that multiple measures and constructs can lend themselves to this process. We summarize the research evidence and include an abridged version of the most comprehensive meta-analysis published on this topic, comprising 58 studies drawn from 49 randomized controlled trials (RCTs) and nine practice-based studies using a wide array of ROM measures and feedback systems (see De Jong et al., 2021). Prior to setting out the main sections of the chapter, we provide a brief background on the context of ROM.

Since the mid-1900s, it has been noted that statistical (i.e., actuarial) models are more accurate than clinical judgments (Meehl, 1954), and research has provided strong evidence to question the empirical accuracy and clinical effectiveness of traditional forms of clinical judgment and intuition (Ægisdóttir et al., 2006). Psychotherapists have been found to be overconfident in evaluating their own effectiveness, with one survey of U.S. practitioners in private practice reporting the belief that 85% of their patients benefitted from therapy (Walfish et al., 2012). However, the return rate of that survey was 26%, and respondents were more likely to be practitioners who thought they had good outcomes. A partial UK-based replication (with a 15% return rate) found broadly similar findings but not so pronounced (i.e., 65% of patients benefitting), suggesting some possible personality and/or cultural differences (Parker & Waller, 2015). This issue is especially critical when it potentially blinds a therapist to a patient's evolving worsening outcomes. Given this context, enhancing the validity of in-session therapist behaviors and client outcomes calls for a measurement-based approach to psychological therapies informed by data and using feedback as a central therapeutic process (Lutz, Schwartz, et al., 2022).

A specific subgroup of patients of concern are those who deteriorate. In an often-cited study, 40 therapists (20 experienced and 20 trainees) were asked to predict their patients who might deteriorate by the end of therapy, having been informed that the overall expected deterioration rate would be in the region of 8% (Hannan et al., 2005). Of the 550 clients (at a university counseling center), 40 actually deteriorated by the end of therapy, but only three were predicted by therapists as likely to deteriorate, and of these, only one was correct. By contrast, actuarial methods correctly identified 36 of the 40 (90%) clients who deteriorated. These collective findings broadly attest to biases

Michael Barkham, Kim de Jong, Jaime Delgadillo, and Wolfgang Lutz, *Routine Outcome Monitoring* In: *Psychotherapy Skills and Methods That Work*. Edited by: Clara E. Hill and John C. Norcross, Oxford University Press. © Clara E. Hill and John C. Norcross 2023.
DOI: 10.1093/oso/9780197611012.003.0015

in clinicians' judgments and support the adoption of data-informed approaches to psychotherapy using ROM.

Routine outcome measurement and feedback have their roots in the paradigm of practice-based evidence and specifically that of *patient-focused research* (Howard et al., 1996). Such an approach encompasses findings from early reviews in the literature (e.g., Bergin, 1963) and subsequent texts (e.g., Lambert, 2010) attesting to the fact that not all patients improve and a proportion of them deteriorate. In such cases, administering outcome measures—either continuously or at regular intervals as a means of checking progress and informing treatment decision-making processes—can yield an *additive* effect over and above the effects of the standard delivery of psychotherapy (De Jong et al., 2021; Lutz et al., 2021).

Hence, it is a method that has the potential to enhance existing therapy outcomes for some patients at a relatively small additional cost (Delgadillo et al., 2021). Developments in information technology and software packages have likely made the adoption of repeated session-by-session outcome measures less onerous for individual practitioners and contributed to standardization in the production of progress graphs (Ogles et al., 2022). The role of software packages is important, with ROM systems using increasingly advanced statistical procedures that identify whether a patient's progress is *not on track* (NOT) compared with historical data from patients presenting with similar clinical profiles (Drapeau, 2012). Notwithstanding these developments, however, the single area of major concern and focus for action relates to implementation (Bovendeerd et al., 2022; Lutz, Rubel, et al., 2022).

Definitions

Routine outcome monitoring has been referred to by various terms, including the following: outcome feedback, progress monitoring, routine outcome monitoring and feedback, measurement-based care, continuous outcome measurement, and feedback informed treatment. Notwithstanding some technical distinctions among these terms, they all contain common features that can be grouped into three distinct phases: (a) collecting patient data on a regular basis; (b) feeding back these data to the therapist and, on many occasions, also to the patient; and (c) when appropriate, adapting the process or focus of therapy in light of the feedback. These three phases have been presented in the literature as a transtheoretical model of measurement-based care, named Collect, Share, Act (Barber & Resnick, 2022).

An inclusive definition of ROM that captures these three phases describes it as involving "the implementation of standardized measures, usually on a session-to-session basis, to guide clinical decision-making, monitor treatment progress, and indicate when treatment adjustment is needed" (Pinner & Kivlighan, 2018, p. 248). ROM has been described as a "relatively straightforward evidence-based practice . . . that the clinician can add to any type of psychotherapy . . . without requiring changes in that psychotherapy" (Persons et al., 2016, p. 25).

ROM data are drawn from patient-reported outcome measures (PROMs), which are psychometric tools providing the foundation and content base for the related practices of ROM and, crucially, progress feedback and adjustment. The extent to which all three

features (data collecting, feeding back, and adapting therapy) are implemented ranges from no use of feedback, to feedback results being available only to the therapist, to the results being fed back to the therapist and patient resulting in a discussion of PROMs that can affect subsequent treatment (Krägeloh et al., 2015). The latter approach, in which regular feedback based on PROMs is integrated into psychotherapy, is the preferred and probably most effective method. This provides a signal, flag, or alert for the therapist that a patient's treatment response is NOT and that some adaptation to the treatment plan is indicated.

Determining the necessary treatment adaptations is a therapeutic method termed *clinical troubleshooting*, using a combination of clinical and research evidence (De Jong et al., 2023). A similar process occurs via the application of *statistical prediction models* and, in some cases, the provision of clinical decision suggestions in areas of therapeutic work that a therapist could address. This is particularly the case for what have been termed *clinical support tools* (CSTs), which are designed to identify and address problems that might be interfering with treatment progress (e.g., Harmon et al., 2007; Lambert, 2010; Lucock et al., 2015).

The American Psychological Association (APA) has long recommended the use of ROM and feedback methodology in routine care (APA, 2006) along with the centrality of outcome measurement to measurement-based care (Wright et al., 2017, 2020). An APA governance-appointed Advisory Committee for Measurement-Based Care and the Mental and Behavioral Health Registry has argued for a professional practice guideline focusing on measurement-based care. The committee proposes the draft statement: "Psychologists aim to routinely assess treatment process and outcomes and integrate that information in ongoing collaboration with their patients" (Boswell et al., 2023, p. 9).

The Joint Commission (2018), a U.S.-based organization focused on quality improvement of health care, while always promoting the assessment of outcomes, now requires organizations

> to accomplish this [assessment of outcomes] through the use of a standardized tool or instrument. Feedback derived through these standardized instruments may be used to inform goals and objectives, monitor individual progress, and inform decisions related to individual plans for care, treatment, or services. (p. 1)

The Roadmap for Mental Health Research in Europe has adopted a similar position (Emmelkamp et al., 2014).

Aside from recommendations of various organizations, national policymakers and regulatory bodies in some countries have made measuring treatment outcomes a requirement, such as in Australia (Burgess et al., 2015), Canada (Tasca et al., 2019), England (Clark, 2018), and Norway (Knapstad et al., 2018). In other countries, national (e.g., the Netherlands; Delespaul, 2015) and local (e.g., Israel; Tzur Bitan et al., 2018) efforts to implement ROM have proved challenging.

ROM has been widely espoused in multiple adult therapeutic modalities—for example, psychodynamic therapy (Winkeljohn Black et al., 2017), couple therapy (Anker et al., 2009), and group therapy (Slone et al., 2015)—as well as with specific presenting problems, such as substance abuse (Crits-Christoph et al., 2012), and in youth settings (Bickman, 2008). It

is pan-theoretical and, as a therapeutic method, combines elements of supervision, continuous assessment, and overall quality assurance. However, in the absence of policy or clinical directives to implement ROM, the willingness to adopt and use this method is likely to be a function of the openness of individual therapists (e.g., Rye et al., 2019).

Clinical Description and Indications

Routine outcome monitoring presents three sequential phases: collection of data, feeding back data, and adapting therapy (Collect, Share, Act [Barber & Resnick, 2022]). Importantly, each phase builds on and is dependent on the preceding phase.

Collecting Patient Data on a Regular Basis

The first phase of ROM comprises consideration and selection of a standardized outcome measure, with the two most frequently researched feedback systems being the OQ-45 System comprising the OQ-45 measures and Analyst software (Lambert et al., 2013) and the PCOMS (Duncan & Reese, 2015), with the latter often referred to by its two separate components—the Outcome Rating Scale (ORS; Miller et al., 2003) and the Session Rating Scale (Duncan et al., 2003). There are also other outcome measures that have been used for ROM (for reviews, see Drapeau, 2012; see also Barkham et al., 2015; Boswell, Kraus, Castonguay, et al., 2015; Brown et al., 2015; Kopta et al., 2015; Miller et al., 2015; Youn et al., 2015).

More recently, newly developed ROM systems incorporate sophisticated prediction and decision-making tools. For example, the Trier Treatment Navigator (TTN; Lutz et al., 2019; Lutz, Deisenhofer, et al., 2022) includes a clinical support tool that provides treatment recommendations at the beginning of treatment as well as adjustments later in therapy in response to NOT signals. Another hallmark of more recent ROM systems is their iterative development in which a system is continually adapted to the needs of practice (e.g., NORSE Feedback System; McAleavey et al., 2021; Moltu et al., 2021).

Whichever measurement tool is adopted, consideration needs to be given to a range of factors, including suitability (i.e., fit) with the patient population, clinical utility (i.e., being clinically meaningful to practitioners), having psychometrically sound measurement properties (i.e., being valid and reliable), being financially viable in terms of resources needed to support implementation, and of minimal burden to patients in light of the need for repeated administration. The selection of the measure will likely have greater subsequent success if the decision on which measure or system to adopt is made collectively within the clinical team. Once a decision has been made, a priority is for all practitioners to be trained in the administration, use, and interpretation of the measure.

Practitioners need a thorough knowledge base of the measure in order for them to introduce ROM at the start of a course of therapy. Crucially, they need to communicate to patients the rationale for completing the outcome measure and explain its role in therapy. ROM will likely have greater utility if introduced and explained at the earliest stage of individual therapy, and it is important to orient patients to the purpose, structure, and use of ROM as part of the routine process of therapy. In particular, therapists

need to explain how the questionnaire is interpreted and how it might inform the treatment process. Explaining the interpretation of questionnaires can enable a transparent and collaborative conversational approach to the process of ROM.

The introduction of ROM by the therapist to a patient will likely reflect a blend of standard and personalized responses based on clinical principles (De Jong et al., 2023). These include offering explanation ("*We're using the measure to track progress . . .*"), personalizing ("*Completing items is an extension of your voice—it says something about you*"), and using language that acts as a bridge between ROM and everyday life ("*So ROM is like a thermometer . . .*"). In addition, regular checking with the patient ("*Does that make sense to you?*"), setting expectations ("*So, filling in that questionnaire is something you'll be doing every session . . .*"), and sharing information about ROM ("*So, when I look at the scores you have given, what I see is . . .*") provide the basis for integrating ROM into routine practice.

ROM can be collected session-by-session or less frequently but at regular intervals (e.g., monthly). The former is likely to occur in contexts that are more resource rich (e.g., research clinics) or where it is nationally mandated (e.g., the English NHS Talking Therapies for Anxiety and Depression program, previously known as the Improving Access to Psychological Therapies program; Clark, 2018). It is likely that session-by-session ROM data collection is more important for shorter term treatments, whereas more intermittent but regular administration of measures fits with longer term treatments. Regardless of the regularity of data collection, it probably works best to review ROM data early in the session in order to leave available time to respond to the most recent data as they may provide a focus for the current session. But it is important for reviewing data not to override any immediate clinical concerns expressed by the patient at the outset of a session.

Some final considerations in relation to data collection. First, the focus need not be totally on symptoms. A patient's quality of life and interpersonal relations are equally as important, and measures tapping these aspects of life might be considered within a ROM system (e.g., Keetharuth et al., 2018). ROM can also include measures of within-session processes (e.g., the therapeutic relationship). Second, almost all ROM measures are nomothetic, and there are increasing calls for idiographic measures to be included (e.g., PSYCLOPS [Ashworth et al., 2004]; see also Sales et al., 2023). Third, as with using any outcome measure, nomothetic or idiographic, consideration should be given to any cultural adaptations that would increase the accessibility of ROM for patients from culturally diverse and ethnic minority populations (e.g., Koslofsky & Rodíguez, 2017).

Overall, the aim of this initial phase is for practitioners to give patients the rationale and appropriate information for them to understand the measure(s) they are routinely completing and the role ROM plays in the therapy process. This initial work creates the most favorable context for patients to engage in and value ROM such that they complete the outcome measure(s) on a regular basis as agreed with their therapist.

Feeding Back Data to Participants

The second phase comprises the therapist obtaining outcomes information directly from the patient or the feedback system and then sharing this information with the patient. This is best achieved by the therapist reviewing outcome measures with the patient

either proactively (i.e., routinely) or in response to an alert generated by a monitoring system. Data need to be reviewed in the form of a conversation rather than simply being told the numbers. Hence, rather than "*The data shows you to be severely depressed this week*," it is probably better to frame as a proposition that allows the patient to confirm, question, or qualify the response: "*The data places you in the severe range this week; does that match with how you have been feeling?*"

The timing of feedback is important, and a balance needs to be taken between following an overly protocolized agenda (e.g., "*We always check on ROM scores at the beginning of a session*") and being responsive to the current state of the patient and their need to talk about a specific clinical issue (Drew et al., 2021). Regardless of when ROM is reviewed, the aim is to ensure not only the continuous reinforcement of outcome monitoring as a shared task but also that potential problems (e.g., symptomatic deterioration) can be prioritized and dealt with as part of the session. If use of the data is to be actionable, then there needs to be sufficient time remaining in the session to respond to the data and their implications.

In most cases, measures will indicate that the patient is progressing as expected, in which case the therapist can offer this interpretation and check the patient's perspective about this interpretation. In addition, a therapist can track responses to individual items within any measure—for instance, noticing which items have changed more substantially than others since the last measurement (Cross et al., 2015). In this way, marked change in individual items, showing either improvement or deterioration, can provide a clinical focus and thereby act as a clinical process tool for therapeutic discussions (Faija et al., 2022). Such practice integrates the outcome measure and its individual items into the therapy session and conversation.

Two main methods are used to generate feedback: rational and empirical. *Rational* methods determine a cutoff point for considering a patient to have worsened based on rational thinking combined with a priori clinical judgment based on clinical rule of thumb, distribution-based rule, or psychometric rule. The clinical rule of thumb adopts a criterion related to the performance on the outcome measure based on clinical impressions, whereas the use a distribution-based rule is usually based on the population norms for an instrument. However, this latter method does not take change into account. Psychometric rules use certain statistical characteristics of the outcome instrument as a criterion, with the Reliable Change Index (RCI) being the most commonly used (Jacobson & Truax, 1991). This method is determined by the specific psychometric properties of a given measure. A more sensitive index for deterioration might be the use of the raw measurement error of the instrument, typically smaller than the RCI (Nugter et al., 2013).

Empirical methods use data from historical cases to generate expected recovery curves (Finch et al., 2001) or expected treatment response models (Lutz et al., 1999). These procedures predict how new patients will progress in therapy compared with patients from the archived data set who presented with similar characteristics (i.e., patients who are, clinically, nearest neighbors; Lutz et al., 2005). During treatment, a patient's actual progress is compared with the prediction, and when a patient deviates by a predetermined amount from the predicted course, a warning signal is given to the therapist that the patient is NOT as opposed to being on track (OT).

In addition to feeding back information derived from visual analyses of outcome measures and numerical data, therapists can use CSTs (Lambert et al., 2015; Slade et al.,

2008), also referred to as *clinical problem-solving tools* (Lutz et al., 2021). Examples of CSTs can be seen in the OQ-System (Lambert et al., 2013) and the TTN (Lutz et al., 2019). The main component of CST feedback within the OQ-System is the Assessment for Signal Clients (ASC; Lambert et al., 2015), a 40-item self-report questionnaire that is completed by patients who are deemed at risk of deterioration and focuses on four domains: (a) therapeutic alliance, (b) motivation, (c) social support, and (d) coping with problematic life events. With domain scores benchmarked against norms, the ASC helps the therapist identify specific areas that may be contributing to a patient's deterioration, with the aim of utilizing this information in clinical decision-making.

