Using feedback from patient-reported outcome measures in mental health services: A scoping study and typology

Christian U Krägeloh, PhD
Auckland University of Technology
Psychology
Auckland
chris.krageloh@aut.ac.nz

Karol Czuba,
Auckland University of Technology - Centre for Person Centred Research

Rex Billington,
Auckland University of Technology - Psychology

Paula Kersten,
Auckland University of Technology - Centre for Person Centred Research
Auckland
New Zealand

Richard Siegert,
Auckland University of Technology - Psychology

Disclosures and Acknowledgements

This study was funded by a contestable grant from the Faculty of Health and Environmental Sciences at Auckland University of Technology. Note that various terms are used to describe people receiving mental health services, including *patients*, *clients*, *consumers*, and *service users* (96). In order to be consistent with the established term PROM, we will occasionally refer to this group as patients, while acknowledging the sensitivity of this term, particularly to those who regard themselves as service users in recovery.

Word count

3,423 (excluding abstract, tables, and references)

Abstract

Objective: Routine evaluation of mental health services has become widespread, and the use of patient-reported outcome measures (PROMs) as clinical aids or discussion tools has been receiving increasing attention. The purpose of this scoping study is to provide a typology of the diverse ways in which studies reporting on PROM use in mental health services have utilized PROMs.

Methods: Iterative scoping searches of the literature identified articles reporting on the use of PROM feedback in mental health settings, which were then categorized to develop a typology along a dimension of intensity of PROM feedback, ranging from no feedback to patient and clinician to clinician-patient discussion that followed a formalized structure.

Results: Of the 172 studies that were identified, 27 were grouped into five categories, ranging from no PROMs feedback to either clinician or patient to studies in which a formalized structure was available by which PROM feedback could be discussed between clinician and patient. Of the 11 studies in the category with formalized feedback, nine studies reported some significant effects of feedback compared to a control condition, and two reported partial effects.

Conclusions: The proposed procedural typology helps explain the diversity of results from studies reporting on the effects of PROM feedback, by highlighting that PROM feedback appears to be more effective when integrated in a formalized and structured manner. Future work is required to isolate these effects from common procedural correlates, such as monitoring of therapeutic alliance.

Introduction

Formal routine evaluation of the outcomes of patient care has become increasingly widespread and plays an important role in mental health service provision (1). Over time, an immense array of Patient-Reported Outcome Measures (PROM) has been developed with the aim of including patients' perspectives within the process of health service provision. PROMs have typically been defined as patient-rated standardized measures of health or functional status, disability, participation, quality of life, well-being, or other specific and relevant outcomes of treatment, such as depression or anxiety (2,3).

Systematic reviews of the literature on the use of PROMs in clinical practice have typically associated PROMs with improvements in some aspects of care outcomes and quality of care. However, clear conclusions are difficult to derive due to methodological limitations of existing studies and lack of clarity regarding the goals and mechanisms of applying PROMs (4,5). Greenhalgh (6) provided an overview of the various ways and purposes of PROM use in clinical practice and presented the following categories: screening tools, monitoring tools, to promote patient-centred care, decision aids, methods to facilitate communication amongst multidisciplinary teams, and to evaluate the effectiveness of routine care and assessing quality of care. The first three uses involve individual-level data, while the last three involve group-level data.

Boyce and Browne (7) systematically reviewed studies that had investigated the effects of providing PROM feedback to healthcare professionals, but found that only one of all 16 eligible 16 studies obtained an overall positive effect. This study (8) reported on the results from an intervention at a hospital-based psychotherapy clinic, using as their PROM the 45-item *Outcome Questionnaire* (OQ-45) (9), which assesses client progress in therapy. The patient-therapist feedback group, in which results from repeated PROM administration were

discussed between patient and therapist, later showed significantly larger improvements in PROM scores than the treatment-as-usual group and another group in which only the therapist received PROM feedback.

Using PROM feedback with patients is consistent with the principles of mental health recovery, which focuses on the transformative aspects of overcoming mental health issues and thus emphasizes self-determination and individuals' sense of agency (10,11). As well as the clinician providing information on the patient's progress, PROMs attempt to capture the patient's view about whether they feel they are progressing, help patients appraise themselves, and reflect on their own recovery journey.

Whilst previous reviews on the effects of PROMs focused on different aspects, such as purpose and nature of applications (6) or the usefulness of PROM feedback at patient- and group-level (7), a systematic description of the range of procedures by which patient feedback is obtained in mental health services is lacking. In particular, the various levels of provider-patient communication associated with these procedures have not been systematically explored. The purpose of the present scoping study is to provide a typology of the ways in which studies reporting on PROM use in mental health services have administered PROMs. Understanding the scope of the literature and categorizing studies by levels of intensity of PROM feedback will highlight new ways of analysis that could help explain the diversity in outcomes when investigating the effects of PROMs (6,7) and provide clarity on whether providing PROM feedback is indeed associated with positive outcomes.

Methods

Scoping study

Scoping studies (12,13) are particularly suitable when the goal is to determine the scope and nature of a field that includes studies with a large range of methods and methodologies. The procedures used are similar to those of systematic literature reviews but tend to focus more on breadth rather than depth of the literature and thus do not exclude studies based on quality criteria. Because of the diversity of methods of studies that are being examined, the common analytical framework used is a descriptive-analytical method within the narrative tradition (12). Scoping studies chart the evidence and procedures of studies to increase conceptual clarity and to map the conceptual boundaries of a specific topic area (14).

Search strategy

The current scoping study was guided by an iterative search strategy (12). Following initial general familiarization with the literature on PROMs, structured searches on the database *Scopus* were conducted for peer-reviewed journal articles, with no restrictions on year of publication or language. Given the variety of terms used to describe this broad topic, search strategies were initially based on a related systematic review in palliative care (15) and also following other recommendations on the most sensitive and specific combination of terms with mental health content (16). The present review focused on PROM use in mental health settings, although this was initially broadly defined to capture a wide range of articles. The database search retrieved 59 articles, of which 13 were retained for more detailed review (3,17-28). Handsearches yielded two further review articles (7,29). After iterative searching of reference lists, citation searches, and specific searches of articles from prominent researchers in the area, a total of 166 articles were obtained. Of these, 109 were excluded as

they employed measures that were not standardized PROMs or were not about research in mental health settings but about mental health aspects in other fields, such as oncology, rehabilitation, general clinical practice, or substance abuse. Studies were also excluded if they merely reported on psychometric properties of PROMs, were surveys on the uptake of PROMs, or opinion pieces. Of the remaining 57 articles, 28 were review articles, and 29 empirical articles were categorized as outlined below. During the peer-review process, the anonymous reviewers identified another six studies that were also included.

Categorization of articles

Scoping studies follow an iterative process (12) that continually refines mapping criteria as new evidence is identified and analysed. Therefore, the author team met regularly for discussion to agree on adequate ways to categorize articles into levels of intensity of PROM feedback used. The final typology is presented in Box 1, containing five categories, ranging from category 1 (PROM scores were not fed back to clinician or patient) to category 5 (PROM feedback to clinician and patient, with a formalized structure to guide clinician-patient discussions).

Although studies in category 1 cannot provide any information on the effects of PROM feedback, retaining this category was useful for the purpose of establishing a typology of PROM feedback provision. Category 2 studies provide PROM feedback to clinicians, and studies in category 3 provide feedback to both clinicians and patients. In categories 2 and 3, discussion of PROM results may take place, although entirely at the discretion of the clinician. Any such discussion would therefore be incidental only. In category 4, clinician-patient PROM discussion is actively encouraged, but no formal structure guides this process.

And finally, category 5 are studies of actively encouraged clinician-patient PROM discussion based on available formal guidelines.

For any study to be allocated to one of the five categories, group consensus was required. Two of the authors (CK and KC) carefully read and categorized the articles independently and iteratively. Disagreement was resolved by discussion, which at times resulted in further refinement of the category wording. The remaining authors assisted with categorization of a selection of articles.

Most studies included control groups (typically category 1), but categorization was based on the procedure of the intervention group. Some studies (8,30-32) included two interventions that belonged to different categories, in which case the study was allocated to the highest category. Of the 35 reviewed studies, four could not be assessed due to incomplete information (24,33-35). Three additional studies were removed as they reported on the same dataset as a study that had already been included (36-38). Two studies (39,40) reported on different sub-groups of the same dataset and were treated as one study.

Results

Table 1 lists the 27 studies included in this review and provides a description of each study's sample, PROM feedback procedure, and results. Two studies belonged to category 1 (41,42), eight to category 2 (39,43-49), four to category 3 (32,50-52), two to category 4 (53,54), and eleven to category 5 (8,30,31,55-62). Almost half (8,30,31,41,42,44,46,47,49,58,59,61,62) of the studies reported on samples from the United States. One article (57) reported on a study conducted in six European countries. Apart from one Australian study (53), the remaining ones were from European countries: Germany (39,43,48,54), United Kingdom (50,52,60), the Netherlands (32,45), Ireland (56), Norway

(55), and Sweden (51). The study populations were diverse, including clinic in-patients (39,43,53,54,62), clinic or service out-patients (8,32,42,45,48,51,55,60,61), and clients at a variety of community-based services (41,44,50,52,57). Eight studies reported on data from clients at university counselling services (30,31,46,47,49,56,58,59), all of which, except for three (56,58,59), were from the same university.

Lambert authored ten of the articles listed in Table 1 (8,30,31,39,42,46,47,49,61,62), and all of these used the OQ-45 (9). Having been used in four additional studies (32,45,51,54), the OQ-45 was the most frequently used PROM. The second most frequently used PROM was the 4-item *Outcome Rating Scale* (ORS) (63). This measure, derived from the OQ-45, was used in four of the studies reviewed in Table 1 (55,56,58,59).

Category 1 functions as a baseline in the typology presented in Box 1. Only two articles (41,42) belonged to this category, largely because the scoping strategy outlined above searched for articles that reported on the use of PROM feedback. Although articles in this category cannot provide any information on the effectiveness of PROM feedback, these two articles are sufficient for the purposes of being exemplars of procedures in which PROMs were taken with no feedback to clinician or client.

All category 2 studies purported to investigate the effects that PROM feedback to clinicians has on patient outcomes. Six of these were randomized controlled trials, while the remaining two were quasi-experimental designs with close resemblance to the design of the other six studies. Table 2 summarizes which studies reported a significant effect of PROM feedback on PROM scores as well as on treatment duration. Two studies reported significant positive effects (43,44), while the remaining studies only reported significantly larger improvements for clients considered “not on track” or “at risk” (39,45-47,49) or no effect (48). Effect sizes were generally small or medium. In four of the studies that reported data on treatment duration (46-49), feedback was associated with significantly longer treatment for

not-on-track clients, and in three of these studies (46,48,49) feedback was also associated with significantly shorter durations for on-track clients. One study (39) reported no effect on treatment duration.

All four category 3 studies (32,50-52) were randomized controlled trials, and none reported a significant effect of PROM feedback to clinicians and patients compared to category 1 control conditions. One of the two category 4 studies reported a significant effect for only a sub-group of the sample and on some measures only (53), while the other category 4 study (54) did not obtain a significant effect. However, while discussion of feedback had been encouraged in that study (54), the authors reported that actual clinician-patient conversations about PROM feedback was rare.

Of the eleven studies in category 5, nine reported a significant effect of structured PROM feedback discussions. Two studies (56,60) obtained partial effects, namely significant results for only a sub-group in their sample or only for some of the outcome measures. Effect sizes were generally either small or medium.

Category 5 generally contained studies with more complex designs, such as multiple experimental groups. Three studies (8,30,31) compared the effects of category 5 feedback to category 2 and category 1 feedback. In all of these studies, feedback resulted in significantly more improved PROM scores than category 1. However, two studies (30,31) did not find a significant difference between the effect of category 2 and category 5 feedback, while one did (8).

Harmon et al. (30) reported significantly longer treatment durations for not-on-track clients, and Slade et al. (31) found that clients in the control condition required significantly more treatment sessions than clients in the feedback conditions. These two studies were also the only quasi-experimental designs. The other category 2 studies were randomized

controlled trials, and, of the six that reported on treatment duration data, none found a significant effect of PROM feedback on treatment duration.

Discussion

The present scoping study mapped previous research studies in mental health according to levels of intensity of PROM feedback use (Box 1), ranging from no feedback (category 1), clinician-only feedback (category 2), feedback to clinicians and patients (category 3), encouragement of mutual PROM discussion (category 4), to availability of formalized mechanisms that could guide such discussion (category 5). Previous systematic reviews concluded that evidence was lacking whether PROM feedback to healthcare professionals improved outcomes, as illustrated by Boyce and Browne's review of systematic reviews (7). In their own systematic review, Boyce and Browne (7) reported that only one of 16 studies had found a positive effect of PROM feedback, and six other partial effects. The present review of the mental health literature revealed that, of the 25 studies that could provide information on the effectiveness of PROM feedback (categories 2 to 5), 11 reported significant effects with generally small to medium effect sizes, 8 partial effects, and 6 no effects. Of the 11 studies in category 5, 9 found significant effects and 2 had partial effects, indicating that formalized clinician-patient PROM feedback was most strongly associated with beneficial outcomes. Compared to studies of categories 2 to 4, category 5 had a significantly higher ratio of studies reporting a statistically significant partial or full effect of feedback versus no effect ($\chi^2(1)=6.20$, $p<.05$) as well as a significantly higher ratio of studies reporting a statistically significant full effect versus only a partial or no significant effect ($\chi^2(1)=11.40$, $p<.01$).

The likelihood of reporting significant effects, however, did not increase in a linear fashion with feedback levels, as two of the category 2 studies found a significant effect, and five of the category 2 studies found a partial effect, while none of the category 3 studies and only one of the two category 4 studies obtained a partial effect. Two studies that examined both category 2 and category 5 experimental conditions did not find a significant difference between outcomes of these two conditions in their sample of clients at a university counseling center (30,31). Hawkins et al. (8), in contrast, reported improved outcomes for category 5 compared to category 2 for hospital outpatients, which could indicate that clinician-patient feedback may be more effective than clinician-only feedback in specific settings only.

With exception of one category 1 study (42), the studies associated with the research programme of Lambert were either of category 2 or 5, and all these studies used the OQ-45. The OQ-45 can be used in conjunction with its associated clinical support tools (CST). Previous studies applied CST with not-on-track patients, resulting in better treatment outcomes than using patient progress feedback with the OQ-45 only (64). Only one study (40) applied CST also for patients on track to recovery and found that this did not substantially enhance treatment. Our typology (Box 1) presents a uni-dimensional outline of intensity of PROM feedback use with clients, and within each category additional variables will be associated with positive therapeutic outcomes, thus creating variability of results within each category of feedback intensity. A formalized structure maximizes the likelihood that feedback is discussed with clients, which appears to be driving the beneficial results of PROM use in studies of category 5. Other aspects of procedural formalization may also be relevant, such as presence of computerized support tools (64), frequency of feedback (44), or whether PROMs are discussed amongst clinicians (65).

The lack of a feedback effect in category 3 and 4 studies is somewhat surprising, but could be related to procedural variations. Newnham et al. (53) speculate whether their delivery of

feedback during group therapy may have been qualitatively different to feedback during individual client-clinician interactions. Therapists' commitment to using PROMs is also related to effectiveness of feedback (45), and the lack of a feedback effect in the other category 4 study (54) may thus be linked the reportedly low frequency of therapist-initiated PROM discussions in that study. Finally, the feedback effects of the category 2 Lambert studies were largely related to clients considered "not on track" (64). With the exception of two studies (32,53), none of the other category 3 and 4 studies reported analyses by sub-groups, which may have revealed some partial feedback effects.