The ASC has also been used as a central component in the TTN, but supplemented with the Affective Style Questionnaire (Hofmann & Kashdan, 2010), a 20-item self-report measure assessing emotional regulation. Both measures are administered at every fifth session. Figure 15.1 displays the output from the TTN showing routinely collected data [A] as against the expected treatment response curve [C] based on a nearest neighbor

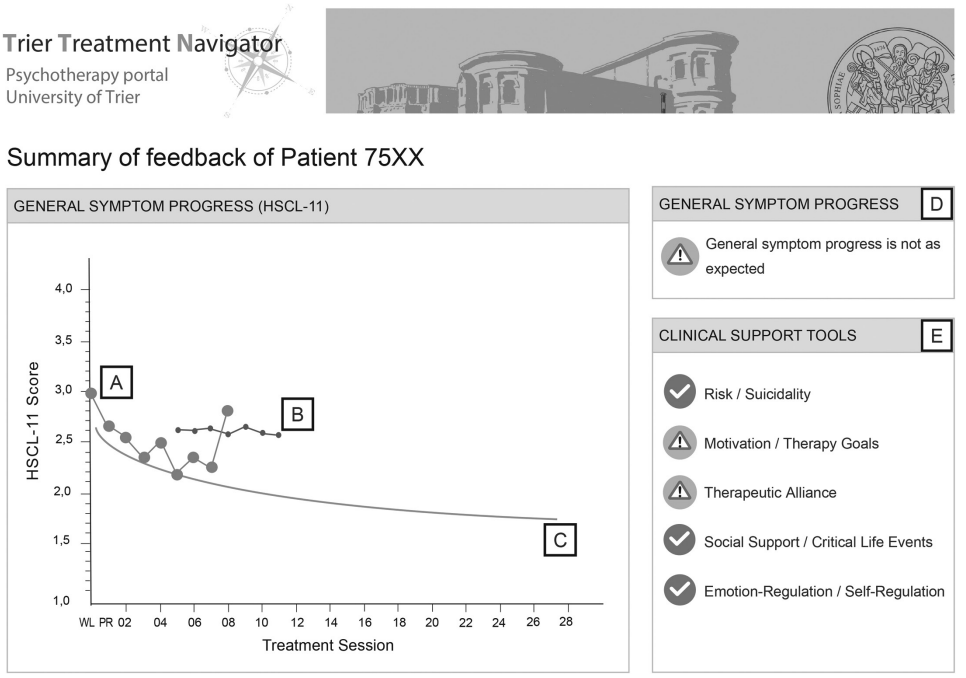

Figure 15.1 Example of feedback graph with clinical support tools. The screenshot shows the symptomatic progress of a patient measured with the Hopkins Symptom Checklist–11 (HSCL-11) within the clinical navigation system by Lutz et al. (2019). A: Individual measurement points for the patient measured at the beginning of each session; B: Dynamic failure boundary; C: Expected treatment response curve; D: As soon as the patient's HSCL-11 score exceeds the failure boundary (marked in the graph with an arrow), the therapist receives a warning signal (top right corner) and more detailed information is provided by the clinical problem-solving tools (CSTs) below; E: CSTs 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 signals that the patient has few or no problems in the respective domain.

approach—that is, based on a defined number of patients from a data archive who most closely resemble the specific patient. The crucial flag occurs at session 9 when the outcome measure exceeds the dynamic failure boundary [B], producing a warning flag [D]. Completion of the clinical support tools [E] shows that the patient is experiencing difficulties regarding motivation/therapy goals and the therapeutic alliance, which then provide a refocusing of therapy to address these domains in the third phase of ROM.

Adapting the Process or Focus of Therapy in Light of Feedback

The third phase comprises, when necessary, adapting or reorienting the focus of psychotherapy in light of feedback. This phase is a clinical process that determines the decisions and actions of the therapist in response to the available feedback data. Hence, as in Figure 15.1, with the aid of clinical support tools, the TTN includes adaptive recommendations during treatment based, in this example, on data from five patient domains of experience, which are identified for patients at risk for becoming NOT cases: (a) risk/suicidality, (b) motivation/therapy goals, (c) therapeutic alliance, (d) social support/critical life events, and (e) emotion regulation/self-regulation. Such systems provide predictions derived from the use of computational algorithms drawing on many previous cases that yield therapeutic suggestions.

When CSTs are not available, the principle is the same, namely addressing issues that appear to be stalling or impeding a patient's improvement, but is achieved via clinical troubleshooting. Troubleshooting is informed by a therapist's knowledge of the clinical and research evidence relating to a patient's response to treatment. This might include contextual (e.g., lack of social support in the patient's life), process (e.g., a difficult therapeutic relationship with the patient), and patient factors (e.g., the presence of comorbid conditions). These domains provide starting points for a therapist to explore and consider factors that may be impinging on a patient's progress and provide clues as to possible adaptations. From here, a therapeutic plan is developed, implemented, and evaluated through subsequent ROM data. Data collected at each subsequent session provide timely feedback on whether the adaptations to the treatment plan have led to scores moving back in the direction of being OT. Finally, the actions and responses need to be reviewed in the context of the overall treatment plan. For example, it may be that a repair to the alliance enables the original treatment plan to resume. Or it may be that adaptations are made to work around cultural issues, or that the treatment plan itself is altered in order to change the focus of therapeutic work.

Assessment

During the Course of Therapy

Assessment of ROM considers both adherence and competence. Adherence refers to the completion of outcome or session measures, involving patients, therapists, and the supporting administrative system. Competence focuses on active use of the feedback by the therapist in adapting the focus or direction of therapy.

The simplest index of adherence regarding data collection is the proportion of completed outcome measures: The actual number will depend on the agreed frequency of completion. This adherence index can also be enhanced by an index of the quality of data completion; for example, a higher rate of missing items will make the data less robust but may also indicate a clinical concern for a patient, highlighting the value of missing data—or be indicative of an issue with the measure itself (e.g., a poor cultural fit regarding an item for certain patients).

In terms of sharing feedback data, several key therapist factors can be assessed. One factor is the therapist's use of externally sourced information, which can be measured by the Internal and External Feedback Propensity Scales (Herold et al., 1996). These scales tap aspects of internal propensity (e.g., "As long as I think I have done something well, I am not too concerned about how other people think I have done") and external propensity (e.g., "I like getting frequent feedback from others regarding my performance"; Herold et al., 1997). Commitment to providing feedback has also been measured using an adapted version of the Goal Commitment Scale (Hollenbeck & Klein, 1987), in which, for example, the item "It's hard to take this goal seriously" was adapted to read "It is hard to take [the idea of using these measures in my clinical practice] seriously."

The assessment of competence in terms of adapting the focus or direction of therapy is more difficult to assess and is largely in the sphere of clinical supervision. The outcome of any adaptations should yield session outcome data moving back below the failure boundary (recall Figure 15.1) or where the direction of change is generally positive.

Distal Treatment Outcomes

Two key distal outcomes of ROM relate to end-state and rate of change. The end-state is indexed by the proportion of patients deteriorating in the feedback condition as contrasted with the comparison group. In effect, this is the main index for determining the effectiveness of ROM and will mean a patient's scores at the end of psychotherapy being within the boundaries of the expected treatment response curve. Additional criteria to assess impact are determining whether the first to last change score exceeds the RCI and whether the end score identifies the patient as being both reliably and clinically significantly improved (Jacobson & Truax, 1991). In addition to the end-state assessment, the rate of change is a further index of change that considers the number of sessions attended to achieve an outcome (Lambert et al., 2001). Another index is a reduction in dropout rates (De Jong et al., 2021). Evidence from meta-analyses has usually reported outcomes based either on all patients (i.e., OT and NOT) combined or separately. For example, the PCOMS has generally considered all cases, whereas the OQ-45 has more often considered the subgroup of NOT cases.

The adoption of the criteria of reliable and clinically significant change in determining the effectiveness of ROM in routine practice has been challenged. This challenge largely rests on the basis of an absence of any counterfactual—that is, what the outcome would have been had feedback *not* been provided (Langkaas et al., 2018). It is argued that although adopting methods of reliable and clinical significance captures change, these methods do not determine *induced difference*—that is, the effect induced specifically by the actions comprising the ROM procedure.

Clinical Example

The clinical example presented here provides an account of the three distinct phases: introducing ROM, feeding the results back to the patient, and adapting therapy in light of feedback information, informed by clinical troubleshooting. This case is based on a real patient who gave permission for it to be used for teaching purposes. The transcript is based on clinical case notes, and some elements of the case have been changed. A fictitious name is used for the patient.

Samira was a 22-year-old woman of color, unemployed, in substantial financial debt, and sharing a house with her partner and his brother. Samira sought psychotherapy after being the victim of an armed violent attack in her house by criminal debt collectors. She was pregnant at the time, and the attack led to a miscarriage that was accompanied by substantial blood loss, which stained the carpet in her house. After the traumatic event, Samira developed obsessive cleaning behavior and disinfecting rituals. She met criteria for post-traumatic stress disorder (PTSD) and contamination-focused obsessive–compulsive disorder (OCD) at the start of treatment. In collaboration with Samira, it was decided that the PTSD symptoms were the most debilitating and would be the primary focus of treatment. Samira wanted to feel like she could speak to people again, and leave the house by herself and eventually get a job and make friends. The recommended treatment was cognitive–behavioral therapy with imagery exposure for PTSD.

Introducing ROM: Early Sessions

Samira (S) was asked by the therapist (T) to complete the Generalized Anxiety Disorder–7 (GAD-7) questionnaire prior to each session through an electronic form. During the first session, Samira was introduced to working with ROM:

T: An important part of the treatment involves keeping track of your anxiety over time, so that we can get an idea of whether the treatment is working and helping you feel better or not. We use a brief questionnaire to do this. Were you able to complete the questionnaire before this appointment?

S: Yes, here it is.

T: Great, thanks. Based on your answers, I can work out a score that tells us how intense the anxiety has been recently. You can think of it as a sort of thermometer of distress. Like your temperature, which fluctuates from day to day, your anxiety symptoms also go up and down, and this measure helps us to track these changes during treatment. Does that make sense?

S: Yes, true. I have good days and bad days.

T: Exactly. If we can keep track of these ups and downs, it can help us to learn about the things that make your anxiety worse and the coping strategies that might help to make things better for you. If therapy helps, we might expect to see these measures going down over time.

S: OK, I hope so. I find that the anxiety is overwhelming lately . . .

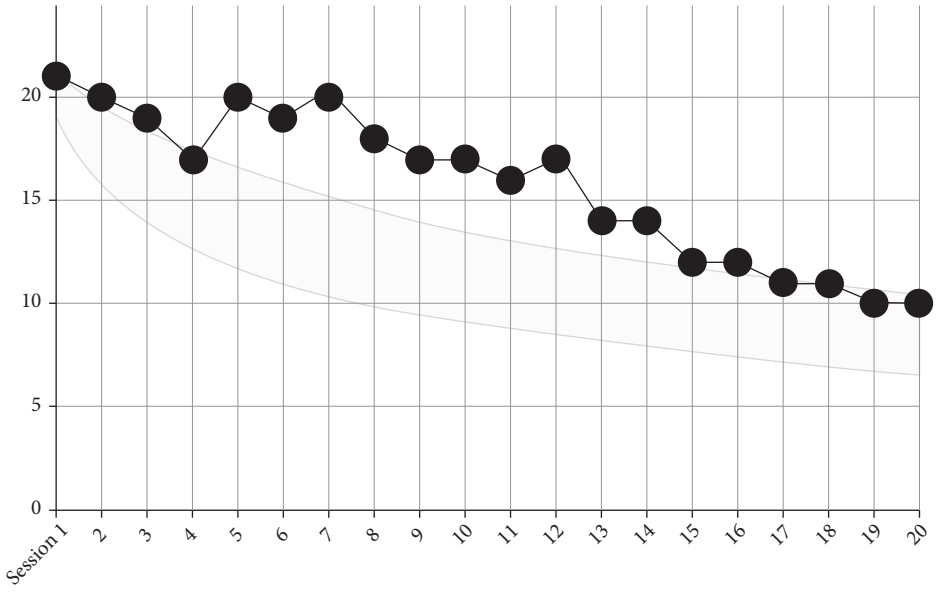

Figure 15.2 Symptoms per session on the Generalized Anxiety Disorder–7 questionnaire.

During the first session, Samira received an explanation about the interpretation of the scores and was informed that the therapist would discuss her progress with her briefly at the start of each session. At the next session, Samira had had more interchanges with her partner over her cleaning behaviors, and she asked the therapist for advice on how to navigate the irritation that her cleaning behavior caused her boyfriend, resulting in frequent arguments. Samira and her therapist spent more than half the session on problem-solving this matter. At the following session, this seemed to have paid off, and her anxiety score had reduced a little.

Feeding Back in Response to an Alert Signal: Sessions 5–7

The ROM system calculated an expected treatment recovery (ETR) curve, which provided an alert to the therapist at session 5 denoting that Samira was NOT as her score clearly exceeded the failure boundary (upper green/gray line in Figure 15.2). There was a substantial increase in anxiety symptoms, and it was evident to the therapist that the treatment was going in the wrong direction and the patient was at risk of ending therapy. So, the therapist discussed potential reasons for Samira's increase in anxious symptoms with her:

T: Shall we have a look at your progress chart? Although the scores in the initial sessions showed improvement, the scores now show your anxiety is going up again this week, and this time quite a bit more than in your second and third session. Does that fit with how you have been feeling in the past week?

S: Um . . . yes. I find it so hard to get through each day. My boyfriend is very annoyed with me about the amount of time cleaning. He says that I should focus on getting a

job instead, so that I can start paying off my debts. These are what got me in trouble in the first place.

T: How does that make you feel when he says that?

S: Well, I don't know. I guess he is right. It is my fault that we got attacked. If I had not borrowed money from those loan sharks, we would not have been attacked, and I would still be carrying my baby. I wish we were not fighting all the time though . . . we used to be able to talk to each other about everything, but since the attack we just can't anymore, and I don't know what to do anymore.

T: It sounds like a lot is going on at the same time, that makes you feel quite anxious. It sounds like the fights with your boyfriend are making you feel very tense. Can you think of other things that have a negative influence on how you are feeling at the moment?

S: My debts really cause me a lot of anxiety as well. At the moment I cannot even leave the house. How am I ever going to start a job and resolve my debts? I just don't know what to do.

T: Yes, I fully understand that your financial problems are very stressful. What else is going on at the moment? Is there anything in the therapy that might not work for you, perhaps?

Samira engaged in a conversation with the therapist and together they noticed that the comorbid OCD symptoms were a source of conflict in her intimate relationship, and this caused Samira to bring it up in treatment a lot.

In the next few sessions, Samira and her therapist tried to get a fuller assessment of whether the trauma was still the appropriate focus and how the OCD symptoms should be discussed in treatment. Through a functional analysis, Samira and her therapist concluded that Samira was using the cleaning as a way of preventing engagement with the traumatic memory. Bringing the OCD symptoms up in session prevented her from having to work on the trauma, and it gave her short-term relief in terms of discussing strategies of dealing with the discourse in her relationship. The behavior itself also made her feel less dirty, and she would feel more relaxed after cleaning.

Making Adaptations to Therapy: Sessions 8–12

In order to disrupt the maintenance cycle of OCD and PTSD, the following adapted treatment plan was agreed between the therapist and patient: (a) Reduce cleaning time (exposure and response prevention), replace the cleaning time with (b) writing and reading a trauma account, (c) a problem-solving window using a structured worksheet, and (d) relaxing activities outside of the house (e.g., walking, drawing, photography). The rationale was that the first strategy would disrupt her OCD and open the possibility of approaching (rather than avoiding) exposure to the trauma memories. The second strategy enabled Samira to start to become exposed to trauma memories in a contained, structured, gradual, and tolerable way. The third strategy enabled her to map out possible solutions to pressing problems that she found stressful and demoralizing (financial problems, relationship problems). The fourth strategy enabled her to gain some sense of relief and relaxation, which she was previously attempting to achieve (ineffectively)

through compulsive cleaning. These strategies were integrated into her routine slowly over several weeks (sessions 8–12).

Therapeutic Impact of Adaptations: Sessions 13–20

Samira's anxiety began to go down slowly as assessed by the patient-completed GAD-7. The improvement in symptoms helped Samira become more engaged with the therapy starting at session 13, and her symptoms continued to improve. By session 16, Samira started searching for a job, and her scores began to fall within the boundaries of the ETR curve. It was agreed to bring the therapy to a close at session 20. At that point, Samira scored 10 on the GAD-7, indicating reliable improvement and within the boundaries of the ETR curve.

Overall, this case illustrates the need to personalize the treatment plan by identifying, formulating, and working with/around complicating factors, until the point when an evidence-based treatment can be resumed (e.g., when the obstacles have been dealt with, at least partly). In this way, feedback-informed treatment balances both evidence-based practice (e.g., applying empirically supported treatment methods) and the flexible personalization of the treatment plans—specifically to address obstacles to improvement. Other detailed clinical case studies documenting the role of ROM in adapting treatment plans have been reported in the literature (for a sample of clinical cases, see De Jong & Aafjes-van Doorn [2022] and associated special issue in *Journal of Clinical Psychology* [2022]).