The ORS questionnaire was the second most frequently used PROM, and here, three studies reported significant effects of category 5 feedback (55,58,59), and one study (56) partial effects. Even more so than the OQ-45 and CST approach, the ORS is rarely offered on its own, but typically together with the *Session Rating Scale* (SRS) (66), which assesses the therapeutic alliance between client and clinician. Of the four studies that used the ORS, only one did not also use the SRS (56). The fact that the latter study "only" obtained a partial effect may thus indicate that other elements in addition to PROM feedback may be responsible for positive therapeutic outcomes.

Feedback is an integral part of meta-therapeutic dialogue, which, in addition to PROMs, often includes assessment of client needs and preferences, as well as therapeutic alliance (67). While the effects of PROM feedback might be difficult to disentangle from other aspects of such dialogue-directed approaches, qualitative reports explicitly point to positive experience of PROM feedback. Cheyne and Kinn (50) did not obtain a significant effect of category 3 PROM feedback, which may have been due to their small sample size. In another article, however, they extensively reported on the positive observations of counsellors when discussing PROM scores (36). Counsellors found that the Schedule for the Evaluation of Quality of Life (SEIQoL) (68) functioned well as an aid for client reflection and to enhance

therapeutic alliance. Similarly, Sundet (35) reported that completing items on the ORS may trigger very specific reactions, thus enhancing client-therapist dialogue by initiating, directing, or focusing conversations.

Limitations

Because of the lack of uniform terms to describe the approach of providing and/or discussing PROM feedback, the scoping method (12,13) was chosen to map out the field and inform our typology. The majority of articles were not obtained through database searches but through extensive iterative searches of citations, reference lists, handsearches, and searches for specific authors. However, because of the tendency of the scoping method's focus on breadth rather than depth, some relevant articles may have been missed. Unlike previous reviews (7), our search was not limited to articles published in English, and while two German-language articles were included (43,48), articles in languages other than English and German may have been missed.

Allocating articles to the categories of Box 1 was at times difficult due to unclear or incomplete information provided. Additionally, category allocation was based on reported procedure and not on how PROM feedback may have actually occurred. Studies in lower categories may have been *de facto* studies of higher categories if therapists frequently discussed PROM feedback with their clients. Similarly, studies of higher categories may have been *de facto* studies of lower categories, such as in the case of Puschner et al. (54), who reported that clinician-patient discussions rarely occurred despite being planned.

The present literature search identified a number of studies that had used the Clinical Outcomes in Routine Evaluation instruments (69). These category 1 articles were not included as the inclusion criteria did not extend to articles reporting on results from primary

care. Future reviews may analyse the extensive literature on primary care using the typology of the present review.

Conclusions

The present scoping study reviewed studies that reported on the effects of PROM feedback in mental health settings and provided a procedural typology of intensity of PROM feedback. Unlike previous reviews that reported little effects of PROM feedback, the present approach of synthesising results with the proposed procedural typology revealed that the availability of formalized guidelines for clinician-patient discussion of PROM feedback was most highly associated with beneficial therapeutic outcomes. Certainly, other variables such as the presence of computerized support tool software (64) or frequency of feedback (44) are also related to positive therapeutic outcomes, and these can be integrated into the present typology as variables that effect variability of results within each category of feedback intensity.

Using PROMs supports patient-centered care (6) as it recognizes patients as participant consumers, who should be active in planning and deciding on treatment options. Qualitative reports favour the use of PROM discussion, such as by enhancing clinician-patient communication and providing clients with mechanisms for reflective practice (36,70). However, as therapeutic approaches of discussing PROM feedback with clients tend to occur in conjunction with general emphasis on therapeutic alliance and meta-therapeutic dialogue, future work is required to isolate the effects of PROM feedback from such procedural correlates.

References

1. Bobbitt BL, Cate RA, Beardsley SD, Azocar F, McCulloch J: Quality improvement and outcomes in the future of professional psychology: Opportunities and challenges. Professional Psychology: Research and Practice 43:551-559, 2012

2. Dawson J, Doll H, Fitzpatrick R, Jenkinson C, Carr AJ: Routine use of patient reported outcome measures in healthcare settings. BMJ 340:464-467, 2010
3. Greenhalgh J, Long AF, Flynn R: The use of patient reported outcome measures in routine clinical practice: lack of impact or lack of theory? Social Science & Medicine 60:833-843, 2005
4. Marshall S, Haywood K, Fitzpatrick R: Impact of patient-reported outcome measures on routine practice: a structured review. Journal of Evaluation in Clinical Practice 12:559-568, 2006
5. Valderas JM, Kotzeva A, Espallargues M, Guyatt G, Ferrans CE, Halyard MY, Revicki DA, Symonds T, Parada A, Alonso J: The impact of measuring patient-reported outcomes in clinical practice: a systematic review of the literature. Quality of Life Research 17:179-193, 2008
6. Greenhalgh J: The applications of PROs in clinical practice: What are they, do they work, and why? Quality of Life Research 18:115-123, 2009
7. Boyce MB, Browne JP: Does providing feedback on patient-reported outcomes to healthcare professional result in better outcomes for patients? A systematic review. Quality of Life Research 22:2265-2278, 2013
8. Hawkins EJ, Lambert MJ, Vermeersch DA, Slade KL, Tuttle KC: The therapeutic effects of providing patient progress information to therapists and patients. Psychotherapy Research 14:308-327, 2004
9. Lambert MJ, Burlingame GM, Umphress V, Hansen NB, Vermeersch DA, Clouse GC, Yanchar SC: The reliability and validity of the Outcome Questionnaire. Clinical Psychology and Psychotherapy 3:249-258, 1996
10. Anthony WA: Recovery from mental illness: The guiding vision of the mental health system in the 1990s. Psychosocial Rehabilitation Journal 16:11-23, 1993
11. Deegan PE: Recovery as a self-directed process of healing and transformation. Occupational Therapy in Mental Health 17:5-21, 2002
12. Arksey H, O'Malley L: Scoping studies: Towards a methodological framework. International Journal of Social Research Methodology 8:19-32, 2005
13. Levac D, Colquhoun H, O'Brien KK: Scoping studies: advancing the methodology. Implementation Science 5:69, 2010
14. Davis K, Drey N, Gould D: What are scoping studies? A review of the nursing literature. International Journal of Nursing Studies 46:1386-1400, 2009
15. Antunes B, Harding R, Higginson IJ, on behalf of EUROIMPACT: Implementing patient-reported outcome measures in palliative care clinical practice: A systematic review of facilitators and barriers. Palliative Medicine 28:158-175, 2014
16. Wilczynski NL, Haynes RB, Team Hedges: Optimal search strategies for identifying mental health content in MEDLINE: an analytic survey. Annals of General Psychiatry 5:4, 2006
17. Burton L-J, Tyson S, McGovern A: Staff perceptions of using outcome measures in stroke rehabilitation. Disability & Rehabilitation 35:828-834, 2013
18. Coombs T, Stapley K, Pirkis J: The multiple uses of routine mental health outcome measures in Australia and New Zealand: experiences from the field. Australasian Psychiatry 19:247-253, 2011