Landmark Studies and Previous Reviews

Landmark Studies

The original investigation of patient feedback was conducted by Lambert and colleagues (2001), noteworthy for its experimental design and replication (Lambert et al., 2002). A previous review of this topic (Lambert et al., 2019) identified four landmark studies, two of which utilized the OQ-45 (Harmon et al., 2007; Simon et al., 2013) and two the PCOMS (Anker et al., 2009; Schuman et al., 2015). We take a complementary approach by prioritizing methodological rigor (reduced risk of bias in studies) as well as methodological advances drawn from the most recent substantive meta-analysis (De Jong et al., 2021), in which only four of 58 studies met the highest threshold for addressing risk of bias. Chronologically, these studies (and nation) were Berking et al. (2006; Switzerland), Davidsen et al. (2017; Denmark), Errázuriz and Zilcha-Mano (2018; Chile), and Bovendeerd et al. (2022; the Netherlands). Here, we consider the three most recent studies, all of which used components of the OQ and/or PCOMS systems, and none were carried out by the measure developers, thereby minimizing any effects of system developer allegiance.

Davidsen et al. (2017) conducted an RCT to determine the effect of client feedback in group psychotherapy for patients diagnosed with eating disorders (bulimia nervosa, binge eating disorder, or eating disorder not otherwise specified). Attendance and

treatment outcomes were the main foci. Randomization was stratified for treatment type and diagnosis, with treatment allocation concealed from the investigators. Eighty participants were allocated to the experimental group, in which participants gave and received feedback about therapy progress and alliance, measured before and after each session using the ORS and Group Session Rating Scale. Seventy-nine participants were assigned to the control condition (no feedback), with both groups receiving 20–25 weekly group psychotherapy sessions. Rate of attendance at treatment sessions was the primary outcome, and severity of eating disorder symptoms measured with the Eating Disorder Examination interview was the secondary outcome. Patient outcome measures were the Symptom Checklist–90-R and the ORS (psychological distress), the Sheehan Disability Scale (social functioning), and the Self-Harm Inventory (episodes of self-harm and suicide).

Results showed feedback compared with control did not affect the rate of attendance, the severity of symptoms, or any of the exploratory outcomes (all *p* values ranged from .06 to .67). Davidsen et al. (2017) concluded that feedback neither increased attendance nor improved outcomes for outpatients in group psychotherapy for eating disorders. Crucially, however, it appeared that although therapists had previously endorsed the feedback procedures and looked at the PCOMS scores, they rarely acted on the information to adapt the course of or length of treatment.

Errázuriz and Zilcha-Mano (2018), in one of the few non-Western studies in the field, used three outcome indices—symptomatology, the working alliance, and both combined—to evaluate a low-cost form of feedback compared with the OQ-45 progress feedback report, thereby yielding four active feedback conditions plus a no feedback condition. They also investigated a range of moderators, including patient, therapist, and process factors, and employed multilevel modeling as well as conducting follow-up interviews with therapists. The sample comprised a total of 547 patients, with each condition comprising between 104 and 116 patients. The total sample was 75% female, average age 41 years, with 95% of the sample being Latino and treated in an outpatient individual psychotherapy setting in Chile.

Results showed that feedback had no effect on outcome, session attendance, or alliance for any of the cases, including NOT cases. Because therapists could decide whether or not to share the feedback they had received with the patient, ascertaining the extent to which feedback was shared with patients was important. Follow-up interviews identified that 64.7% of therapists used feedback sometimes or always, and when only this sample was analyzed, the results were the same. That is, there was no main effect for feedback based on the whole sample, on just the NOT cases, or for therapists who did report using the feedback.

However, Errázuriz and Zilcha-Mano (2018) did find that in NOT cases, baseline severity moderated feedback effects and that patients with a low baseline severity benefited more in most of the feedback conditions. Furthermore, for highly severe patients, the outcomes in the control group and the group receiving feedback on the alliance were superior to outcomes for patients receiving feedback regarding their symptoms. This effect was largely due to there being adverse effects for this specific patient group (i.e., more severe). Crucially, however, because the study was focusing on evaluating a low-cost implementation of feedback, clinical support tools were not included in the study. The authors concluded that providing feedback to therapists without offering them tools to

improve treatment may be ineffective and even detrimental. Therapists receiving negative feedback in relation to patients with more severe disorders may be placed at a disadvantage without access to the clinical support tools that appear to play a crucial role in delivering effective feedback.

A significant portion of the literature has been characterized by biases arising from researcher allegiance and from a lack of independent outcome measures (i.e., independent of the feedback measure). Bovendeerd et al. (2022) addressed these two key issues in a cluster randomized trial of systematic client feedback (SCF) in which feedback using an adapted form of the PCOMS was compared with treatment-as-usual (TAU). The independent outcome measure was the OQ-45. Patients were drawn from four treatment centers, and power calculations were based on assumptions of a small effect, resulting in a minimum of 208 patients per center and 16 patients per therapist. A total of 1,933 patients were recruited into the trial, but 200 patients did not attend therapy, yielding an analyzed sample of 1,733 patients. Patient presentation was mild to moderate depression and anxiety (OQ-45 baseline scores for SCF and TAU conditions were 76.12 and 73.89, respectively).

Two analyses were conducted—intent-to-treat (ITT) and per protocol (PP)—and a three-level model was used (patient, therapist, center). For the ITT analysis, the gains in the feedback condition exceeded those for the TAU condition at all assessment points. At 1 year, the gains were more than 25% greater for the feedback condition on the OQ-45 compared with the TAU condition, representing an additional gain of 5.25 OQ-45 points (24.88 vs. 19.63). Although patients in both conditions made the majority of their gains within the initial 3 months, the key finding was the additional gains made across 12 months in the feedback condition compared with TAU. A virtually similar result was obtained in the PP analysis. No significant effects were found on the other outcome variables.

These three high-quality studies provide a representative summary of the range of outcomes for ROM. The null findings from group (Davidsen et al., 2017) and individual therapy (Errázuriz and Zilcha-Mano, 2018) indicate that ROM is not a panacea for all patients under all conditions. The former may be a result of applying feedback within highly structured therapies in which therapist flexibility in individually adapting therapy is limited. It may be that when applied to more intensive treatments, there is a ceiling effect that mitigates against additional gains from ROM.

In terms of the Errázuriz and Zilcha-Mano (2018) study, the result clearly identifies the importance of moderators, in particular patient severity, with the salutary finding that negative feedback can have an aversive effect. What is unknown from the study is whether the inclusion of CSTs, which are associated with the largest feedback effects, would have obviated the aversive impact. Although the attempt to test low-cost variants of feedback in a routine non-Western clinical setting is one of the most real-world controlled designs in the field, it may be that the exclusion of a key component in effective feedback compromised its impact. However, it showed the potential role of moderators, albeit in the context of an overall null finding. The study confirmed the importance of determining under what specific conditions feedback is effective, which may allow for more targeted efforts.

The positive findings using samples from national health care settings with an independent outcome measure (Bovendeerd et al., 2022) endorse findings, for example,

from the English national NHS Talking Therapies for Anxiety and Depression program comprising a large analyzed sample (1,733 patients) in which, after controlling for therapist effects, patients classified as NOT had less severe symptoms ($d = 0.22$) after treatment as measured by an independent outcome measure when they received feedback compared with those in the control group. The study showed that supplementing psychological therapy with low-cost outcome feedback improved outcomes for patients who were at risk of poor response to treatment (see Delgadillo et al., 2018).

Meta-Analytic Reviews

At least 50 controlled, quasi-experimental, observational, and implementation studies have investigated the effectiveness of ROM feedback systems applied in several countries during the past 20 years. These studies have been synthesized across numerous systematic reviews (e.g., Carlier et al., 2012; Davidson et al., 2014; Gondek et al., 2016; Krägeloh et al., 2015; Mackrill & Sørensen, 2020) and at least 11 meta-analyses (Bergman et al., 2018; De Jong et al., 2021; Kendrick et al., 2016; Knaup et al., 2009; Lambert et al., 2003, 2018; Østergård et al., 2020; Pejtersen et al., 2020; Rognstad et al., 2023; Shimokawa et al., 2010; Tam & Ronan, 2017).

Early meta-analyses suggested that ROM feedback improves distal treatment outcomes relative to usual psychological care. However, trials published at that time (approximately a dozen) were predominantly from the United States and located in university counseling settings (Knaup et al., 2009; Lambert et al., 2003; Shimokawa et al., 2010). A seminal meta-analysis reported a pooled (intention-to-treat) effect size of $g = 0.28$ ($p = .003$) favoring feedback-informed treatment relative to usual care in cases classed as NOT using the OQ-45 measure (Shimokawa et al., 2010). The pooled (intention-to-treat) effect size was larger when examining data from studies that supplemented feedback with clinical support tools ($g = 0.44$). These promising results motivated feedback studies in other countries, treatment settings, age groups, and outcome measures.

Recent meta-analyses have examined the effects of feedback with specific measures, principally the OQ-45 and PCOMS (Lambert et al., 2018; Østergård et al., 2020; Pejtersen et al., 2020), and on specific populations such as children and young people (Bergman et al., 2018; Tam & Ronan, 2017) and presenting conditions (e.g., common mental health disorders; Rognstad et al., 2023). As the literature from trials accumulated, meta-analyses of ROM with heterogeneous populations and outcome measures emerged. Such broad meta-analyses reported discrepant findings. For example, one review reported a pooled effect size of $g = 0.10$ favoring feedback in short-term therapies but no significant effect in long-term therapies (Knaup et al., 2009), whereas another review reported no significant overall effect of feedback but reported a significant effect size of $g = 0.22$ favoring feedback in NOT cases (Kendrick et al., 2016). The most comprehensive meta-analysis to date, including an examination of potential moderators, is reported in the next section (De Jong et al., 2021). A further recent meta-analysis of only RCTs reported a virtually identical effect size reported by De Jong et al. (2021) of 0.14 for a total sample and slightly larger effect size of 0.29 for NOT cases (Rognstad et al., 2023).

Overall, the additive effect sizes from these meta-analyses of psychological therapy with feedback compared to psychological therapy without feedback can be summarized as ranging from small ($g = 0.14$; Lambert et al., 2018; Rognstad et al., 2023) when based on all cases, to medium ($g = 0.33$; e.g., Lambert et al., 2018) when based on NOT cases, and increasing to 0.49 when CST feedback was used (Lambert et al., 2018). These mixed findings seem to be a function of the analyzed samples and their heterogeneity, which requires a comprehensive examination of potential moderators of effects sizes across studies.

Research Review

In this section, we provide an abridged version of the largest meta-analysis conducted to date on ROM (De Jong et al., 2021). Readers are directed to the original source for the full details, which is available under Creative Commons License. Given that the search was carried out in September 2020, within the preparation time frame of this chapter, together with the extensive inclusion criteria of the study, we considered this review to be sufficiently up-to-date so as not to warrant a new search and meta-analysis. It comprised both RCTs and practice-based studies, did not limit the focus to selected feedback systems, included studies in several languages, and was not conducted by a developer of a feedback system.

Search Strategy and Inclusion Criteria

The starting search date was set at 2001 (this marked the introduction of progress feedback) and included literature to September 30, 2020. The sources for the searches comprised PsycInfo, PubMed, and Web of Science, with additional searches of the Current Controlled Register Trials Register and Google Scholar. Indicative search terms for PsycInfo included the terms "feedback" OR "outcome monitoring" OR "routine outcome measurement" OR "progress monitoring" within the context of psychotherapy, counseling, and mental health care. Exclusion terms included NOT education NOT (video-feedback OR video feedback) NOT (bio-feedback OR biofeedback OR "EEG feedback" OR brain) NOT (schizophrenia OR psychosis) NOT motor. Full search terms for all databases are available as supplemental material linked to the original publication. In addition, existing meta-analyses were screened for relevant publications as well as calls using the listserv of the Society for Psychotherapy Research and emails to known researchers in the field seeking unpublished studies. Inclusion criteria were published and unpublished studies that (a) examined effects of feedback on patient outcome; (b) compared one or more feedback groups and a no feedback control group or cohort; (c) focused on psychological interventions in a psychological therapy, psychiatry, or counseling setting; (d) treated patients with mental health or substance abuse problems; (e) assessed end-of-treatment outcomes; and (f) in which therapists received ROM information about their patients on at least three occasions during treatment. Documents in English, German, and Dutch were included, a feature that enhanced the inclusivity of the meta-analysis in comparison with previous meta-analytic reviews.

Progress feedback was defined as "providing information on treatment progress from standardized measures to a clinician and/or patient on a regular basis throughout the course of treatment." Studies were coded for a range of feedback characteristics (e.g., instrument, type of feedback) and study characteristics (e.g., treatment duration, setting). Of 40 authors approached by the team for additional information about their data, 31 provided it.

Primary and Secondary Outcomes

The primary outcome was the difference in post-therapy symptom reduction on a standardized outcome measure between patients in the feedback group and patients in the treatment as usual/without feedback group. For cohort studies, the standardized mean difference in change score was taken as the effect size measure (i.e., accounting for the pretreatment scores). To meet criteria for inclusion at follow-up, data were required to be at least 1 month following the end of treatment or study and represent a standardized mean difference in change score from pre-treatment.

Secondary outcomes were post-treatment differences between the feedback and control groups in dropout rate, deteriorated cases, and treatment duration. Dropout included treatment dropouts and study dropouts as well as cases with only one or two sessions. Reliable deterioration was defined according to standard criteria (Jacobson & Truax, 1991), and treatment duration was defined by the number of sessions received.

Study Quality

The quality of RCTs was assessed using the Risk of Bias (Higgins et al., 2020), and the cohort studies were rated using the Risk of Bias in Non-randomized Studies of Interventions template (Sterne et al., 2016). In addition, the potential for allegiance bias was coded in the affirmative if the developer of the feedback system was a co-author on the study.

Moderating Variables

Multiple moderating variables were investigated. These included voluntary versus mandatory participation, treatment intensity, adult versus youth, therapy modality (individual, group), and length (fixed vs. open). Outcome assessment was differentiated according to whether it was independent or not of the measure used to provide feedback. Type of feedback was coded (raw data, ERT, or CSTs), including additional information on aspects of therapy (e.g., alliance). Particular features of feedback were distinguished according to timing (before or within session vs. delayed) and frequency (continuous vs. intermittent). Data were coded whether patients had direct access to feedback or not. Studies were distinguished between those in which therapists received training in the feedback tool versus not.

Meta-Analytic Approach

A three-level meta-analysis was conducted, thereby taking account of three levels of variability in effect sizes (Cheung, 2014): (a) the sampling variance of the individual extracted effect sizes (level 1), (b) the variance between effect sizes extracted from the same study (level 2), and (c) the variance between effect sizes extracted from different studies (level 3). The overall effect size and variance estimates were derived from a random effects three-level model. The likelihood ratio tests determined whether the within and between effect size estimates were significantly different from zero.

The three-level random effects analyses were conducted twice: first for full samples (both NOT and OT cases) using data from all studies, and second for only a subsample of studies providing data for NOT cases. The I^2 statistic was used as a standardized measure of heterogeneity (Nakagawa & Santos, 2012). Because the secondary outcomes (dropout, percentage of deteriorated cases, treatment duration) did not have nested data, standard random and mixed effects meta-analysis models were conducted. Moderation analyses were conducted using the same moderators used in the models as the primary outcome variables.

Egger's regression test of funnel plot asymmetry (Egger et al., 1997) was used to evaluate publication bias for the results of the standard random effect models. An adapted version of Egger's regression test was used to assess publication bias in the three-level random effect models.

For the overall effect of feedback, the fail-safe *N* was estimated. The fail-safe *N* refers to the number of potentially missing studies with a *z* value of 0 that would need to be added in order to make the overall effect size statistically insignificant.

Sample

The final sample comprised 58 studies, 21,699 patients (41.9% controls; 58.1% feedback), and 110 effect sizes. The vast majority of studies were conducted in counseling or outpatient settings (84%), with adults (91%), in individual or couple therapies (71%). A total of 97% of the studies were accounted for by Europe and the United States (52% and 45%, respectively), with the OQ System and PCOMS being the most frequently used feedback systems (38% and 36%, respectively). Use of ETRs as the means for providing feedback was the most common (45%), followed by use of raw scores (38%); ETRs together with CSTs accounted for the least (17%). Feedback was provided immediately (59%) more often than delayed (41%), with 88% of the studies providing session-by-session feedback.

Study Quality

The overall risk of bias was high for 31% of RCTs and 100% for all cohort studies due to the high risk of confounding arising from the absence of randomization. Anticipating the moderator analyses (below), larger effects were obtained in studies in which the developer of the feedback system was a co-author ($d = 0.21$, 95% confidence interval [CI]

[0.12, 0.31]) than not (d = 0.12, 95% CI [0.06, 0.19]). That suggests the possibility of researcher allegiance.

Results for Full Sample

The average sample for the 58 studies was 321.4 (SD = 433.6; median = 184) per study, and using an average of 2.04 (SD = 1.73) outcome measures per study. The overall effect size was d = 0.15, 95% CI [0.10, 0.20] (Figure 15.3). The between-study heterogeneity was 72.6%, and the within-study variance was 0.

Five moderator variables had a significant impact on outcomes. Larger effects were found for the following: (a) the ORS (d = 0.34) compared with the OQ-45 (d = 0.11) or other outcome instruments (d = 0.12); (b) the PCOMS feedback system (d = 0.24) in contrast to the OQ System (d = 0.13) or other feedback systems (d = 0.07); (c) studies conducted in the United States (d = 0.23) compared with those conducted elsewhere (d = 0.11); (d) use of the feedback instrument as the outcome measure (d = 0.19) compared with use of an independent measure (d = 0.08); and (e) an effect of publication date, with later studies yielding an effect size of 0.02 lower per year since the initial study in 2001. Larger effects were also obtained in studies in which the developer of the feedback system was a co-author (d = 0.21, 95% CI [0.12, 0.31]) than not (d = 0.12, 95% CI [0.06, 0.19]).

Ten effect sizes were derived from four studies providing follow-up data. The three-level random effects model yielded an overall effect size of feedback of 0.18 (95% CI [0.03, 0.39]) compared to no-feedback control groups. Due to the small number of studies, this was not statistically significant. Moreover, a sensitivity analysis excluding one study (Anker et al., 2009) substantially reduced the pooled effect size (d = 0.09, 95% CI [0.13, 0.32]).

Results for Not on Track Cases

The NOT subgroup comprised 27 studies, 15,146 patients (40.1% controls; 59.9% feedback), and 43 effects. The analysis yielded an overall effect size of d = 0.17 (95% CI [0.09, 0.25]; Figure 15.4). The effect size increased slightly as a result of sensitivity analyses excluding non-RCTs (d = 0.20, 95% CI [0.11, 0.28]), as well as excluding studies using a control group adopting some form of outcome monitoring (d = 0.20, 95% CI [0.12, 0.27]). The between-group heterogeneity was 28%, and it was 36% for the within-study variance. In the NOT sample, feedback type was a significant moderator. Specifically, studies utilizing CSTs (d = 0.36) were more effective than feedback systems using raw scores (d = 0.04) or ETRs (d = 0.12): $F(2, 40) = 5.08$, $p = .01$.

Results for Secondary Outcomes

Information on dropout was available for 39 studies, comprising 14,369 patients (48.9% controls; 51.1% feedback). Analyses showed differential dropout rates for feedback

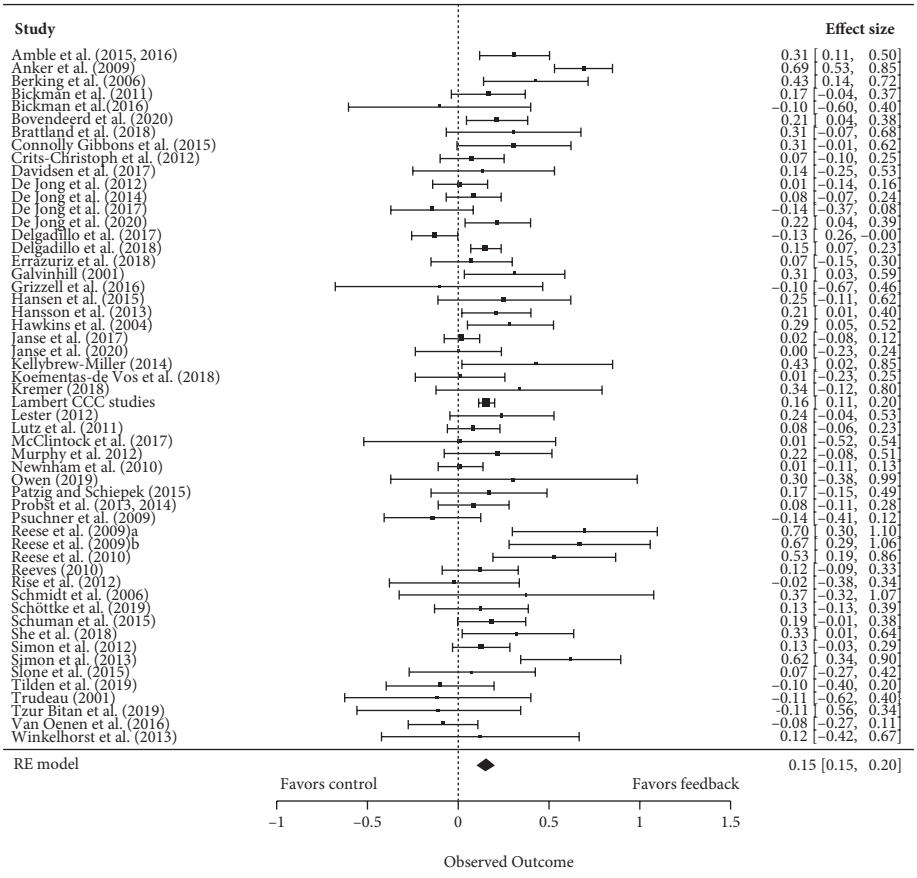

Figure 15.3 Effects of feedback in full sample (on-track and not-on-track combined). One extreme effect size was excluded from the analyses. Some studies by Lambert and colleagues used the same outcome data as an archival control group (Harmon et al., 2007; Lambert et al., 2001, 2002; Slade et al., 2008; Whipple et al., 2003). Therefore, the studies conducted in the Brigham Young University college counseling center were combined into one effect size, referred to as *Lambert CCC studies*.

groups of 20.9% compared to 24.5% for control groups, amounting to an approximately 20% greater chance of dropout in the absence of feedback (odds ratio [*OR*] = 1.19, 95% CI [1.03, 1.38], $p < .01$). Heterogeneity across studies was 47.8%. The effect was moderated by feedback instrument, with studies using the PCOMS (*OR* = 1.48) yielding significantly higher effects of feedback on reducing dropout than for the OQ System (*OR* = 1.21) and for studies using other feedback systems (*OR* = 1.08). Effect sizes were also larger for studies carried out in the United States (*OR* = 1.77) compared with outside the United States (*OR* = 1.07): $Q_M(1) = 8.99$, $p = .003$.

Data on the percentage of deteriorated cases was available for 26 studies, comprising 10,413 patients (46.4% control; 53.6% feedback), and analyses showed no significant differences between feedback and control conditions. The heterogeneity across studies was 7.1%. Moderating effects were found for training and type of feedback. Studies in

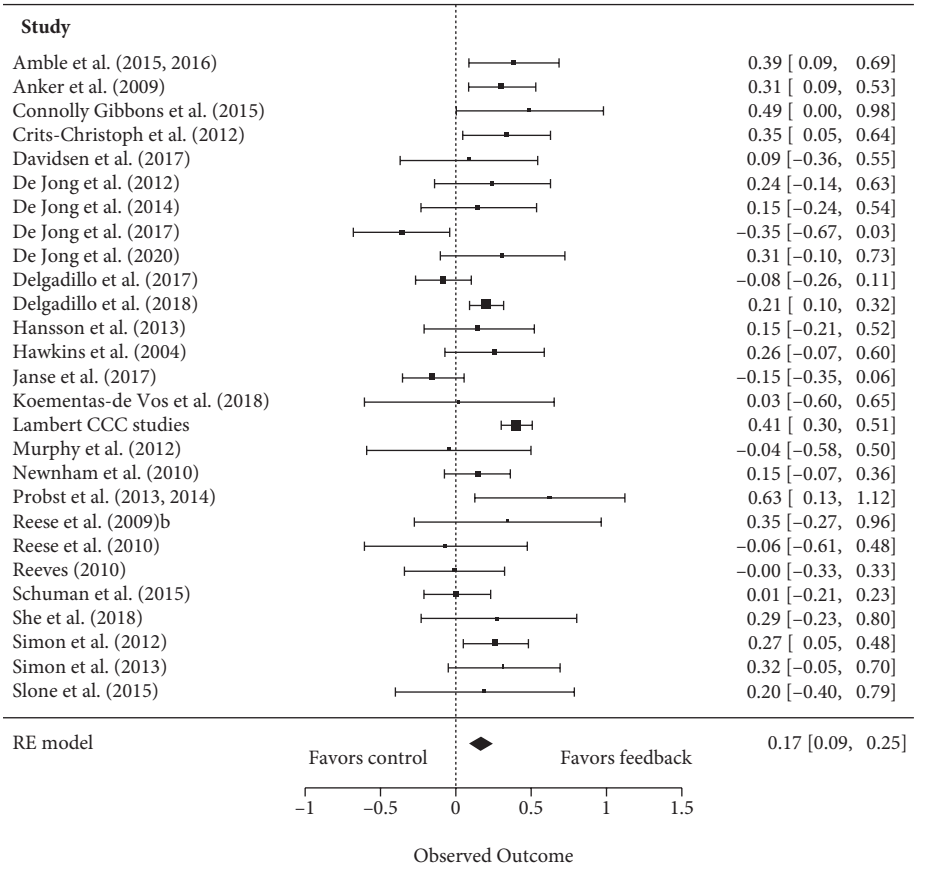

Figure 15.4 Effect of feedback in not-on-track subgroup.

which therapists received training in feedback yielded larger effects (*OR* = 1.28) than when no training was available (*OR* = 0.81), $Q_M = 5.45$, $p = .02$. Studies using ETRs in the feedback resulted in larger effects (*OR* = 1.36) than studies using CSTs (*OR* = 1.29) or raw scores (*OR* = 0.81), $Q_M = 6.70$, $p = .04$.

Feedback had no overall effect on the number of sessions conducted ($d = 0.04$, 95% CI [−0.06, 0.15]), a finding associated with high heterogeneity across studies (86.3%). No significant moderators of the effect of feedback on treatment duration were found.

Publication Bias

The results suggest that there was no publication bias regarding the results for the meta-analytic models. The *p* values were .55 (full sample) and .58 (NOT sample). In addition, the fail-safe *N* for the effect of feedback in the full sample was 3,695, meaning that an additional 3,695 studies would be required to overturn this result, suggesting that the effect of feedback is robust.

Different Feedback Systems

There were clear outcome differences between the two main feedback systems (OQ System and PCOMS) in the full sample. PCOMS yielded larger effect sizes in the full sample, but it had little effect in the NOT subgroup. By contrast, the OQ System seemed more effective with NOT cases, particularly in combination with CSTs. Because PCOMS did not show a differential effect benefitting NOT cases, it can be viewed as a more generic feedback system useful in OT cases, whereas the OQ System is more effective in NOT cases.

Such a result is consistent with the developmental rationale for each system. Namely, the PCOMS was constructed to be completed and discussed in session, thereby promoting better communication between patient and therapist. By contrast, the OQ System was premised on the basis of targeting patients who were at risk of failing in treatment. The PCOMS and its component outcome measure, the ORS, yielded larger effects, likely due to the latter being more sensitive to change than other outcome measures. Overall, the most parsimonious conclusion is that ROM was modestly helpful in general but not with all patients.

Cumulative and Causal Evidence

Locating the results of this meta-analysis in the broader evidence base provides some clarity on the impact of feedback. The effect sizes of psychological therapy with client feedback compared to psychological therapy without client feedback confirm the lower estimate ($g = 0.15$; De Jong et al., 2021) as previously reported in the literature ($g = 0.14$; Lambert et al., 2018). The largest effect obtained in the NOT sample for feedback enhanced by CSTs ($g = 0.36$; De Jong et al., 2021), an effect slightly lower than previously reported for NOT with CSTs ($g = 0.49$; Lambert et al., 2018), but similar to both the effect of PCOMS for all samples ($g = 0.40$) and NOT samples using the OQ-45 ($g = 0.33$; Lambert et al., 2018). These effects, we should remember, are *additive* to the effects of standard treatment (i.e., when there is no feedback).

Effect sizes regarding improvement for NOT patients with feedback versus TAU have been reported in seven meta-analyses (De Jong et al., 2021; Kendrick et al., 2016; Lambert et al., 2003, 2018; Østergård et al., 2020; Rognstad et al., 2022; Shimokawa et al., 2010). The effects range between $g = 0.17$ (De Jong et al., 2021) and $g = 0.53$ (Shimokawa et al., 2010). Three meta-analyses have assessed the effects of feedback with CSTs compared to no-feedback controls in NOT cases and have found effect sizes ranging from $g = 0.36$ to $g = 0.53$ (De Jong et al., 2021; Lambert et al., 2018; Shimokawa et al., 2010).

In addition to the effects of feedback on symptom reduction, meta-analyses have assessed the effects of feedback on the odds of treatment success, treatment duration, and the odds of dropout. Feedback increased the rate of improvement in NOT cases when CSTs were included ($OR = 2.40$ vs. $OR = 1.89$ without CSTs; Lambert et al., 2018). In terms of the effect of feedback on treatment duration, contrary to the findings of the present meta-analysis, differential effects have been reported between OT and NOT cases. OT cases have been reported to receive on average 0.69 sessions fewer when

feedback was provided, whereas NOT cases receive 0.73 sessions more when feedback was provided (Kendrick et al., 2016). Although these results concur with earlier findings (e.g., Lambert et al., 2003), results need further investigation. And regarding dropout, one salient finding from the meta-analysis of De Jong et al. (2021) is the effect of feedback on reducing dropout (recall by approximately 20%). Reducing dropout rates needs to be a major focus of effort for clinics and services.

Recall that these effects are *additive* to the effects of standard treatment with no feedback. The average effect size of feedback ($g = 0.15$) corresponds to a success rate difference (SRD; Furukawa & Leucht, 2011) of 8.45%. This effect size increases when clinical support tools for NOT cases are used to $g \approx 0.36–0.53$, which corresponds to an SRD between 20.09% and 29.22%.

Interpretation of effect sizes is crucial, particularly when between-group effect sizes (as cited in this chapter) are compared with effect sizes from correlational studies (Kraft, 2020). Between-group effect sizes reflect causal effects, whereas effects from correlational studies represent only descriptive relations between two variables. Correlational designs and deducted effect sizes are, on average, substantially larger than those from between-group designs and their respective effects sizes (Kraft, 2020).

In summary, patient feedback is a relatively simple method provided in addition to psychological treatment, where the effects compared to a range of control conditions are moderate to large (e.g., Barkham & Lambert, 2021). In other words, feedback offers value above and beyond the general effectiveness of psychotherapy. Overall, client feedback is effective in many settings, primarily with patients at risk of deterioration and with the use of CSTs. However, effect sizes depend on the comparison group and the feedback method used.

Research Review on Components of ROM

We have complemented the meta-analysis with evidence drawn from a selective narrative review of studies addressing components in each of the three ROM phases. Evidence relating to the first phase of ROM (data collection) indicates that patients generally support monitoring outcomes during therapy (see reanalysis of data from Lutz et al. [2011] as reported in Castonguay et al. [2013]). Reports show that patients prefer using a short measure to monitor progress if the therapist believes it to be useful (Thew et al., 2015; Zimmerman & McGlinchey, 2008). Yet, qualitative data have also revealed patients' doubts about completing measures when their providers either did not look at or did not make use of the information (Talib et al., 2018).

A systematic review comprising 26 studies noted four meta-themes identified by patients (Solstad et al., 2019): concern about motives for adopting ROM (i.e., suspicion that ROM was being used to determine service effectiveness rather than used for patients' benefit); dominance of symptom focus (i.e., the need to broaden out to include other domains of experience such as social functioning that better reflect the complexity of patients' lives); the need to provide a rationale, engaging with patients, and explaining how the data will be used (Börjesson & Boström, 2020); and developing a collaborative practice in which ROM becomes a clinical process tool to direct and deepen the therapeutic dialogue (Faija et al., 2022). Detailed analysis of a small sample

of patients using the NORSE feedback system has shown that although ROM enhances their awareness of emotions and experiences (see also Greenhalgh et al., 2018), patients can also be uncertain about whether the assignment of a specific rating on a scale actually captures their felt experience (Solstad et al., 2021). However, there was an overall view that the use of ROM was an enabling process that helped direct the focus of therapeutic dialogue.

Research into the frequency of data completion suggests no significant difference between continuous feedback (i.e., every session) and less frequent feedback (De Jong et al., 2021). A comparison between a basic (e.g., every fifth session) and a high-intensive (every session) delivery of feedback showed no difference in outcomes, but the latter resulted in fewer therapy sessions and a lower patient dropout rate (Janse et al., 2020). Prediction modeling of patient outcomes and dropout have been used on basic ROM data collected at intervals of between five and 15 sessions to show that the use of more sophisticated analyses is not dependent on session-by-session data (Mütze et al., 2022). Although such a result provides support for a less intense program of obtaining ROM data, securing useful data on the course of treatment was only viable due to the longer term duration of treatments, in which the mean number of sessions was 39 (range: 5–95). Continuous data collection also facilitates clinical decision-making, by increasing the odds of identifying early that a patient is not progressing well.

Obtaining baseline data is crucial therapeutically but also because most computerized feedback systems generate interpretation rules, such as expected treatment response curves or early response signals, based on the pretreatment baseline measure (Lutz et al., 2002). The limitations of relying on fixed predictions based solely on an initial assessment are being addressed by moves toward more dynamic modeling in which predictions are continually updated and have been shown to exceed the performance of previous methods (e.g., Bone et al., 2021).