19. Delaffon V, Anwar Z, Noushad F, Ahmed AS, Brugha TS: Use of Health of the Nation Outcome Scales in psychiatry. Advances in Psychiatric Treatment 18:173-179, 2012
20. Greenhalgh J, Flynn R, Long AF, Tyson S: Tacit and encoded knowledge in the use of standardised outcome measures in multidisciplinary team decision making: A case study of in-patient neurorehabilitation. Social Science & Medicine 67:183-194, 2008
21. Hatfield DR, Ogles BM: The use of outcome measures by psychologists in clinical practice. Professional Psychology: Research and Practice 35:485-491, 2004
22. James M, Kehoe R: Using the Health of the Nation Outcome Scales in clinical practice. Psychiatric Bulletin 23:536-538, 1999
23. Patterson P, Matthey S, Baker M: Using mental health outcome measures in everyday clinical practice. Australasian Psychiatry 14:133-136, 2006
24. Prabhu R, Oakley Browne M: The Use of the Health of the Nation Outcome Scale in an Outreach Rehabilitation Program. Australasian Psychiatry 16:195-199, 2008
25. Rey JM, Grayson D, Mojarrad T, Walter G: Changes in the rate of diagnosis of major depression in adolescents following routine use of a depression rating scale. Australian and New Zealand Journal of Psychiatry 36:229-233, 2002
26. Skinner A, Turner-Stokes L: The use of standardized outcome measures in rehabilitation centres in the UK. Clinicial Rehabilitation 20:609-615, 2006
27. Stevens AM, Gwilliam B, A'Hern R, Broadley K, Hardy J: Experience in the use of the palliative care outcome scale. Supportive Care in Cancer 13:1027-1034, 2005
28. Tavabie JA, Tavabie OD: Improving care in depression: qualitative study investigating the effects of using a mental health questionnaire. Quality in Primary Care 17:251-261, 2009
29. Sprangers MAG, Hall P, Morisky DE, Narrow WE, Dapueto J: Using patient-reported measurement to pave the path towards personalized medicine. Quality of Life Research 22:2631-2637, 2013
30. Harmon SC, Lambert MJ, Smart DM, Hawkins E, Nielsen SL, Slade K, Lutz W: Enhancing outcome for potential treatment failures: Therapist-client feedback and clinical support tools. Psychotherapy Research 17:379-392, 2007
31. Slade K, Lambert MJ, Harmon SC, Smart DW, Bailey R: Improving psychotherapy outcome: The use of immediate electronic feedback and revised clinical support tools. Clinical Psychology and Psychotherapy 15:287-303, 2008
32. de Jong K, Timman R, Hakkaart-van Roijen L, Vermeulen P, Kooiman K, Passchier J, van Busschbach J: The effect of outcome monitoring feedback to clinicians and patients in short and long-term psychotherapy: A randomized clinical trial. Psychotherapy Research in press
33. Asay TP, Lambert MJ, Gregersen AT, Goates MK: Using patient-focused research in evaluating treatment outcome in private practice. Journal of Clinical Psychology 58:1213-1225, 2002
34. Oades L, Deane F, Crowe T, Lambert WG, Kavanagh D, Lloyd C (2005). Collaborative recovery: an integrative model for working with individuals who experience chronic and recurring mental illness. Australasian Psychiatry 13(3): 279-284, 2005

35. Sundet R: Postmodern-oriented practices and patient-focused research: possibilities and hazards. Australian and New Zealand Journal of Family Therapy 33:299-308, 2012
36. Cheyne A, Kinn S: Counsellors' perspectives on the use of the Schedule for the Evaluation of Individual Quality of Life (SEIQoL) in an alcohol counselling setting. British Journal of Guidance & Counselling 29:35-46, 2001
37. Lambert MJ, Whipple JL, Bishop MJ, Vermeersch DA, Gray GV, Finch AE: Comparison of empirically-derived and rationally-derived methods for identifying patients at risk for treatment failure. Clinical Psychology and Psychotherapy 9:149-164, 2002
38. Halford, WK, Hayes S, Christensen A, Lambert M, Baucom DH, Atkins DC: Toward making progress feedback an effective common factor in couple therapy. Behavior Therapy 43:49-60, 2012
39. Probst T, Lambert MJ, Loew TH, Dahlbender RW, Göllner R, Tritt K: Feedback on patient progress and clinical support tools for therapists: Improved outcome for patients at risk of treatment failure in psychosomatic in-patient therapy under the conditions of routine practice. Journal of Psychosomatic Research 75:255-261, 2013
40. Probst T, Lambert MJ, Dahlbender RW, Loew TH, Tritt K: Providing patient progress feedback and clinical support tools to therapists: Is the therapeutic process of patients on-track to recovery enhanced in psychosomatic in-patient therapy under the conditions of routine practice?. Journal of Psychosomatic Research 76:477-484, 2014
41. Christensen A, Atkins DC, Berns S, Wheeler J, Baucom DH, Simpson LE: Traditional versus integrative behavioral couple therapy for significantly and chronically distressed married couples. Journal of Consulting and Clinical Psychology 72:176-191, 2004
42. Hannan C, Lambert MJ, Harmon C, Nielsen SL, Smart DW, Shimokawa K, Sutton SW: A lab test and algorithms for identifying clients at risk for treatment failure. Journal of Clinical Psychology/In Session 61:155-163-60, 2005
43. Berking M, Orth U, Lutz W: Wie effektiv sind systematische Rückmeldungen des Therapieverlaufs an den Therapeuten? Eine empirische Studie in einem stationär-verhaltenstherapeutischen Setting [How effective is systematic feedback of treatment progress to the therapist? An empirical study in a cognitive-behavioural-oriented impatient setting]. Zeitschrift für Klinische Psychologie und Psychotherapie 35:21-29, 2006
44. Bickman L, Kelley SD, Breda C, de Andrade AR, Riemer M: Effects of routine feedback to clinicians on mental health outcomes of youths: Results of a randomized trial. Psychiatric Services 62:1423-1429, 2011
45. de Jong K, van Sluis P, Nugter MA, Heiser WJ, Spinhoven P: Understanding the differential impact of outcome monitoring: Therapist variables that moderate feedback effects in a randomized clinical trial. Psychotherapy Research 22:464-474, 2012
46. Lambert MJ, Whipple JL, Smart DW, Vermeersch DA, Nielsen SL, Hawkins EJ: The effects of providing therapists with feedback on patient progress during psychotherapy: Are outcomes enhanced? Psychotherapy Research 11:49-68, 2001

47. Lambert MJ, Whipple JL, Vermeersch DA, Smart DW, Hawkins EJ, Nielsen SL, Goates M: Enhancing psychotherapy outcomes via providing feedback on client progress: A replication. Clinical Psychology and Psychotherapy 9:91-103, 2002
48. Lutz W, Wittmann WW, Böhnke JR, Rubel J, Steffanowski A: Zu den Ergebnissen des Modellprojektes der Techniker-Krankenkasse zum Qualitätsmonitoring in der ambulanten Psychotherapie aus Sicht des wissenschaftlichen Evaluationsteams [Results from the pilot project of the Techniker Krankenkasse (TK) 'quality monitoring in outpatient psychotherapy': The evaluators' perspective]. Psychotherapie, Psychosomatik, medizinische Psychologie 62:413-417, 2012
49. Whipple JL, Lambert MJ, Vermeersch DA, Smart DW, Nielsen SL, Hawkins EJ: Improving the effects of psychotherapy: The use of early identification of treatment failure and problem-solving strategies in routine practice. Journal of Counseling Psychology 50:59-68, 2003
50. Cheyne A, Kinn S: A pilot study for a randomised controlled trial of the use of the Schedule for the Evaluation of Individual Quality of Life (SEIQoL) in an alcohol counselling setting. Addiction Research & Theory 9:165-178, 2001
51. Hansson H, Rundberg J, Österling A, Öjehagen A, Berglund M: Intervention with feedback using Outcome Questionnaire 45 (OQ-45) in a Swedish psychiatric outpatient population. A randomized controlled trial. Nordic Journal of Psychiatry 67:274-281, 2013
52. Slade M, McCrone P, Kuipers E, Leese M, Cahill S, Parabiaghi A, Priebe S, Thornicroft G: Use of standardised outcome measures in adult mental health services. British Journal of Psychiatry 189:330-336, 2006
53. Newnham EA, Hooke GR, Page AC: Progress monitoring and feedback in psychiatric care reduces depressive symptoms. Journal of Affective Disorders 127:139-146, 2010
54. Puschner B, Schöfer D, Knaup C, Becker T: Outcome management in in-patient psychiatric care. Acta Psychiatrica Scandinavica 120:308-319, 2009
55. Anker MG, Duncan BL, Sparks JA: Using client feedback to improve couple therapy outcomes: A randomized clinical trial in a naturalistic setting. Journal of Consulting and Clinical Psychology 77:693-704, 2009
56. Murphy KP, Rashleigh CM, Timulak L: The relationship between progress feedback and therapeutic outcome in student counselling: A randomised control trial. Counselling Psychology Quarterly 25:1-18, 2012
57. Priebe S, McCabe R, Bullenkamp J, Hansson L, Lauber C, Martinez-Leal R et al: Structured patient-clinician communication and 1-year outcome in community mental healthcare. British Journal of Psychiatry 191:420-426, 2007
58. Reese RJ, Norsworthy LA, Rowlands SR: Does a continuous feedback system improve psychotherapy outcome? Psychotherapy Theory, Research, Practice, Training 46:418-431, 2009
59. Reese RJ, Toland MD, Slone NC, Norsworthy LA: Effect of client feedback on couple psychotherapy outcomes. Psychotherapy Theory, Research, Practice, Training 47:616-630, 2010
60. Schmidt U, Landau S, Pombo-Carril MG, Bara-Carril N, Reid Y, Murray K, Treaure J, Katzman M: Does personalized feedback improve the outcome of cognitive-