Research focusing on the second phase (feeding back data) has found that although completion of self-report measures can result in greater self-learning by some patients, the impact on outcomes and treatment likely depends on how the information is subsequently shared and used in therapy (Greenhalgh et al., 2018). Data need to be reviewed in the form of a conversation rather than simply being told the numbers that can be viewed within the context of shared decision-making. An RCT showed no evidence that ROM enhanced shared decision-making overall across a range of patient presenting problems (Metz et al., 2019). However, it was associated with better outcomes for patients experiencing mood disorders, with the suggestion that the feedback was not sufficiently targeted to the other presenting conditions.

There is greater benefit in viewing ROM data collaboratively between patient and therapist (Hepner et al., 2019) rather than feeding back directly to patients (e.g., Slade et al., 2008). Research suggests both patients and therapists have similar preferences for how feedback is presented, preferring greater specificity in relation to the nature of predictors as well as of advice (Hilhorst et al., 2022). Both have preferences for feedback to be presented as either a continuous outcome or an outcome that is expressed in terms of a probability, with the feedback representation comprising both text and images. More generally, in terms of technical support to aid ROM and feedback, evidence from a systematic review shows the use of communication technologies (e.g.,

smartphones, tablets) to be practical and feasible in psychological therapies (Gual-Montolio et al., 2020).

Evidence supports the use of CSTs to augment feedback. In studies focusing on NOT patients in which a feedback arm was compared with one accessing CSTs, results have consistently favored CSTs ($d = 0.36$) compared with expected treatment response ($d = 0.12$) or use of raw scores ($d = 0.04$; De Jong et al., 2021). Results from other meta-analyses have yielded slightly larger effects for CSTs (e.g., 0.49; Lambert et al., 2018).

One key CST is the ASC (Lambert et al., 2015). In a study of 107 off-track outpatients, 58% had sufficient problems detected by one of the ASC scales to trigger an alarm, with 29% of all NOT patients meeting the criterion level for an alarm in the area of (lack of) social support (White et al., 2015). Lack of social support was more often identified as a flag for deterioration than were problems in the therapeutic alliance. Indeed, the dearth of social support as the single best predictor of patient deterioration has been replicated (Probst et al., 2020). In contrast to the prominence of external factors to therapy, scales relating to internal therapy processes (therapeutic alliance) either did not predict deterioration or, for the motivation scale, did so but inconsistently. In another study, an analysis comparing 273 OT and 143 NOT cases showed suicidality, motivation, and life events to be better predictors of subsequent deterioration in NOT cases (Schilling et al., 2021). These findings relating to the role of life events and the lack of social support in the lives of patients are a salutary reminder of the importance of the social and interpersonal world of patients outside of therapy and contrast with the considerable attention paid to in-session concepts, such as the therapeutic alliance.

In the third phase (adapting), evidence of adapting the focus or direction of therapy has been shown from analyses of patient–therapist dyads, with specific examples within the course of therapy and across services (Brooks Holliday et al., 2021). The within-session adjustments included the following: setting and monitoring treatment goals; determining the most appropriate therapeutic approach for a given patient; adjusting the pace of therapy; focusing the nature of the discussion during a given session and/or assigning treatment "homework"; or adjusting therapeutic modalities. Adjustments focusing on specific service parameters included making referrals to additional forms of care; making transition, terminations, or discharge decisions; facilitating communication with other clinicians and joint decision-making; and initiating/adjusting medication. Overall, a set of best practices for discussing feedback proposed providing a strong rationale for ROM, discussing ROM every time measures are administered; actively engaging patients in the discussion of ROM; and using a graph to show progress to patients.

Moderators

A number of potential moderators of ROM's effects have been investigated, but with little consistency across studies, and with many studies being underpowered, the research yield has been limited. However, the use of clinical support tools enhances the impact of feedback (Lutz, Deisenhofer, et al., 2022), particularly in NOT cases (De Jong et al., 2021). A number of studies have also shown therapist effects to moderate feedback effects (e.g., Bovendeerd et al., 2022; Janse et al., 2020). A reanalysis of six earlier

ROM studies using the OQ-45 found that the provision of feedback reduced the size of the therapist effect, thereby leveling the variability between more and less effective therapists (Delgadillo et al., 2022). Therapist positive attitude to feedback and using specific modifications in light of feedback have been found to be associated with enhanced effects (Lutz et al., 2015), whereas therapist-rated usefulness of feedback has been found to be a significant moderator of feedback outcome associations (Lutz, Deisenhofer, et al., 2022).

The direct study of therapist effects and ROM has found female therapists and those with a greater commitment to feedback to show higher probability of using the information provided by a feedback system and therapists who used the feedback system also to be more effective for NOT patients (De Jong et al., 2012). Specifically, therapists with a low internal feedback propensity who were more committed to using the feedback at the beginning of the study saw patients who improved more quickly. By contrast, therapists with a high internal feedback propensity (i.e., were more likely to trust their own opinion than that of feedback from external sources) saw patients with a slower rate of change (De Jong et al., 2012).

Implementation has been shown to impact ROM effects (Bovendeerd et al., 2022; Simon et al., 2013; van Sonsbeek et al., 2021). Several multicenter studies found differential effects of feedback within their trial, with some locations showing medium effects and other locations showing no effect at all (e.g., Bovendeerd et al., 2022). The effects of ROM have been found to increase over time, with one study showing post-treatment scores for patients recruited at the end of a study to have an effect size 0.467 greater than the post-treatment score of patients who started treatment at the commencement of the trial (Brattland et al., 2018). Although no specific factors were investigated to account for the improvement, monthly ROM meetings and biannual 1-day workshops were a feature of the implementation. Such a finding underscores the centrality of investing in resources to support high-quality ROM implementation.

Implementation Science

Although the literature on the benefits of ROM has yielded both positive and null findings, it is almost unanimous in identifying implementation as the main barrier for successful ROM. Obstacles to implementation of outcome measurement in routine practice have long been documented in the literature (e.g., Marks, 1998). These barriers can be grouped into three main categories: (a) the people (i.e., patient or therapist), (b) organizational aspects, and (c) systems (Lewis et al., 2019; Van Wert et al., 2021). Unlike many, or most, therapeutic methods, effective use of ROM is dependent on a climate of organizational support. Indeed, the aspiration has been referred to as developing a "culture for feedback" (Bertolino & Miller, 2012). In addition, a distinction has been made between practice issues and attitudinal issues, with cultural and philosophical issues being potentially greater obstacles than practical issues to successful implementation (Boswell, Kraus, Miller, et al., 2015; Boyce et al., 2014).

Commonly reported barriers include ease of accessing and using ROM systems, guidance on selecting outcome measures, and organizational accountability (Van Wert et al., 2021). A systematic review of ROM implementation identified a wide range of

factors comprising leadership, interorganizational factors, feedback culture, implementation team, coordinators and champions, supervision, training, measures, and generating a language for ROM use in clinical practice (Mackrill & Sørensen, 2020).

Attitudinal barriers to ROM have been explicitly investigated using the Evidence-based Practice Attitude Scale–Routine Outcome Monitoring (Rye et al., 2019). These studies showed that holding more positive attitudes regarding the adoption of ROM predicted greater use of standardized instruments. Limitations centered on ROM being seen as too narrowly focused, not suitable for patients presenting with multiple problems, and hindering the relationship between patient and therapist. Such concerns predicted poorer uptake of standardized measures (Rye et al., 2019). Therapists with a higher commitment to client feedback also had a higher probability to make use of the psychometric feedback system, and those therapists were more effective for NOT patients (De Jong et al., 2021). Similarly, therapists' satisfaction with the ROM system and the use of the feedback information tend to predict the magnitude of feedback effects (Lutz et al., 2015).

Strategies for improving implementation have been described by many authors (e.g., Bear et al., 2022; Lewis et al., 2019; Lutz et al., 2021; Mellor-Clark et al., 2016; Nilsen, 2015; Nilsen & Birken, 2020). Drawing from these sources and case examples, psychotherapists and organizations that plan to adopt ROM can consider the following actions: Nominate an implementation leader, supported by an implementation team; secure a costed plan for implementation, including technology, resources, and clinical time necessary to train therapists; share good practices in monitoring and feedback; automate the collection of data; and adopt an available feedback system that uses signaling technologies.

Possible Negative Effects and Harm

Evidence suggests two groups of patients might experience negative effects of ROM: patients with severe psychopathology and those experiencing cluster B personality disorders. As reported previously, patients experiencing greater severity experienced an aversive effect when in receipt of negative feedback in which clinical support tools were not available (Errázuriz & Zilcha-Mano, 2018). In this instance, the absence of CSTs may have been crucial, such that it is not severe psychopathology per se that is the issue but, rather, this factor combined with the absence of resources to support therapists with this specific patient group.

Adverse effects have also been reported both for NOT patients and for those presenting with cluster B personality disorders and also personality disorders not otherwise specified (De Jong et al., 2018). The patient sample comprised day patients and inpatients aged between 17 and 36 years, all diagnosed with severe personality disorders and receiving a long-term intensive psychotherapeutic program. The study compared a no-feedback control group with a group in which the therapist team received feedback and a group in which both the therapist team and the client received feedback. There was a significant difference in treatment course in the first 6 months of treatment, with the group in which patients and therapists received feedback showing higher levels of symptom distress than the other two groups. Crucially, this result seemed to be a

function of NOT patients who met criteria for a cluster B personality disorder or PD-NOS, but it did not apply to patients presenting with cluster C personality disorder. However, after this initial increase, the symptoms in this group decreased and were no longer significantly different from the other two groups by 39 weeks.

These collective findings confirm that feedback arising from ROM is not a panacea for enhancing treatment and may not be best suited to people who are more vulnerable psychologically. It may be the combination of discouraging feedback and greater severity or vulnerability makes negative feedback an unwelcome component, particularly in situations in which there are not the resources to support delivery of CSTs or where the flexibility of the therapist may be limited (e.g., in inpatient settings). Notwithstanding these cautionary notes, we are not aware of any reports documenting data on patients being harmed by ROM.

Diversity Considerations

Diversity, and in particular an intersectional approach, has been poorly covered in ROM and feedback literature. Diverse populations tend to be underrepresented in studies, and the reporting of socioeconomic and ethnic minority variables is rare. In addition, only two studies have been conducted in non-Western countries (Errázuriz & Zilcha-Mano, 2018 [Chile]; She et al, 2018 [China]). The study in Chile was presented earlier (see the section on Landmark Studies). In a study carried out in a Chinese college counseling center (N = 157), feedback based on the PCOMS was found to have a significant positive effect on treatment outcomes both in the full sample and in the NOT students. The latter group had six times the rate of reliable change in the feedback group than in the no-feedback condition (She et al., 2018). Importantly, the administration and use of the PCOMS was culturally adapted such that counselors did not administer the measure and did not have to make use of it in the session (Sun et al., 2021).

Limitations of the Research

In addition to the predominantly Western samples and minimal research extending beyond that cultural space, the major limitations concern poor implementation and lack of statistical power. Null findings highlight research dilemmas in these two specific areas (e.g., van Sonsbeek et al., 2021), along with complex designs (e.g., Errázuriz & Zilcha-Mano, 2018), mismatch between feedback system and the patient population (e.g., van Oenen et al., 2016), and infrequent application of feedback (e.g., Schöttke et al., 2020).

Regarding statistical power, only 20% of the studies in the main meta-analysis reported here (De Jong et al., 2021) comprised a total patient sample greater than 500. Power calculations consistently fail to account for therapists despite evidence attesting to their effect in routine practice (Baldwin & Imel, 2013; Johns et al., 2019). Adequate power is also essential to advance our understanding of potential moderators, which, along with a greater focus of mechanisms and theoretical models (e.g., Sapyta et al., 2005), is urgently needed.

A portion of research is closely aligned (e.g., via co-authorship) with measure developers. Reliance on highly selected samples of committed practitioners, sometimes with leading international experts as advisors, lessens the generalizability of such studies. Independence of authorship and ROM systems as well as adequately powered studies of clinical as opposed to student populations are required together with follow-up data. In the reported meta-analysis, only four studies included follow-up data greater than 1 month, and the maximum was at 6 months (De Jong et al., 2021). There is a need for follow-up data to be collected at upwards of 12 months and preferably longer. However, practical and possible ethical issues of clinical responsibility arise when extending the concept of routine monitoring beyond the time at which direct care is provided.

All the routine outcome measures reported in the current review are based on nomothetic principles. Some practitioners would prefer using idiographic outcome measures (Jensen-Doss et al., 2018) accompanied by a growing evidence base regarding the psychometric standing of idiographic measures, which can be either problem focused or goal focused (Cooper & Xu, 2023; Sales et al., 2023). However, adopting an experimental design to determine preferences between idiographic and nomothetic approaches found no difference (Bugatti & Boswell, 2022). A balance between standardization and personalization in ROM may yield better rates of adoption by practitioners (Bjaastad et al., 2019).

Finally, outcome measures used in ROM assume structural invariance over time; that is, the interpretation of items remains the same over time. Analyses of several outcome measures have yielded differing results, with longitudinal invariance being found in some measures and not in others (e.g., Coleman et al., 2022; Rosenström et al., 2022).

Training Implications

Training in ROM has been examined as a potential moderator but has not been found to significantly impact feedback effects on symptom reduction or dropout, although it has resulted in decreasing the percentage of deteriorated cases (De Jong et al., 2021). However, training in feedback was highly varied between studies, ranging from a 2-hour introduction to a full day of training with added booster sessions and monthly supervision. In other studies, results have shown that positive attitudes toward ROM (Edbrooke-Childs et al., 2016) and rates of ROM adoption (Persons et al., 2016) can be both enhanced following planful and well-designed training. A notable finding of the latter study was that at 12-month follow-up, practitioners were still adhering to using ROM in 57% of sessions as compared with 40% initially. But, importantly, these data related to the use of any progress monitoring measure, thereby generalizing from the specific measure used in the training program.

Findings from a survey of U.S. psychology training clinics found the top two ranked reasons for using ROM were to help trainees determine when there was a need to adapt treatment and to help them make better treatment decisions (Peterson & Fagan, 2017). And the two top-ranked reasons why ROM was not adopted in training clinics were lack of resources (e.g., personnel, money) and hesitancy by supervisors to adopt measures with which they were unfamiliar.

In terms of learning ROM, whereas passive learning (e.g., reading, workshops) increases knowledge, strategies that include behavioral rehearsal and modeling of practical actions are likely to enhance the delivery of ROM (Beidas et al., 2014). These include components that are primarily interactions with patients (e.g., explaining the rationale, responding to patients' reactions to data, and addressing adaptations to treatment). The adoption of deliberate practice is likely to support a more active and method-based approach to learning key components of ROM (Rousmaniere, 2017).

In addition, actions for improving the adoption of ROM include training the trainers as well as trainees. ROM as a tool in supervision has been espoused, essentially emphasizing the view that supervisors and the supervision process are integral to securing the adoption of ROM by trainees (Swift et al., 2015). Accounts of differing training programs premised on ROM have been reported in the literature—for example, in the United States (e.g., Cooper et al., 2021) and Germany (e.g., Lutz et al., 2023).

To achieve successful implementation requires collaborative leadership and time (Goldberg et al., 2016). The time taken to successfully implement ROM depends on the target by which it is being evaluated and the level of resource available. And once achieved, it is about maintaining and sustaining the practice. Some ROM activities can be achieved in months, whereas others may take longer (e.g., Cooper et al., 2021; see also De Jong et al., 2023).

Therapeutic Practices

- Consider adopting ROM as an evidence-based addition to psychotherapy, regardless of theoretical orientation. It provides transparency regarding patient outcomes; is viewed favorably by patients; and acts as a complement to clinician judgments of patient outcomes, particularly of deterioration.
- Remember that although the effect of ROM for all patients, on average, is small, it is additive to the general effectiveness of psychological therapies and slightly greater for patients who are NOT.
- Develop a positive attitude toward ROM because attitude is a significant factor in the success of ROM and cultivates a community of ROM champions. This supports a positive learning and implementation environment.
- Ensure that the outcome measure is both psychometrically sound and clinically practical. Therapists need to be familiar with the measure and its interpretation so that patients understand how a measure is scored, what it means, and how that information can be used to help them progress in therapy.
- Introduce ROM together with a clear rationale for its use to each patient early in the course of therapy, preferably in the first session so that the expectations are set out clearly and ROM is presented as an integral part of therapy and not an add-on.
- Engage patients in discussions about ROM with a particular emphasis on the data using teach-back methods to facilitate dialogue about the match or mismatches between the data and their psychological health.
- Take account of the clinical population and setting because ROM is not a panacea. Patients presenting with high levels of severity or cluster B personality disorders may not benefit due to the impact of repeated negative feedback. Certain clinical

settings may also limit the benefits of ROM where there is insufficient flexibility for the therapist to adapt the focus or direction of treatment.

- Administer outcome measures frequently. Although it is not essential to capture data at every session, there needs to be sufficient data density to recognize patterns of deterioration at the earliest opportunity so that treatment adaptations can be considered. Hence, if planned treatment durations are likely to be shorter, then more frequent administration of ROM is advisable (e.g., at each session).
- Frame ROM in-session activity in terms of three actionable phases: administering measures, feeding back results, and adapting therapy (where indicated). The first and second phases can apply to all patients, but the third applies only to those who are identified as NOT and when there is evidence of deterioration.
- Use a graph to show progress to patients because visualization enables them to have a better understanding of the overall concept of ROM as well as connect data with possible treatment options. Use a computer program that will generate expected treatment curves if available; if not, using graphs generated by Excel can be informative. The use of expected treatment curves also results in lower deterioration rates.
- Use ROM to adjust treatment when patients are NOT in terms of their predicted outcomes or when outcome measures are showing patient deterioration.
- Supplement ROM with clinical support tools, particularly for NOT patients, because these have consistently yielded the largest effects in research on ROM. These provide therapists with a clinical focus to address potential reasons for a lack of progress or deterioration.
- Implement training in ROM for therapists because this enhances reduction in patient deterioration rates. Active models of training involving role-plays and deliberate practice are likely to enhance the interactive components of ROM above and beyond passive knowledge acquisition. Provision of ROM training for the trainers (i.e., supervisors) is an important component.
- Provide sufficient time and resources to support the adoption of ROM. The time taken to implement ROM will depend on the specific stated aims of the project, the resources available to support it, and the compliance of all stakeholders involved.
- Consider culturally appropriate adaptations to the content or processes of ROM to ensure that it does not disadvantage culturally diverse populations.
- Identify and address obstacles to successful implementation of ROM before attempting to initiate it.

References

Ægisdóttir, S., White, M. J., Spengler, P. M., Maugherman, A. S., Anderson, L. A., Cook, R. S., Nichols, C. N., Lampropoulos, G. K., Walker, B. S., Cohen, G., & Rush, J. D. (2006). The Meta-Analysis of Clinical Judgment Project: Fifty-six years of accumulated research on clinical versus statistical prediction. *The Counseling Psychologist*, *34*(3), 341–382. https://doi.org/10.1177%2F0011000005285875

American Psychological Association, Presidential Task Force on Evidence-Based Practice. (2006). Evidence-based practice in psychology. *American Psychologist*, *61*(4), 271–285. https://doi.org/10.1037/0003-066X.61.4.271

Anker, M. G., Duncan, B. L., & Sparks, J. A. (2009). Using client feedback to improve couple therapy outcomes: A randomized clinical trial in a naturalistic setting. *Journal of Consulting and Clinical Psychology*, *77*(4), 693–704. https://doi.org/10.1037/a0016062

Ashworth, M., Shepherd, M., Christey, J., Matthews, V., Wright, K., Parmentier, H., Robinson, S., & Godfrey, E. (2004). A client-generated psychometric instrument: The development of "PSYCHLOPS." *Counselling & Psychotherapy Research*, *4*(2), 27–31. https://doi.org/10.1080/14733140412331383913

Baldwin, S. A., & Imel, Z. (2013). Therapist effects: Findings and methods. In M. J. Lambert (Ed.), *Bergin and Garfield's handbook of psychotherapy and behavior change* (6th ed., pp. 258–296). Wiley.

Barber, J., & Resnick, S. G. (2022). Collect, Share, Act: A transtheoretical clinical model for doing measurement-based care in mental health treatment. *Psychological Services*. Advance online publication. https://doi.org/10.1037/ser0000629

Barkham, M., & Lambert, M. J. (2021). The efficacy and effectiveness of psychological therapies. In M. Barkham, W. Lutz, & L. G. Castonguay (Eds.), *Bergin and Garfield's handbook of psychotherapy and behavior change* (7th ed., pp. 135–189). Wiley.

Barkham, M., Mellor-Clark, J., & Stiles, W. B. (2015). A CORE approach to progress monitoring and feedback: Enhancing evidence and improving practice. *Psychotherapy*, *52*(4), 402–411. https://doi.org/10.1037/pst0000030

Bear, H. A., Dalzell, K., Edbrooke-Childs, J., & Wolpert, M. (2022). Applying behaviour change theory to understand the barriers to implementing routine outcome monitoring. *British Journal of Clinical Psychology*, *61*(3), 557–578. https://doi.org/10.1111/bjc.12322

Beidas, R. S., Cross, W., & Dorsey, S. (2014). Show me, don't tell me: Behavioral rehearsal as a training and analogue fidelity tool. *Cognitive and Behavioral Practice*, *21*, 1–11. http://dx.doi.org/10.1016/j.cbpra.2013.04.002

Bergin, A. E. (1963). The effects of psychotherapy: Negative results revisited. *Journal of Counseling Psychology*, *10*, 224–250.

Bergman, H., Kornør, H., Nikolakopoulou, A., Hanssen-Bauer, K., Soares-Weiser, K., Tollefsen, T. K., & Bjørndal, A. (2018). Client feedback in psychological therapy for children and adolescents with mental health problems. *Cochrane Database of Systematic Reviews*, *8*, CD011729. https://doi.org/10.1002/14651858.CD011729.pub2

Berking, M., Orth, U., & Lutz, W. (2006). How effective is systematic feedback of treatment progress to the therapist? An empirical study in a cognitive–behavioural-oriented inpatient setting. *Zeitschrift fur Klinische Psychologie und Psychotherapie*, *35*, 21–29. https://doi.org/10.1026/1616-3443.35.1.21

Bertolino, B., & Miller, S. D. (2012). *ICCE manuals on feedback-informed treatment* (Vols. 1–6). ICCE Press.

Bickman, L. (2008). A measurement feedback system (MFS) is necessary to improve mental health outcomes. *Journal of the American Academy of Child and Adolescent Psychiatry*, *47*(10), 1114–1119. https://doi.org/10.1097/CHI.0b013e3181825af8

Bjaastad, J. F., Jensen-Doss, A., Moltu, C., Jakobsen, P., Hagenberg, H., & Joa, I. (2019). Attitudes toward standardized assessment tools and their use among clinicians in a public mental health service. *Nordic Journal of Psychiatry*, *73*(7), 387–396. https://doi.org/10.1080/08039488.2019.1642383

Bone, C., Simmonds-Buckley, M., Thwaites, R., Sandford, D., Merzhvynska, M., Rubel, J., Deisenhofer, A. K., Lutz, W., & Delgadillo, J. (2021). 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. https://doi.org/10.1016/S2589-7500(21)00018-2

Börjesson, S., & Boström, P. K. (2020). "I want to know what it is used for": Clients' perspectives on completing a routine outcome measure (ROM) while undergoing psychotherapy. *Psychotherapy Research*, *30*(3), 337–347. https://doi.org/10.1080/10503307.2019.1630780

Boswell, J. F., Hepner, K. A., Lysell, K., Rothrock, N. E., Bott, N., Childs, A. W., Douglas, S., Owings-Fonner, N., Wright, C. V., Stephens, K. A., Bard, D. E., Aajmain, S., & Bobbitt, B. L.

(2023). The need for a measurement-based care professional practice guideline. *Psychotherapy*, *60*(1), 1–16. https://doi.org/10.1037/pst0000439

Boswell, J. F., Kraus, D. R., Castonguay, L. G., & Youn, S. J. (2015). Treatment Outcome Package: Measuring and facilitating multidimensional change. *Psychotherapy*, *52*(4), 422–431. https://doi.org/10.1037/pst0000028

Boswell, J. F., Kraus, D. R., Miller, S. D., & Lambert, M. J. (2015). Implementing routine outcome monitoring in clinical practice: Benefits, challenges, and solutions. *Psychotherapy Research*, *25*(1), 6–19. https://doi.org/10.1080/10503307.2013.817696

Bovendeerd, B., De Jong, K., de Groot, E., Moerbeek, M., & de Keijser, J. (2022). Enhancing the effect of psychotherapy through systematic client feedback in outpatient mental healthcare: A cluster randomized trial. *Psychotherapy Research*, *32*(6), 710–722. https://doi.org/10.1080/10503307.2021.2015637

Boyce, M. B., Browne, J. P., & Greenhalgh, J. (2014). The experiences of professionals with using information from patient-reported outcome measures to improve the quality of healthcare: A systematic review of qualitative research. *BMJ Quality and Safety*, *23*(6), 508–518. https://doi.org/10.1136/bmjqs-2013-002524

Brattland, H., Koksvik, J. M., Burkeland, O., Gråwe, R. W., Klöckner, C., Linaker, O. M., Ryum, T., Wampold, B., Lara-Cabrera, M. L., & Iversen, V. C. (2018). The effects of routine outcome monitoring (ROM) on therapy outcomes in the course of an implementation process: A randomized clinical trial. *Journal of Counseling Psychology*, *65*(5), 641–652. https://doi.org/10.1037/cou0000286

Brooks Holliday, S., Hepner, K. A., Farmer, C. M., Mahmud, A., Kimerling, R., Smith, B. N., & Rosen, C. (2021). Discussing measurement-based care with patients: An analysis of clinician–patient dyads. *Psychotherapy Research*, *31*(2), 211–223. https://doi.org/10.1080/10503307.2020.1776413

Brown, G. S., Simon, A., Cameron, J., & Minami, T. (2015). A collaborative outcome resource network (ACORN): Tools for increasing the value of psychotherapy. *Psychotherapy*, *52*(4), 412–421. https://doi.org/10.1037/pst0000033

Bugatti, M., & Boswell, J. F. (2022). Clinician perceptions of nomothetic and individualized patient-reported outcome measures in measurement-based care. *Psychotherapy Research*, *32*(7), 898–909. https://doi.org/10.1080/10503307.2022.2030497

Burgess, P., Pirkis, J., & Coombs T. (2015). Routine outcome measurement in Australia. *International Review of Psychiatry*, *27*(4), 264–275. https://doi.org/10.3109/09540261.2014.977234

Carlier, I. V., Meuldijk, D., Van Vliet, I. M., Van Fenema, E., Van der Wee, N. J., & Zitman, F. G. (2012). Routine outcome monitoring and feedback on physical or mental health status: Evidence and theory. *Journal of Evaluation in Clinical Practice*, *18*, 104–110. https://doi.org/10.1111/j.1365-2753.2010.01543.x

Castonguay, L. G., Barkham, M., Lutz, W., & McAleavy, A. (2013). Practice oriented research: Approaches and applications. In M. J. Lambert (Ed.), *Bergin and Garfield's handbook of psychotherapy and behavior change* (6th ed., pp. 85–133). Wiley.

Cheung, M. W.-L. (2014). Modeling dependent effect sizes with three-level meta-analyses: A structural equation modeling approach. *Psychological Methods*, *19*(2), 211–229. https://doi.org/10.1037/a0032968

Clark, D. M. (2018). Realizing the mass public benefit of evidence-based psychological therapies: The IAPT program. *Annual Review of Clinical Psychology*, *14*, 159–183. https://doi.org/10.1146/annurev-clinpsy-050817-084833

Coleman, J. J., Dong, Y., Dumas, D., Owen, J., & Kopta, M. (2022). Longitudinal measurement invariance of the Behavioral Health Measure in a clinical sample. *Journal of Counseling Psychology*, *69*(1), 100–110. https://doi.org/10.1037/cou0000524

Cooper, L. D., Murphy, H. G., Delk, L. A., Fraire, M. G., Van Kirk, N., Sullivan, C. P., Waldron, J. C., Halliburton, A. E., Schiefelbein, F., & Gatto, A. (2021). Implementing routine outcome monitoring in a psychology training clinic: A case study of a process model. *Training and Education in Professional Psychology*, *15*(2), 87–96. https://doi.org/10.1037/tep0000298

Cooper, M., & Xu, D. (2023). Goals Form: Reliability, validity, and clinical utility of an idiographic goal-focused measure for routine outcome monitoring in psychotherapy. *Journal of Clinical Psychology*, *79*(3), 641–666. https://doi.org/10.1002/jclp.23344

Crits-Christoph, P., Ring-Kurtz, S., Hamilton, J. L., Lambert, M. J., Gallop, R., McClure, B., Kulaga, A., & Rotrosen, J. (2012). A preliminary study of the effects of individual patient-level feedback in outpatient substance abuse treatment programs. *Journal of Substance Abuse Treatment*, *42*(3), 301–309. https://doi.org/10.1016/j.jsat.2011.09.003

Cross, S., Mellor-Clark, J., & Macdonald, J. (2015). Tracking responses to items in measures as a means of increasing therapeutic engagement in clients: A complementary clinical approach to tracking outcomes. *Clinical Psychology & Psychotherapy*, *22*(6), 698–707. https://doi.org/10.1002/cpp.1929

Davidsen, A. H., Poulsen, S., Lindschou, J., Winkel, P., Tróndarson, M. F., Waaddegaard, M., & Lau, M. (2017). Feedback in group psychotherapy for eating disorders: A randomized clinical trial. *Journal of Consulting and Clinical Psychology*, *85*(5), 484–494. https://doi.org/10.1037/ccp0000173

Davidson, K., Perry, A., & Bell, L. (2014). Would continuous feedback of patient's clinical outcomes to practitioners improve NHS psychological therapy services? Critical analysis and assessment of quality of existing studies. *Psychology and Psychotherapy: Theory, Research and Practice*, *88*(1), 21–37. https://doi.org/10.1111/papt.12032

De Jong, K., & Aafjes-van Doorn, K. (2022). Routine outcome monitoring: The need for case examples. *Journal of Clinical Psychology*, *78*(10), 1963–1972. https://doi.org/10.1002/jclp.23441

De Jong, K., Conijn, J. M., Gallagher, R., Reshetnikova, A. S., Heij, M., & Lutz, M. C. (2021). Using progress feedback to improve outcomes and reduce drop-out, treatment duration, and deterioration: A multilevel meta-analysis. *Clinical Psychology Review*, *85*, 102002. https://doi.org/10.1016/j.cpr.2021.102002

De Jong, K., Delgadillo, J., & Barkham, M. (2023). *Routine outcome monitoring and feedback in psychological therapies*. Open University Press.

De Jong, K., Segaar, J., Ingenhoven, T., van Busschbach, J., & Timman, R. (2018). Adverse effects of outcome monitoring feedback in patients with personality disorders: A randomized controlled trial in day treatment and inpatient settings. *Journal of Personality Disorders*, *32*(3). https://doi.org/10.1521/pedi_2017_31_297

De Jong, K., van Sluis, P., Nugter, M. A., Heiser, W. J., & Spinhoven, P. (2012). Understanding the differential impact of outcome monitoring: Therapist variables that moderate feedback effects in a randomized clinical trial. *Psychotherapy Research*, *22*(4), 464–474. https://doi.org/10.1080/10503307.2012.673023

Delespaul, P. A. E. G. (2015). Routine outcome measurement in the Netherlands—A focus on benchmarking. *International Review of Psychiatry*, *27*(4), 320–328. https://doi.org/10.3109/09540261.2015.1045408

Delgadillo, J., Deisenhofer, A. K., Probst, T., Shimokawa, K., Lambert, M. J., & Kleinstäuber, M. (2022). Progress feedback narrows the gap between more and less effective therapists: A therapist effects meta-analysis of clinical trials. *Journal of Consulting and Clinical Psychology*, *90*(7), 559–567. https://doi.org/10.1037/ccp0000747

Delgadillo, J., De Jong, K., Lucock, M., Lutz, W., Rubel, J., Gilbody, S., Ali, S., Aguirre, E., Appleton, M., Nevin, J., O'Hayon, H., Patel, U., Sainty, A., Spencer, P., & McMillan, D. (2018). Feedback-informed treatment versus usual psychological treatment for depression and anxiety: A multisite, open-label, cluster randomised controlled trial. *Lancet Psychiatry*, *5*(7), 564–572. https://doi.org/10.1016/S2215-0366(18)30162-7

Delgadillo, J., McMillan, D., Gilbody, S., De Jong, K., Lucock, M., Lutz, W., Rubel, J., Aguirre, E., & Ali, S. (2021). Cost-effectiveness of feedback-informed psychological treatment: Evidence from the IAPT-FIT trial. *Behaviour Research and Therapy*, *142*, 103873. https://doi.org/10.1016/j.brat.2021.103873

Drapeau, M. (2012). Ten tools for progress monitoring in psychotherapy. *Integrating Science and Practice*, *2*, 1–45. https://www.mcgill.ca/psy/files/psy/10_tools_pdf_final.pdf

Drew, P., Irvine, A., Barkham, M., Faija, C., Gellatly, J., Ardern, K., Armitage, J. C., Brooks, H., Rushton, K., Welsh, C., Bower, P., & Bee, P. (2021). Telephone delivery of psychological interventions: Balancing protocol with patient-centred care. *Social Science & Medicine*, *277*, 113818. https://doi.org/10.1016/j.socscimed.2021.113818

Duncan, B. L., Miller, S. D., Sparks, J., Claud, D., Reynolds, L., Brown, J., & Johnson, L. (2003). The Session Rating Scale: Preliminary psychometric properties of a "working" alliance measure. *Journal of Brief Therapy*, *3*(1), 3–12.