behavioural guided self-care in bulimia nervosa? A preliminary randomized controlled trial. British Journal of Clinical Psychology 45:111-121, 2006
61. Simon W, Lambert MJ, Harris MW, Busath G, Vazquez A: Providing patient progress information and clinical support tools to therapists: Effects on patients at risk of treatment failure. Psychotherapy Research 22:638-647, 2012
62. Simon W, Lambert MJ, Busath G, Vazquez A, Berkeljon A, Hyer K, Granley M, Berrett M: 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 23:287-300, 2013
63. Miller SD, Duncan BL, Brown J, Sparks JA, Claud DA: 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, 2003
64. Shimokawa K, Lambert MJ, Smart DW: 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, 2010
65. Reese RJ, Usher EL, Bowman DC, Norsworthy LA, Halstead JL, Rowlands SR, Chisholm RR: Using client feedback in psychotherapy training: An analysis of its influence on supervision and counselor self-efficacy. Training and Education in Professional Psychology 3:157-168, 2009
66. Duncan BL, Miller SD, Sparks JA, Claud DA, Reynolds LR, Brown J, Johnson LD: The Session Rating Scale: Preliminary psychometric properties of a “working” alliance measure. Journal of Brief Therapy 3:3-12, 2003
67. Bowens M, Cooper M: Development of a client feedback tool: A qualitative study of therapists’ experiences of using the Therapy Personalisation Forms. European Journal of Psychotherapy & Counselling 14:47-62, 2012
68. Joyce CRB, Hickey A, McGee HM, O’Boyle CA: A theory-based method for the evaluation of individual quality of life: The SEIQoL. Quality of Life Research 12:275-280, 2003
69. Evans C, Mellor-Clark J, Margison F, Barkham M, Audin K, Connell J, McGrath G: CORE: Clinical Outcomes in Routine Evaluation. Journal of Mental Health 9:247-255, 2000
70. Sundet R: Collaboration: family and therapist perspectives of helpful therapy. Journal of Marital & Family Therapy 37:236-249, 2011
71. Spanier GB: Measuring dyadic adjustment: New scales for assessing the quality of marriage and similar dyads. Journal of Marriage and the Family 38:15-28, 1976
72. Negy C, Snyder DK: Ethnicity and acculturation: Assessing Mexican American couples’ relationships using the Marital Satisfaction Inventory – Revised. Psychological Assessment 9:414-421, 1997
73. Weiss RL, Cerreto MC: The marital status inventory: Development of a measure of dissolution potential. American Journal of Familty Therapy 8:80-85, 1980
74. Lueger RJ, Howard KI, Martinovich Z, Lutz W, Anderson EE, Grissom G: Assessing treatment progress of individual patients using expected treatment response models. Journal of Consulting and Clinical Psychology 69:150-158, 2001

75. Endermann M: The Brief Symptom Inventory (BSI) as a screening tool for psychological disorders in patients with epilepsy and mild intellectual disabilities in residential care. Epilepsy & Behavior 7:85-94, 2005
76. Horowitz LM, Rosenberg SE, Baer BA, Ureño G, Villaseñor VS: Inventory of Interpersonal Problems: Psychometric properties and clinical applications. Journal of Consulting and Clinical Psychology 56:885-892, 1988
77. Grosse Holtforth M, Grawe K: Der Inkongruenzfragebogen (INK): Ein Messinstrument zur Analyse motivationaler Inkongruenz [The incruence questionnaire (INK): An instrument for the analysis of motivational incongruence]. Zeitschrift für Klinische Psychologie und Psychotherapie 32:315-323, 2003
78. Michalak J, Kosfelder J, Meyer F, Schulte D: Messung des Therapieerfolgs – Veränderungsmaße oder retrospektive Erfolgsbeurteilung [Measuring therapy outcome – pre-post effect sizes and retrospective measurement]. Zeitschrift für Klinische Psychologie und Psychotherapie 32:94-102, 2003
79. Athay MM, Riemer M, Bickman L: The Symptoms and Functioning Severity Scale (SFSS): Psychometric evaluation and discrepancies among youth, caregiver, and clinician ratings over time. Adminstration and Policy in Mental Health 39:13-29, 2012
80. de Jong K, Nugter MA, Lambert MJ, Burlingame GM: Handleiding voor afname en scoring van de Outcome Questionnaire (OQ-45), Salt Lake City, UT, OQ Measures LLC, 2009
81. Müller-Nordhorn J, Roll S, Willich SN: Comparison of the short form (SF)-12 health status instrument with the SF-36 in patients with coronary heart disease. Heart 90:523-527, 2004
82. Wittmann WW, Lutz W, Steffanowski A, Kriz D, Glahn EM, Völkle MC et al.: Qualitätsmonitoring in der ambulanten Psychotherapie: Abschlussbericht. Hamburg, Germany, Techniker Krankenkasse, 2011
83. Lutz W, Wittmann WW, Böhnke JR, Rubel J, Steffanowski A: Das Modellprojekt der TK zum Qualitätsmonitoring in der ambulanten Psychotherapie aus Sicht der Evaluatoren – Ein Plädoyer für mehr Psychotherapieforschung in Deutschland [The TK-Project quality monitoring in outpatient psychotherapy from the perspective of the evaluation team – A plea for more psychotherapy research in Germany]. Psychotherapie, Psychosomatik, medizinische Psychologie 63:225-228, 2013
84. Lambert MJ, Hannöver W, Nisslmüller K, Richard M, Kordy H: Fragebogen zum Ergebnis von Psychotherapie: Zur Reliabilität und Validität der deutschen Übersetzung des Outcome Questionnaire 45.2 (OQ-45.2) [Questionnaire on the results of psychotherapy: Reliability and validity of the German translation of the Outcome Questionnaire 45.2 (OQ-45.2)]. Zeitschrift für Klinische Psychologie und Psychotherapie 1:40-47, 2002
85. Wennberg P, Philips B, de Jong K: The Swedish version of the Outcome Questionnaire (OQ-45): Reliability and factor structure in a substance abuse sample. Psychology and Psychotherapy: Theory, Research and Practice 83:325-329, 2010