Duncan, B. L., & Reese, R. J. (2015). The Partners for Change Outcome Management System (PCOMS): Revisiting the client's frame of reference. *Psychotherapy*, *52*, 391–401. https://doi.org/10.1037/pst0000026

Edbrooke-Childs, J., Wolpert, M., & Deighton, J. (2016). Using Patient Reported Outcome Measures to Improve Service Effectiveness (UPROMISE): Training clinicians to use outcome measures in child mental health. *Administration and Policy in Mental Health*, *43*(3), 302–308. https://doi.org/10.1007/s10488-014-0600-2

Egger, M., Smith, G. D., Schneider, M., & Minder, C. (1997). Bias in meta-analysis detected by a simple, graphical test. *BMJ*, *315*(7109), 629–634. https://doi.org/10.1136/bmj.315.7109.629

Emmelkamp, P. M., David, D., Beckers, T., Muris, P., Cuijpers, P., Lutz, W., Andersson, G., Araya, R., Banos Rivera, R. M., Barkham, M., Berking, M., Berger, T., Botella, C., Carlbring, P., Colom, F., Essau, C., Hermans, D., Hofmann, S. G., Knappe, S., . . . Vervliet, B. (2014). Advancing psychotherapy and evidence-based psychological interventions. *International Journal of Methods in Psychiatric Research*, *23*(Suppl. 1), 58–91. https://doi.org/10.1002/mpr.1411

Errázuriz, P., & Zilcha-Mano, S. (2018). In psychotherapy with severe patients discouraging news may be worse than no news: The impact of providing feedback to therapists on psychotherapy outcome, session attendance, and the alliance. *Journal of Consulting and Clinical Psychology*, *86*(2), 125–139. https://doi.org/10.1037/ccp0000277

Faija, C. L., Bee, P., Lovell, K., Lidbetter, N., Gellatly, J., Ardern, K., Rushton, K., Brooks, H., McMillan, D., Armitage, C. J., Woodhouse, R., & Barkham, M. (2022). Using routine outcome measures as clinical process tools: Maximising the therapeutic yield in the IAPT programme when working remotely. *Psychology and Psychotherapy: Theory, Research and Practice*, *95*(3), 820–837. https://doi.org/10.1111/papt.12400

Finch, A. E., Lambert, M. J., & Schaalje, B. G. (2001). Psychotherapy quality control: The statistical generation of expected recovery curves for integration into an early warning system. *Clinical Psychology and Psychotherapy*, *52*, 391–401. https://doi.org/10.1037/pst0000026

Furukawa, T. A., & Leucht, S. (2011). How to obtain NNT from Cohen's d: Comparison of two methods. *PLoS ONE*, *6*(4), e19070. https://doi.org/10.1371/journal.pone.0019070

Goldberg, S. B., Rousmaniere, T., Miller, S. D., Whipple, J., Nielsen, S. L., Hoyt, W. T., & Wampold, B. E. (2016). Do psychotherapists improve with time and experience? A longitudinal analysis of outcomes in a clinical setting. *Journal of Counseling Psychology*, *63*(1), 1–11. https://doi.org/10.1037/cou0000131

Gondek, D., Edbrooke-Childs, J., Fink, E., Deighton, J., & Wolpert, M. (2016). Feedback from outcome measures and treatment effectiveness, treatment efficiency, and collaborative practice: A systematic review. *Administration and Policy in Mental Health and Mental Health Services Research*, *43*(3), 325–343. https://doi.org/10.1007/s10488-015-0710-5

Greenhalgh, J., Gooding, K., Gibbons, E., Dalkin, S., Wright, J., Valderas, J., & Black, N. (2018). How do patient reported outcome measures (PROMs) support clinician–patient communication and patient care? A realist synthesis. *Journal of Patient-Reported Outcomes*, *2*, Article 42. https://doi.org/10.1186/s41687-018-0061-6

Gual-Montolio, P., Martínez-Borba, V., Bretón-López, J. M., Osma, J., & Suso-Ribera, C. (2020). How are information and communication technologies supporting routine outcome monitoring and measurement-based care in psychotherapy? A systematic review. *International Journal of Environmental Research and Public Health*, *17*(9), 3170. https://doi.org/10.3390/ijerph17093170

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

Harmon, C. S., Lambert, M. J., Smart, D. M., Hawkins, E., Nielsen, S. L., Slade, K., & Lutz, W. (2007). Enhancing outcome for potential treatment failures: Therapist–client feedback and clinical support tools. *Psychotherapy Research*, *17*(4), 379–392. https://doi.org/10.1080/10503300600702331

Hepner, K. A., Farmer, C. M., Brooks Holliday, S., Bharil, S., Kimerling, R. E., McGee-Vincent, P., McCaslin, S. E., & Rosen, C. (2019). *Displaying behavioral health measurement-based care data: Identifying key features from clinician and patient perspectives*. RAND. https://www.rand.org/pubs/research_reports/RR3078.html

Herold, D. M., Parsons, C. K., & Fedor, D. B. (1997). *Individual feedback propensities and their effects on motivation, training success and performance*. U.S. Army Research Institute for the Behavioral and Social Sciences.

Herold, D. M., Parsons, C. K., & Rensvold, R. B. (1996). Individual differences in the generation and processing of performance feedback. *Educational and Psychological Measurement*, *56*(1), 5–25.

Higgins, J. P. T., Savović, J., Page, M. J., Elbers, R. G., & Sterne, J. A. C. (2020). Assessing risk of bias in a randomized trial. In J. P. T. Higgins, J. Thomas, J. Chandler, M. Cumpston, T. Li, M. J. Page, & V. A. Welch (Eds.), *Cochrane handbook for systematic reviews of interventions* (version 6.1, updated September 2020). Cochrane. https://training.cochrane.org/handbook

Hilhorst, L., Stappen, J. V., Lokkerbol, J., Hiligsmann, M., Risseeuw, A. H., & Tiemens, B. G. (2022). Patients' and psychologists' preferences for feedback reports on expected mental health treatment outcomes: A discrete-choice experiment. *Administration and Policy in Mental Health*, *49*, 707–721. https://doi.org/10.1007/s10488-022-01194-2

Hofmann, S. G., & Kashdan, T. B. (2010). The Affective Style Questionnaire: Development and psychometric properties. *Journal of Psychopathology and Behavioral Assessment*, *32*(2), 255–263. https://doi.org/10.1007/s10862-009-9142-4

Hollenbeck, J. R., & Klein, H. J. (1987). Goal commitment and the goal-setting process: Problems, prospects, and proposals for future research. *Journal of Applied Psychology*, *72*, 212–220.

Howard, K. I., Moras, K., Brill, P. L., Martinovich, Z., & Lutz, W. (1996). Evaluation of psychotherapy: Efficacy, effectiveness, and patient progress. *American Psychologist*, *51*(10), 1059–1064. https://doi.org/10.1037//0003-066x.51.10.1059

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

Janse, P. D., De Jong, K., Veerkamp, C., van Dijk, M. K., Hutschemaekers, G., & Verbraak, M. (2020). The effect of feedback-informed cognitive behavioral therapy on treatment outcome: A randomized controlled trial. *Journal of Consulting and Clinical Psychology*, *88*(9), 818–828. https://doi.org/10.1037/ccp0000549

Jensen-Doss, A., Haimes, E. M. B., Smith, A. M., Lyon, A. R., Lewis, C. C., Stanick, C. F., & Hawley, K. M. (2018). Monitoring treatment progress and providing feedback is viewed favorably but rarely used in practice. *Administration and Policy in Mental Health and Mental Health Services Research*, *45*(1), 48–61. https://doi.org/10.1007/s10488-016-0763-0

Johns, R. G., Barkham, M., Kellett, S., & Saxon, D. (2019). A systematic review of therapist effects: A critical narrative update and refinement to review. *Clinical Psychology Review*, *67*, 78–93. https://doi.org/10.1016/j.cpr.2018.08.004

Keetharuth, A. D., Brazier, J, Connell, J., Bjorner, J. B., Carlton, J., Taylor Buck, E., Ricketts, T., McKendrick, K., Browne, J., Croudace, T., & Barkham, M., on behalf of the ReQoL Scientific Group. (2018). Recovering Quality of Life (ReQoL): A new generic self-reported outcome measure for use with people experiencing mental health difficulties. *British Journal of Psychiatry*, *212*, 42–49.

Kendrick, T., El-Gohary, M., Stuart, B., Gilbody, S., Churchill, R., Aiken, L., Bhattacharya, A., Gimson, A., Brütt, A. L., De Jong, K., & Moore, M. (2016). Routine use of patient reported outcome measures (PROMs) for improving treatment of common mental health disorders in adults. *Cochrane Database of Systematic Reviews*, *7*(7), CD011119. https://doi.org/10.1002/14651858.CD011119.pub2

Knapstad, M., Nordgreen, T., & Smith, O. R. (2018). Prompt mental health care, the Norwegian version of IAPT: Clinical outcomes and predictors of change in a multicenter cohort study. *BMC Psychiatry*, *18*(1), 1–16. https://doi.org/10.1186/s12888-018-1838-0

Knaup, C., Koesters, M., Schoefer, D., Becker, T., & Puschner, B. (2009). Effect of feedback of treatment outcome in specialist mental healthcare: Meta-analysis. *British Journal of Psychiatry*, *195*, 15–22. https://doi.org/10.1192/bjp.bp.108.053967

Kopta, M., Owen, J., & Budge, S. (2015). Measuring psychotherapy outcomes with the Behavioral Health Measure–20: Efficient and comprehensive. *Psychotherapy*, *52*(4), 442–448. https://doi.org/10.1037/pst0000035

Koslofsky, S., & Rodíguez, M. M. D. (2017). Cultural adaptations to psychotherapy: Real-world applications. *Clinical Case Studies*, *16*(1), 3–8. https://doi.org/10.1177/1534650116668273

Kraft, M. A. (2020). Interpreting effect sizes of education interventions. *Educational Researcher*, *49*(4), 241–253. htttps://doi.org/10.3102/0013189X20912798

Krägeloh, C. U., Czuba, K. J., Billington, D. R., Kersten, P., & Siegert, R. J. (2015). Using feedback from patient-reported outcome measures in mental health services: A scoping study and typology. *Psychiatric Services*, *66*(3), 224–241. https://doi.org/10.1176/appi.ps.201400141

Lambert, M. J. (2010). *Prevention of treatment failure: The use of measuring, monitoring, & feedback in clinical practice*. American Psychological Association Press.

Lambert, M. J., Bailey, R. J., White, M., Tingey, K. M., & Stevens, E. (2015). *Clinical support tool manual (Brief Version–40)*. OQMeasures.

Lambert, M. J., Kahler, M., Harmon, C., Burlingame, G. M. Shimokawa K., & White, M. M. (2013). *Administration and scoring manual: Outcome Questionnaire OQ-45.2*. OQMeasures.

Lambert, M. J., Whipple, J. L., Hawkins, E. J., Vermeersch, D. A., Nielsen, S. L., & Smart, D. W. (2003). Is it time for clinicians to routinely track patient outcome? A meta-analysis. *Clinical Psychology: Science and Practice*, *10*(3), 288–301. https://psycnet.apa.org/doi/10.1093/clipsy.bpg025

Lambert, M. J., Whipple, J. L., & Kleinstäuber, M. (2018). Collecting and delivering progress feedback: A meta-analysis of routine outcome monitoring. *Psychotherapy*, *55*(4), 520–537. https://doi.org/10.1037/pst0000167

Lambert, M. J., Whipple, J. L., & Kleinstäuber, M. (2019). Collecting and delivering client feedback. In J. C. Norcross & M. J. Lambert (Eds.), *Relationships that work* (3rd ed., pp. 580–630). Oxford University Press.

Lambert, M. J., Whipple, J. L., Smart, D. W., Vermeersch, D. A., Nielsen, S. L., & Hawkins, E. J. (2001). The effects of providing therapists with feedback on client progress during psychotherapy: Are outcomes enhanced? *Psychotherapy Research*, *11*, 49–68. https://doi.org/10.1093/ptr/11.1.49

Lambert, M. J., Whipple, J. L., Vermeersch, D. A., Smart, D. W., Hawkins, E. J., Nielsen, S. L., & Goates, M. (2002). Enhancing psychotherapy outcomes via providing feedback on client progress: A replication. *Clinical Psychology & Psychotherapy*, *9*(2), 91–103. https://doi.org/10.1002/cpp.324

Langkaas, T. F., Wampold, B. E., & Hoffart, A. (2018). Five types of clinical difference to monitor in practice. *Psychotherapy*, *55*(3), 241–254. https://doi.org/10.1037/pst0000194

Lewis, C. C., Boyd, M., Puspitasari, A., Navarro, E., Howard, J., Kassab, H., Hoffman, M., Scott, K., Lyon, A., Douglas, S., Simon, G., & Kroenke, K. (2019). Implementing measurement-based care in behavioral health: A review. *JAMA Psychiatry*, *76*(3), 324–335. https://doi.org/10.1001/jamapsychiatry.2018.3329

Lucock, M., Halstead, J., Leach, C., Barkham, M., Tucker, S., Randall, C., Lloyd, J., Khan, W., Catlow, H., Waters, E., & Saxon, D. (2015). A mixed method investigation of patient monitoring and feedback in routine practice: Barriers and facilitators. *Psychotherapy Research*, *25*(6), 633–646. https://doi.org/10.1080/10503307.2015.1051163

Lutz, W., Böhnke, J. R., & Köck, K. (2011). Lending an ear to feedback systems: Evaluation of recovery and non-response in psychotherapy in a German outpatient setting. *Community Mental Health Journal*, *47*(3), 311–317. https://doi.org/10.1007/s10597-010-9307-3

Lutz, W., Deisenhofer, A.-K., Rubel, J., Bennemann, B., Giesemann, J., Poster, K., & Schwartz, B. (2022). Prospective evaluation of a clinical decision support system in psychological therapy. *Journal of Consulting and Clinical Psychology*, *90*(1), 90–106. https://doi.org/10.1037/ccp0000642

Lutz, W., Deisenhofer, A.-K., Weinmann-Lutz, B., & Barkham, M. (2023). Data and feedback-informed clinical training and practice. In L. G. Castonguay & C. E. Hill (Eds.), *Becoming better therapists: Advancing training and supervision in psychotherapy*. American Psychological Association.

Lutz, W., De Jong, K., Rubel, J. A., & Delgadillo, J. (2021). Measuring, predicting, and tracking change in psychotherapy. In M. Barkham, W. Lutz, & L. G. Castonguay (Eds.), *Bergin and Garfield's handbook of psychotherapy and behavior change* (7th ed., pp. 89–133). Wiley.

Lutz, W., Leach, C., Barkham, M, Lucock, M., Stiles, W. B., Evans, C., Noble, R., & Iveson, S. (2005). Predicting change for individual psycho therapy clients on the basis of nearest neighbors. *Journal of Consulting and Clinical Psychology*, *73*(5), 904–913. https://doi.org/10/1037/0022-006X.73.5.904

Lutz, W., Martinovich, Z., & Howard, K. I. (1999). Patient profiling: An application of random coefficient regression models to depicting the response of a patient to outpatient psychotherapy. *Journal of Consulting and Clinical Psychology*, *67*(4), 571–577. https://doi.org/10.1037//0022-006x.67.4.571

Lutz, W., Martinovich, Z., Howard, K. I., & Leon, S. C. (2002). Outcomes management, expected treatment response, and severity-adjusted provider profiling in outpatient psychotherapy. *Journal of Clinical Psychology*, *58*(10), 1291–1304. https://doi.org/10.1002/jclp.10070

Lutz, W., Rubel, J., Deisenhofer, A. K., & Moggia, D. (2022). Continuous outcome measurement in modern data-informed psychotherapies. *World Psychiatry*, *21*(2), 215–216. https://doi.org/10.1002/wps.20988

Lutz, W., Rubel, J., Schiefele, A. K., Zimmermann, D., Böhnke, J. R., & Wittmann, W. W. (2015). Feedback and therapist effects in the context of treatment outcome and treatment length. *Psychotherapy Research*, *25*(6), 647–660. https://doi.org/10.1080/10503307.2015.1053553

Lutz, W., Rubel, J. A., Schwartz, B., Schilling, V., & Deisenhofer, A.-K. (2019). Towards integrating personalized feedback research into clinical practice: Development of the Trier Treatment Navigator (TTN). *Behaviour Research and Therapy*, *120*, 103438. https://doi.org/10.1016/j.brat.2019.103438

Lutz, W., Schwartz, B., & Delgadillo, J. (2022). Measurement-based and data-informed psychological therapy. *Annual Review of Clinical Psychology*, *18*, 71–98. https://doi.org/10.1146/annurev-clinpsy-071720-014821

Mackrill, T., & Sørensen, K. M. (2020) Implementing routine outcome measurement in psychosocial interventions—A systematic review. *European Journal of Social Work*, *23*(5), 790–808. https://doi.org/10.1080/13691457.2019.1602029

Marks, I. (1998). Overcoming obstacles to routine outcome measurement: The nuts and bolts of implementing clinical audit. *British Journal of Psychiatry*, *173*, 281–286. https://doi.org/10.1192/bjp.173.4.281

McAleavey, A. A., Nordberg, S. S., & Moltu, C. (2021). Initial quantitative development of the Norse Feedback system: A novel clinical feedback system for routine mental healthcare. *Quality of Life Research*, *30*(11), 3097–3115. https://doi.org/10.1007/s11136-021-02825-1

Meehl, P. E. (1954). *Clinical versus statistical prediction: A theoretical analysis and a review of the evidence*. University of Minnesota Press.