86. Priebe S, Huxley P, Knight S, Evans S: Application and results of the Manchester Short Assessment of Quality of Life (MANSA). International Journal of Social Psychiatry 45:7-12, 1999
87. Newnham EA, Hooke GR, Page AC: Monitoring treatment response and outcomes using the World Health Organization's Wellbeing Index in psychiatric care. Journal of Affective Disorders 122:133-138, 2010
88. Ware JE, Gandek B, IQOLA Project Group: The SF-36 health survey: Development and use in mental health research and the IQOLA Project. International Journal of Mental Health 23:49-73, 1994
89. Lovibond PF, Lovibond SH: The structure of negative emotional states: Comparison of the Depression Anxiety Stress Scales (DASS) with the Beck Depression and Anxiety Inventories. Behaviour Research and Therapy 33:335-343, 1995
90. Byrne SL, Hooke GR, Newnham EA, Page AC: The effects of progress monitoring on subsequent readmission to psychiatric care: A six-month follow-up. Journal of Affective Disorders 137:113-116, 2012
91. Freeston MH, Pléchaty M: Reconsideration of the Locke-Wallace Marital Adjustment Test: Is it still relevant for the 1990s? Psychological Reports 81:419-434, 1997
92. Bauer S, Winn S, Schmidt U, Kordy H: Construction, scoring and validation of the Short Evaluation of Eating Disorders (SEED). European Eating Disorders Review 13:191-200, 2005
93. Bjelland I, Dahl AA, Haug TT, Neckelmann D: The validity of the Hospital Anxiety and Depression Scale – An updated literature review. Journal of Psychosomatic Research 52:69-77, 2002
94. Kordy H, Richard M, Herrmann A, Murphy F, Treasure J, Charpentier P: A computer assisted eating disorder specific quality management system: EQUAL-TREAT. European Eating Disorders Review 7:239-258, 1999
95. Fritz CO, Morris PE, Richler JJ: Effect size estimates: Current use, calculations, and interpretation. Journal of Experimental Psychology: General 141:2-18, 2012
96. Simmons P, Hawley CJ, Gale TM, Sivakumaran T: Service user, patient, client, user or survivor: describing recipients of mental health services. The Psychiatrist 34:20-23, 2010

Box 1: Description of the criteria used to categorize articles by levels of intensity of PROM feedback discussion.

Category 1:	PROMs taken with no feedback provided to clinician or patient
Studies that used PROMs to assess the effect of treatment or an intervention, typically by	

comparing pre- with post-measures. The outcome reports were not fed back to the clinicians or the clients and in no way informed the intervention or treatment.	
Category 2:	PROM results reported back to clinician
Studies in which clients completed PROMs at some stage of their treatment, often at baseline and after treatment. The outcome reports were routinely fed back to clinicians but not the client, although clinicians were able to provide PROM feedback to their clients at their own discretion. This way of using PROM feedback enabled the clinicians to make decisions regarding the treatment plan.	
Category 3:	PROM results reported back to clinician and client
Studies that used PROMs to monitor the treatment outcome and fed back the outcome reports not only to clinicians, but also to the clients. Clinicians were able to react to clients' progress, but no process of including the outcome report in a discussion between clinician and client was proposed, and if discussions occurred, they were therefore incidental.	
Category 4:	PROM results reported back to clinician and client, with opportunities created for discussion
Studies that reported on PROM feedback to both clinician and client, and opportunities were created for outcomes to be discussed. This discussion was able to influence subsequent treatment, but such discussion was unstructured, or no structure or process was reported by the authors.	
Category 5:	PROM results reported back to clinician and client, with a formal procedure in which a discussion of PROMs can affect subsequent treatment
Studies that obtained PROMs, which were then fed back to clinician and client and were available for discussion for the purpose of informing subsequent treatment. The procedure for including PROMs in any such discussion was formalized and structured in forms of guidelines and recommendations.	

Table 1: Studies identified by the present scoping study are listed under the five categories presented in Box 1. For each study, information is provided on the design, number of participants and characteristics of the sample that provided the basis for statistical analyses, standardized or validated PROM(s) used (with reference to studies that reported on the psychometric properties of that PROM), procedure of the study related to the use of PROMs, and, lastly, a brief summary of the results.

Authors	Design	Sample	PROM(s) used	Procedure on use of PROM(s)	Results
Category 1					
Christensen et al. (2004) (41)	Randomized controlled trial	134 seriously and chronically distressed married couples undergoing a free therapy program in two cities in the United States	- 32-item *Dyadic Adjustment Scale* (DAS) (71), a self-report measure of marital satisfaction - 3 sub-scales from the *Marital Satisfaction Inventory—Revised* (MSI-R) (72): the 22-item *Global Distress Scale* (GDS), 19-item *Problem-Solving Communication* (PSC), and 13-item *Affective Communication* (AFC) scale. - 14-item *Marital Status Inventory* (MSI) (73), measuring thoughts, tentative, and actual steps undertaken toward	Couples were randomly assigned to one of two treatment conditions (comparing two treatment types). All clients completed various screening measures before and at intake. At intake, 13 weeks, and 26 weeks, couples completed all PROMs. At the end of treatment, clients completed relationship satisfaction and client evaluation of services measures.	The two treatment types were compared in terms of change in PROM scores.

			divorce - 68-item *Mental Health Index* (MHI), measuring current symptoms, life satisfaction, and well-being; this is a sub-scale of the *Compass Outpatient Treatment Assessment System* (74)		
Hannan et al. (2005) (42)	Single-group post-test design	618 clients at a university out-patient clinic in the United States	45-item *Outcome Questionnaire* (OQ-45) (9), measuring client progress along three dimensions: subjective discomfort (25 items), interpersonal relationships (9 items), and social role performance (11 items).	Clients completed an outcomes questionnaire before each therapy session. Routine feedback to therapists was suspended for a period of three weeks to investigate therapists' ability to estimate client progress.	Therapists tended to overpredict improvement of their clients and not to predict deterioration.
Category 2					
Berking et al. (2006) (43)	Randomized controlled trial	118 in-patients at a psychosomatics, psychotherapy, and behavioral medicine clinic	- 11-item German version of the *Brief Symptom Inventory* (BSI) (75) - 12-item German version of the *Inventory of Interpersonal*	Patients receiving cognitive-behavioral therapy were randomly allocated to either a feedback or no-feedback condition. All patients completed the EMI-B, BSI, IIP, and INK on the first day, two days later, and weekly from then onwards. In the feedback condition,	Average improvement on all outcome measures was significantly larger in the feedback group.

		in Germany	*Problems* (IIP) (76), a self-rated measure of interpersonal difficulties - 10-item *Inkongruenzfragebogen* (INK) (77), assessing extent of congruence of current situation with one's motivations and goals - 42-item *Veränderungfragebogen des Erlebens und Verhaltens* (VEV) (78), measuring therapy-induced changes in experience and behavior	therapists received the results the following day. At the end of therapy, patients completed the VEV.	
Bickman et al. (2011) (44)	Randomized controlled trial; note the substantial attrition in the study	340 youth (between 11 to 18 years of age) receiving home-based services from a private, for-profit, behavioral health organization at	32-item *Symptoms and Functioning Severity Scale* (SFSS) (79), which assesses the frequency of emotions of behaviors linked to typical mental health disorders in youth	Clients were randomly allocated to an experimental or a control group. At the end of a treatment session, clients completed a paper questionnaire. Clinicians of clients in the experimental group received weekly feedback (mean scores and alerts) and cumulative feedback every 90 days; clinicians of the control group only received the 90-day feedback.	Client-reported PROMs in the experimental group improved significantly faster than those in the control group.

		28 sites in the United States			
de Jong et al. (2012) (45)	Randomized controlled trial	413 out-patient receiving psychiatric treatment at a medium-sized healthcare institution in the Netherlands	OQ-45, Dutch version (80)	Patients were randomly allocated to an experimental feedback group or a no-feedback control group. All patients completed the PROM after sessions 1, 3, 5, and subsequently every 5th session. After each time a PROM was taken, therapists in the feedback group received an email containing information on the patient's PROM progress. No alarms were used, but therapists were able to identify "not on track" cases themselves. The study also investigated to what extent therapist characteristics may moderate the effects of feedback, and thus therapists completed a use-of-feedback questionnaire at the end of the study.	For clients identified as "not on track", feedback resulted in a significant positive effect on the PROMs when therapists indicated they had used the feedback with their patients.
Lambert et al. (2001) (46)	Randomized controlled trial	609 clients at a university counseling center in the United States	OQ-45	Clients were randomly allocated to an experimental or a control group. All clients completed the OQ-45 at intake and prior to each treatment session. Data from the control group were not shared with clients and therapists. In the experimental group, therapists were given the results on a graph and	For clients identified as "not on track", feedback resulted in significantly better outcome scores and significantly longer duration of treatment. For clients "on track", there were no

				alerted to the client's progress using a color-coding system. Clinicians' reactions to the feedback were not managed, with no mechanism to use the feedback in any systematic way.	significant differences in outcome measures, and number of treatment sessions was significantly less for the feedback condition.
Lambert et al. (2002) (47)	Quasi-experimental design where intervention was conducted after data for the control group had been collected	1,020 clients at a university counseling center in the United States	OQ-45	Intended as a replication of Lambert et al. (46) with a larger sample size. Clients during summer and fall semesters of 1999 were assigned to the control group, clients in winter and spring semesters of 2000 to the experimental (feedback) condition. All clients completed the OQ-45 at intake and prior to each treatment session. Data from the control group were not shared with clients and therapists. In the experimental group, therapists were given the results on a graph and were alerted to the client's progress using a color-coding system. Clinicians' reactions to the feedback were not managed, with no mechanism to use the feedback in any systematic way. However, therapists whose clients were in the feedback group received a tracking form, which was suggestive of possible clinician actions in response to	For clients identified as "not on track", feedback resulted in significantly better outcome scores and significantly longer duration of treatment. For clients "on track", there were no significant differences in outcome measures or treatment duration.

				feedback.	
Lutz et al. (2012) (48)	Randomized controlled trial	1,708 clients receiving out-patient psychotherapy treatment in one of three regions in Germany	- German version of the BSI - German version of the IIP - 12-item SF-12 health status questionnaire (81) - Some additional measures were taken depending on patients' main diagnosis.	Clinicians were randomly allocated to an experimental or a control group. In both groups, PROMs were taken at intake, discharge, and one year later. Patients in the experimental group also completed PROMs five times throughout treatment. Patients in the control group received treatment as usual. In the experimental group, therapists received immediate PROM feedback (summary and graphs) about their patients. There were no prescriptive guidelines on PROM feedback use, which meant that therapists could incorporate this information into therapy at their own discretion. Some of the more detailed information shown here was extracted from the final report of the so-called TK-model (82). Lutz et al. (83) note that the results of the study need to be interpreted with caution due to some compromising externally-imposed design modifications.	Feedback did not affect PROM scores. The groups also did not differ in terms of length of treatment.
Probst et al.	Randomized	252 in-patients	OQ-45, German version	Patients were randomly allocated to an	For patients at risk of

(2013) (39)	controlled trial	recruited from a psychosomatics department of a hospital and a psychosomatics hospital, both located in Germany. Probst et al. (39) reported on results from 43 patients at risk of outcome deterioration, and Probst et al. (40) reported on 209 patients considered on track.	(EB-45) (84)	experimental or a control group. All patients completed the OQ-45 every weekend. On Mondays, the feedback reports were given to the therapists of patients in the experimental group. Therapists could choose freely to discuss the feedback with their patients. Also included was the *Assessment of Signal Cases* scale, which measures therapeutic alliance, motivation for change, social support, and critical life events. This is part of clinical support tools (CST), which provide empirically-based problem-solving strategies.	deterioration, feedback significantly improved outcome scores (39). For patients "on track", the feedback condition did not have a significant effect (40).
Whipple et al. (2003) (49)	Quasi-experimental study where assignment of participants to experimental and intervention groups was	358 adult clients in a university counseling center in the United States	OQ-45	Clients were randomly allocated to an experimental (feedback) or a control group. All clients completed the OQ-45 at intake and prior to each treatment session. In the feedback group, results were presented to therapists in form of graphs and a color-coding system to signal client progress, as well as suggested decision rules. Therapists	For clients "not on track", feedback+CST resulted in significantly higher outcome scores than feedback only, which in turn resulted in significantly higher scores than no feedback. For clients

	determined randomly, but assignment to one of the experimental groups was nonrandom			whose clients were in the feedback group and considered "not on track" received a tracking form, which was suggestive of possible clinician actions in response to feedback. The experimental group was further divided into a feedback-only group and a feedback+CST group. However, this happened nonrandomly, as therapists decided the extent to which they opted to use CSTs.	"on track", there were no significant group differences. Clients considered "not on track" and who were in one of the two feedback groups remained in therapy significantly longer than "not on track" clients in the control group. For "on track" clients, therapy duration was significantly longer for the control group than the two feedback groups.
Category 3					
Cheyne & Kinn (2001) (50)	Pilot randomized controlled trial	42 consecutive referrals for alcohol counseling at a range of local community-based cognitive-behavioral counseling	*Schedule for the Evaluation of Individual Quality of Life* (SEIQoL) (68), which allows respondents to rate the importance of life areas to their overall quality of life.	Clients were randomly allocated to an experimental or a control group. Clients in the experimental group completed the SEIQoL together with the therapist at the first and at the end of the final counseling session as well as at 4- and 8-week review appointments. Four weeks after completion of treatment, all participants were posted a	The experimental condition resulted in a larger proportion of favorable outcomes, but this effect was not statistically significant. A separate publication (36) reports qualitative data on the positive

		services in the United Kingdom		questionnaire on satisfaction with service and outcomes achieved (42% response rate).	experiences of completing the SEIQoL with clients.
de Jong et al. (in press) (32)	Randomized controlled trial	475 out-patients at private psychotherapy practices and mental health institutes in the Netherlands	OQ-45, Dutch version (80)	Patients were randomly allocated to a no-feedback control group, a therapist-only feedback group, or a therapist-patient feedback group. All patients completed the OQ-45 online (typically on a laptop in the therapist's waiting room) prior to each therapy session, but not more often than once per week. In the two feedback conditions, PROM scores and feedback messages were generated immediately, and subsequent discussion of feedback was at the therapists' discretion.	Group differences of OQ-45 scores at treatment end were not significant, although the therapist-client group had the smallest number of deteriorated cases. For "not on track" clients, feedback was preventive of negative outcomes.
Hansson et al. (2013) (51)	Randomized controlled trial	262 patients in two general psychiatry out-patient clinics in Sweden	OQ-45, Swedish version (85)	Patients were randomly allocated to an experimental or a control group. At intake, all patients completed the OQ-45, as well as at each further visit to the clinic, but not more often than once a week. Therapists of patients in the experimental group received their clients' OQ-45 scores via a web application before each subsequent visit, which was also handed to the patient. In the control group, neither	Patients in the experimental group had larger improvements in their outcome scores, but this difference did not reach statistical significance.

				patient nor therapist received feedback.	
Slade et al. (2006) (52)	Randomized controlled trial	160 patients of eight community mental health teams in the United Kingdom	12-item *Manchester Short Assessment* (MANSA) (86), QOL instrument	Patients were randomly allocated to an experimental or a control group. Both groups received treatment as usual. Patients and therapists in the experimental group also completed a monthly postal questionnaire and were sent identical feedback every three months in form of graphics and text that also highlighted areas of disagreement between patient and therapist.	There were no significant group differences in quality of life scores, as well as no significant differences in scores of patient-rated unmet needs and other secondary measures that were rated by therapists. The intervention, however, resulted in significantly reduced psychiatric in-patient days.
Category 4					
Newnham et al. (2010) (53)	Historical cohort design	1,308 consecutive in-patients and day patients participating in 10-day cognitive behavioral group therapy at a private psychiatric hospital in	- 5-item *World Health Organization Wellbeing Index* (WHO-5) (87), a measure of positive mental health - Four subscales (4-item *vitality*, 2-item *social functioning*, 3-item *role emotion*, and 5-item *mental*	Patients in the first cohort (*n*=461) received treatment as usual. Patients in the second cohort (*n*=439) completed the WHO-5 every second day but did not receive feedback (scores and a graph with accompanying explanation) until the final day of therapy, where they were then given an opportunity to discuss their scores during the group session. Patients in the third cohort (*n*=408) completed the WHO-5 every	There was no effect of feedback on WHO-5 scores. For patients "not on track", feedback was significantly associated with decreased depressive symptoms (DASS-21) and the *vitality* and *role emotion* subscales of the SF-36, but not for any of the