Mellor-Clark, J., Cross, S., Macdonald, J., & Skjulsvik, T. (2016). Leading horses to water: Lessons from a decade of helping psychological therapy services use routine outcome measurement to improve practice. *Administration and Policy in Mental Health*, *43*(3), 279–285. https://doi.org/10.1007/s10488-014-0587-8

Metz, M. J., Veerbeek, M. A., Twisk, J. W. R., van der Feltz-Cornelis, C. M., de Beurs, E., & Beekman, A. T. F. (2019). Shared decision-making in mental health care using routine outcome monitoring: Results of a cluster randomised-controlled trial. *Social Psychiatry and Psychiatric Epidemiology*, *54*(2), 209–219. https://doi.org/10.1007/s00127-018-1589-8

Miller, S. D., Duncan, B. L., Brown, J., Sparks, J., & Claud, D. (2003). The Outcome Rating Scale: A preliminary study of the reliability, validity, and feasibility of a brief visual analog measure. *Journal of Brief Therapy*, *2*, 91–100.

Miller, S. D., Hubble, M. A., Chow, D., & Seidel, J. (2015). Beyond measures and monitoring: Realizing the potential of feedback-informed treatment. *Psychotherapy*, *52*(4), 449–457. https://doi.org/10.1037/pst0000031

Moltu, C., McAleavey, A. A., Helleseth, M. M., Møller, G. H., & Nordberg, S. S. (2021). How therapists and patients need to develop a clinical feedback system after 18 months of use in a practice–research network: A qualitative study. *International Journal of Mental Health Systems*, *15*(1), Article 43. https://doi.org/10.1186/s13033-021-00465-z

Mütze, K., Witthöft, M., Lutz, W., & Bräscher, A. K. (2022). Matching research and practice: Prediction of individual patient progress and dropout risk for basic routine outcome monitoring. *Psychotherapy Research*, *32*(3), 358–371. https://doi.org/10.1080/10503307.2021.1930244

Nakagawa, S., & Santos, E. S. A. (2012). Methodological issues and advances in biological meta-analysis. *Evolutionary Ecology*, *26*(5), 1253–1274.

Nilsen, P. (2015). Making sense of implementation theories, models and frameworks. *Implementation Science*, *10*, Article 53. https://doi.org/10.1186/s13012-015-0242-0

Nilsen, P., & Birken, S. A. (Eds.). (2020). *Handbook on implementation science*. Elgar.

Nugter, M. A., De Jong, K., & Kortrijk, H. E. (2013). Meten van veranderingen bij patiënten. In V. J. A. Buwalda, M. A. Nugter, W. van Tilburg, & A. T. F. Beekman (Eds.), *Praktijkboek ROM in de ggz II: Implementatie en gebruik bij verschillende doelgroepen* (pp. 149–164). De Tijdstroom.

Ogles, B. M., Goates-Jones, M. K., & Erekson, D. M. (2022). Treatment success or failure? Using a narrative interview to supplement ROM. *Journal of Clinical Psychology*, *78*(10), 1986–2001. https://doi.org/10.1002/jclp.23345

Østergård, O. K., 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. *Psychotherapy Research*, *30*(2), 195–212. https://doi.org/10.1080/10503307.2018.1517949

Parker, Z. J., & Waller, G. (2015). Factors related to psychotherapists' self-assessment when treating anxiety and other disorders. *Behaviour Research and Therapy*, *66*, 1–7. https://doi.org/10.1016/j.brat.2014.12.010

Pejtersen, J. H., Viinholt, B. C. A., & Hansen, H. (2020). Feedback-informed treatment: A systematic review and meta-analysis of the Partners for Change Outcome Management System. *Journal of Counseling Psychology*, *67*(6), 723–735. https://doi.org/10.1037/cou0000420

Persons, J. B., Koerner, K., Eidelman, P., Thomas, C., & Liu, H. (2016). Increasing psychotherapists' adoption and implementation of the evidence-based practice of progress monitoring. *Behaviour Research and Therapy*, *76*, 24–31. https://doi.org/10.1016/j.brat.2015.11.004

Peterson, A. P., & Fagan, C. (2017). Training the next generation in routine outcome monitoring: Current practices in psychology training clinics. *Training and Education in Professional Psychology*, *11*(3), 182–189. https://doi.org/10.1037/tep0000148

Pinner, D. H., & Kivlighan, D. M., III. (2018). The ethical implications and utility of routine outcome monitoring in determining boundaries of competence in practice. *Professional Psychology: Research and Practice*, *49*(4), 247–254. https://doi.org/10.1037/pro0000203

Probst, T., Kleinstäuber, M., Lambert, M. J., Tritt, K., Pieh, C., Loew, T. H., Dahlbender, R. W., & Delgadillo, J. (2020). Why are some cases not on track? An item analysis of the Assessment for Signal Cases during inpatient psychotherapy. *Clinical Psychology & Psychotherapy*, *27*(4), 559–566. https://doi.org/10.1002/cpp.2441

Rognstad, K., Wentzel-Larsen, T., Neumer, P.-S., & Kjøbli, J. (2023). A systematic review and meta-analysis of measurement feedback systems in treatment for common mental health disorders. *Administration and Policy in Mental Health and Mental Health Services Research*, *50*, 269–282. https://doi.org/10.1007/s10488-022-01236-9

Rosenström, T. H., Mylläri, S., Malkki, V., & Saarni, S. E. (2022). Feasibility of generic, short, and easy-to-use assessment of psychological distress during psychotherapy: Longitudinal

measurement invariance of CORE-10 and -OM. *Psychotherapy Research*, *32*(8), 1090–1099. https://doi.org/10.1080/10503307.2022.2074807

Rousmaniere, T. (2017). *Deliberate practice for psychotherapists: A guide to improving clinical effectiveness*. Routledge.

Rye, M., Rognmo, K, Aarons, G. A., & Skre, I. (2019). Attitudes towards the use of routine outcome monitoring of psychological therapies among mental health providers: The EBPAS-ROM. *Administration and Policy in Mental Health*, *46*(6), 833–846. https://doi.org/10.1007/s10488-019-00968-5

Sales, C., Ashworth, M., Ayis, S., Barkham, M., Edbrooke-Childs, J., Faísca, L., Jacob, J., Xu, D., & Cooper, M. (2023). Idiographic patient reported outcome measures (I-PROMs) for routine outcome monitoring in psychological therapies: Position paper. *Journal of Clinical Psychology*, *79*(3), 596–621. https://doi.org/10.1002/jclp.23319

Sapyta, J., Riemer, M., & Bickman, L. (2005). Feedback to clinicians: Theory, research, and practice. *Journal of Clinical Psychology*, *61*(2), 145–153. https://doi.org/10.1002/jclp.20107

Schilling, V. N. L. S., Zimmermann, D., Rubel, J. A., Boyle, K. S., & Lutz, W. (2021). Why do patients go off track? Examining potential influencing factors for being at risk of psychotherapy treatment failure. *Quality of Life Research*, *30*, 3287–3298. https://doi.org/10.1007/s11136-020-02664-6

Schöttke, H., Unrath, M., & Uhlmann, C. (2020). The effect of patient progress feedback on psychotherapy outcome. *Verhaltenstherapie*, *30*(3). https://doi.org/10.1159/000503765

Schuman, D. L., Slone, N. C., Reese, R. J., & Duncan, B. (2015). Efficacy of client feedback in group psychotherapy with soldiers referred for substance abuse treatment. *Psychotherapy Research*, *25*(4), 396–407. https://doi.org/10.1080/10503307.2014.900875

She, Z., Duncan, B. L., Reese, R. J., Sun, Q., Shi, Y., Jiang, G., Wu, C., & Clements, A. L. (2018). Client feedback in China: A randomized clinical trial in a college counseling center. *Journal of Counseling Psychology*, *65*(6), 727–737. https://doi.org/10.1037/cou0000300

Shimokawa, K., Lambert, M. J., & Smart, D. W. (2010). Enhancing treatment outcome of patients at risk of treatment failure: Meta-analytic and mega-analytic review of a psychotherapy quality assurance system. *Journal of Consulting and Clinical Psychology*, *78*, 298–311. https://doi.org/10.1037/a0019247

Simon, W., Lambert, M. J., Busath, G., Vazquez, A., Berkeljon, A., Hyer, K., Granley, M., & Berrett, M. (2013). Effects of providing patient progress feedback and clinical support tools to psychotherapists in an inpatient eating disorders treatment program: A randomized controlled study. *Psychotherapy Research*, *23*(3), 287–300. https://doi.org/10.1080/10503307.2013.787497

Slade, K., Lambert, M. J., Harmon, S. C., Smart, D. W., & Bailey, R. (2008). Improving psychotherapy outcome: The use of immediate electronic feedback and revised clinical support tools. *Clinical Psychology & Psychotherapy*, *15*(5), 287–303. https://doi.org/10.1002/cpp.594

Slone, N. C., Reese, R. J., Mathews-Duvall, S., & Kodet, J. (2015). Evaluating the efficacy of client feedback in group psychotherapy. *Group Dynamics*, *19*, 122–136. https://doi.org/10.1037/gdn0000026

Solstad, S. M., Castonguay, L. G., & Moltu, C. (2019). Patients' experiences with routine outcome monitoring and clinical feedback systems: A systematic review and synthesis of qualitative empirical literature. *Psychotherapy Research*, *29*(2), 157–170. https://doi.org/10.1080/10503307.2017.1326645

Solstad, S. M., Kleiven, G. S., & Moltu, C. (2021). Complexity and potentials of clinical feedback in mental health: An in-depth study of patient processes. *Quality of Life Research*, *30*(11), 3117–3125. https://doi.org/10.1007/s11136-020-02550-1

Sterne, J. A., Hernán, M. A., Reeves, B. C., Savović, J., Berkman, N. D., Viswanathan, M., Henry, D., Altman, D. G., Ansari, M. T., Boutron, I., Carpenter, J. R., Chan, A. W., Churchill, R., Deeks, J. J., Hróbjartsson, A., Kirkham, J., Jüni, P., Loke, Y. K., Pigott, T. D., . . . Higgins, J. P. (2016). ROBINS-I: A tool for assessing risk of bias in non-randomised studies of interventions. *BMJ*, *355*, i4919. https://doi.org/10.1136/bmj.i4919

Sun, Q., Wu, C., Wang, C., Yu, L., She, Z., & Falkenström, F. (2021). Alliance–outcome relation and progress feedback: Secondary data analyses of a randomized clinical trial study in China. *Psychotherapy Research*, *31*(2), 145–156. https://doi.org/10.1080/10503307.2020.1772517

Swift, J. K., Callahan, J. L., Rousmaniere, T. G., Whipple, J. L., Dexter, K., & Wrape, E. R. (2015). Using client outcome monitoring as a tool for supervision. *Psychotherapy*, *52*(2), 180–184. https://doi.org/10.1037/a0037659

Talib, T. L., DeChant, P., Kean, J., Monahan, P. O., Haggstrom, D. A., Stout, M. E., & Kroenke, K. (2018). A qualitative study of patients' perceptions of the utility of patient-reported outcome measures of symptoms in primary care clinics. *Quality of Life Research*, *27*(12), 3157–3166. https://doi.org/10.1007/s11136-018-1968-3

Tam, H. E., & Ronan, K. (2017). The application of a feedback-informed approach in psychological service with youth: Systematic review and meta-analysis. *Clinical Psychology Review*, *55*, 41–55. https://doi.org/10.1016/j.cpr.2017.04.005

Tasca, G. A., Angus, L., Bonli, R., Drapeau, M., Fitzpatrick, M., Hunsley, J., & Knoll, M. (2019). Outcome and progress monitoring in psychotherapy: Report of a Canadian Psychological Association task force. *Canadian Psychology*, *60*(3), 165–177. https://doi.org/10.1037/0022-006X.73.5.904

The Joint Commission. (2018, February 6). *R_3 report: Requirement, rationale, reference*. https://www.jointcommission.org/-/media/tjc/documents/standards/r3-reports/r3_13_outcome_measures_1_30_18_final.pdf

Thew, G., Fountain, L., & Salkovskis, P. (2015). Service user and clinician perspectives on the use of outcome measures in psychological therapy. *The Cognitive Behaviour Therapist*, *8*, e23. https://doi.org/10.1017/S1754470X15000598

Tzur Bitan, D., Ganor, O., Biran, L., & Bloch, Y. (2018). Implementing routine outcome monitoring in public mental health services in Israel: Shared and unique challenges. *Journal of Evaluation in Clinical Practice*, *24*, 323–330. https://doi.org/10.1111/jep.12839

van Oenen, F. J., Schipper, S., Van, R., Schoevers, R., Visch, I., Peen, J., & Dekker, J. (2016). Feedback-informed treatment in emergency psychiatry: A randomised controlled trial. *BMC Psychiatry*, *16*(1), Article 110. https://doi.org/10.1186/s12888-016-0811-z

van Sonsbeek, M. A. M. S., Hutschemaekers, G. J. M., Veerman, J. W., Vermulst, A., Kleinjan, M., & Tiemens, B. G. (2021). Challenges in investigating the effective components of feedback from routine outcome monitoring (ROM) in youth mental health care. *Child & Youth Care Forum*, *50*(2), 307–332. https://doi.org/10.1007/s10566-020-09574-1

Van Wert, M. J., Malik, M., Memel, B., Moore, R., Buccino, D., Hackerman, F., Kumari, S., Everett, A., & Narrow W. (2021). Provider perceived barriers and facilitators to integrating routine outcome monitoring into practice in an urban community psychiatry clinic: A mixed-methods quality improvement project. *Journal of Evaluation and Clinical Practice*, *27*(4), 767–775. https://doi.org/10.1111/jep.13457

Walfish, S., McAlister, B., O'Donnell, P., & Lambert, M. J. (2012). An investigation of self-assessment bias in mental health providers. *Psychological Reports*, *110*(2), 639–644. https://doi.org/10.2466/02.07.17.PR0.110.2.639-644

Whipple, J. L., Lambert, M. J., Vermeersch, D. A., Smart, D. W., Nielsen, S. L., & Hawkins, E. J. (2003). Improving the effects of psychotherapy: The use of early identification of treatment and problem-solving strategies in routine practice. *Journal of Counseling Psychology*, *50*(1), 59–68. https://doi.org/10.1037/0022-0167.50.1.59

White, M. M., Lambert, M. J., Ogles, B. M., Mclaughlin, S. B., Bailey, R. J., & Tingey, K. M. (2015). Using the Assessment for Signal Clients as a feedback tool for reducing treatment failure. *Psychotherapy Research*, *25*(6), 724–734. https://doi.org/10.1080/10503307.2015.1009862

Winkeljohn Black, S., Owen, J., Chapman, N., Lavin, K., Drinane, J. M., & Kuo, P. (2017). Feedback informed treatment: An empirically supported case study of psychodynamic treatment. *Journal of Clinical Psychology*, *73*(11), 1499–1509. https://doi.org/10.1002/jclp.22529

Wright, C. V., Beattie, S. G., Galper, D. I., Church, A. S., Bufka, L. F., Brabender, V. M., & Smith, B. L. (2017). Assessment practices of professional psychologists: Results of a national survey. *Professional Psychology: Research and Practice*, *48*(2), 73–78. https://doi.org/10.1037/pro0000086

Wright, C. V., Goodheart, C., Bard, D., Bobbitt, B. L., Butt, Z., Lysell, K., McKay, D., & Stephens, K. (2020). Promoting measurement-based care and quality measure development: The APA

mental and behavioral health registry initiative. *Psychological Services*, *17*(3), 262–270. https://doi.org/10.1037/ser0000347

Youn, S. J., Castonguay, L. G., Xiao, H., Janis, R., McAleavey, A. A., Lockard, A. J., Locke, B. D., & Hayes, J. A. (2015). The Counseling Center Assessment of Psychological Symptoms (CCAPS): Merging clinical practice, training, and research. *Psychotherapy*, *52*(4) 432–441. https://doi.org/10.1037/pst0000029

Zimmerman, M., & McGlinchey, J. B. (2008). Depressed patients' acceptability of the use of self-administered scales to measure outcome in clinical practice. *Annals of Clinical Psychiatry*, *20*(3), 125–129. https://doi.org/10.1080/10401230802177680