		Australia	*health*) of the SF-36 health status questionnaire (88) - 21-item *Depression Anxiety Stress Scale* (DASS-21) (89), a measure of negative emotional symptoms	second day and received the same WHO-5 feedback from their therapists midway through treatment (Day 5) and on the final day, again with opportunities to discuss scores. Therapists were not given specific instructions on the use of feedback. Patients in all groups also completed the DASS-21 and SF-36 at admission and discharge.	other subscale measures. Byrne et al. (90) reported that, post-treatment, "on track" patients in the third cohort were significantly less likely to be readmitted than "on track" patients of the second cohort.
Puschner et al. (2009) (54)	Randomized controlled trial	264 adults receiving in-patient treatment at a psychiatric hospital in Germany	OQ-45, German version (EB-45)	Clinicians were randomly allocated to an experimental or a control group. All patients completed the EB-45 at intake, every week thereafter, and at discharge. In the experimental group, patients and clinicians received summary information a day or two after completion of the PROM. This information consisted of graphs, text with treatment recommendation and possible alert messages, and encouragement for patients and clinicians to discuss the results. However, no guidelines for such discussion were provided. Patients and clinicians in the control group received no feedback.	There was no significant effect of feedback on treatment outcome as measured by the EB-45. Most patients found the feedback useful for motivation, but there were mixed views on their effectiveness. Most patients reported that they rarely discussed the feedback with professionals or carers.

Category 5					
Anker et al. (2009) (55)	Randomized controlled trial	205 couples seeking out-patient couple therapy at a family counseling agency in Norway	- 4-item *Outcome Rating Scale* (ORS) (63), derived from the OQ-45 - 15-item Locke-Wallace Marital Adjustment Test (LW) (91), covering aspects of marital functioning and satisfaction	Participants were randomly allocated to an experimental (feedback) or a control (treatment as usual) group. Participants completed the ORS and LW before the first session, the ORS prior to each subsequent session, and the ORS and LW again six months after the final session. In the control group, ORS was completed by participants in the presence of a secretary, and results were not fed back to either participants or therapist. In the experimental group, ORS was rated in the presence of the therapist prior to each session and scored immediately. Therapist were trained to incorporate into their treatment the ORS feedback and associated computer-generated treatment and progress feedback. They were also advised to show the results to the clients and initiate discussions, although this was not monitored. Clients also completed the *Session Rating Scale* (SRS) (66), a measure of therapeutic alliance	The improvements in ORS scores were significantly higher in the experimental (feedback) group than the control group, which was maintained at six-month follow-up.
Harmon et al.	Quasi-	1,374 adult	OQ-45	Due to attrition, not all clients could be	Mean OQ-45 scores

(2007) (30)	experimental design with nonrandom group allocation and comparison group from archival data	clients seeking treatment at a large university counseling center in the United States		allocated randomly to the two intervention groups (feedback to both therapists and clients versus feedback to therapists only). Archival data (*n*=1,445) from the same clinic and therapists served as a no-feedback control group. Clients completed the OQ-45 and weekly thereafter. Prior to each session, previous week's scores were made available as feedback in form of graphs as well as using a color-coding system to categorize client progress. In both groups, clients considered "not on track" were further randomly allocated to either CST feedback (where results from additional measures of therapeutic alliance, stages of change, and social support were taken) versus no CST feedback. Clients who received feedback and were not responding well to treatment were encouraged to discuss their concerns about lack of progress and idea for therapy modifications. Clinicians' reactions to the PROM feedback were not managed. Therapists who received feedback+CST were able to consult a CST manual for treatment suggestions	improved significantly more for the feedback groups compared to the archival no-feedback control group. There was no significant difference between the two intervention groups of feedback to both therapists and clients versus feedback to therapists only. However, CST feedback (in addition to PROM feedback to therapist only or therapist and client) resulted in significantly improved outcomes than feedback without CST. Clients considered "not on track" received significantly more sessions in the feedback conditions than clients in the control group.

				based on feedback data.	
Hawkins et al. (2004) (8)	Randomized controlled trial	201 adults seeking out-patient psychotherapy services at a hospital-based clinic in the United States	OQ-45	Clients were nonrandomly assigned to therapists, based on their availability, but were subsequently assigned randomly to one of two treatment conditions (feedback to both therapist and client or feedback to therapist only) or the control condition (treatment as usual with no PROM feedback). All clients completed the OQ-45 at intake and after each treatment session. In the feedback conditions, previous week's scores were made available prior to each session in form of graphs as well as using a color-coding system to categorize client progress and make treatment recommendations (similar to 46,47). However, clinicians' reactions to the PROM feedback were not managed or monitored. In the client-therapist feedback condition, clients also received written feedback messages, and, if identified as not progressing, they were encouraged to discuss personal concerns about their progress and potential treatment modifications. Again, a format was available to discuss treatment progress, although	The largest improvement of OQ-45 scores was for clients in the client-therapist feedback condition, followed by therapist-only feedback, and finally the control condition. These differences were statistically significant. For clients considered "not on track" only, there were no significant group differences, although this may likely have been due to small sample size. There were no significant group effects on duration of treatment received.

				interactions with patients were not monitored.	
Murphy et al. (2012) (56)	Randomized controlled trial	110 adult clients at a university counseling service in Ireland	ORS	The ORS is typically administered in conjunction with the SRS, a measure of therapeutic alliance. The purpose of this study was to test the effects of ORS on its own. Clients were randomly allocated to an experimental (feedback to both therapist and client) or a no-feedback control group. All clients completed the ORS at intake and before each subsequent session. In the control group, clients completed the ORS in the presence of a researcher (except for the very first administration), and neither client nor therapist received feedback on ORS scores. In the experimental group, clients completed the ORS in front of the therapist using a software program, which instantly generated score feedback, such as in form of progress graphs. Therapists could decide freely on to react to this feedback and such as whether to discuss it with clients, but were provided with an ORS and SRS manual that provided them with strategies and recommendations for appropriate	Feedback resulted in significant differences for clients with anxiety issues, but not for clients with depression, relationship issues, or other concerns. There was no effect of feedback on treatment duration.

				course of action in response to ORS scores.	
Priebe et al. (2007) (57)	Randomized controlled trial	507 patients with severe and enduring mental illness who used community psychiatric services in one of six European countries (Germany, the Netherlands, Spain, Sweden, Switzerland, and the United Kingdom)	MANSA	Clinicians were randomly allocated to an experimental or control group. Clinicians in the control group provided treatment as usual. Clinicians in the experimental group implemented a manualized computer-mediated intervention. In this feedback intervention, patients rated their QOL approximately every two months during routine care, which was then followed up by questions whether patients wished additional support for particular domains. Patients in the control group completed the QOL questionnaire prior treatment and 12 months later. Other measures included satisfaction with treatment and unmet care needs.	QOL scores were significantly higher for the experimental group 12 months later, despite the presence of ceiling effects in the measure. The effect size of this group difference was higher when only analyzing results of participants with a low initial QOL score.
Reese et al. (2009) (58)	Randomized controlled trial	Study 1: 74 clients at a university counseling center in the United States Study 2: 74 clients receiving	ORS	Study 1: Clients were randomly assigned to an experimental (feedback) or control group. Clients in the control group were given the ORS at intake and end of treatment. Responses were not analyzed by the therapist, nor were any scores made available to the therapist. In the feedback condition, clients	In both studies, clients in the experimental (feedback) group received significantly larger gains in ORS scores than clients in the control group, indicating improved

		individual therapy at a graduate training clinic for a marriage and family therapy master's program in the United States		completed the ORS at the beginning of each session and the SRS at the end of each session. ORS graphs were generated as feedback, and general guidelines were available on how the therapist may proceed, although this was not monitored or managed. Study 2: Unlike in Study 1, therapists rather than clients were randomly allocated to either feedback or no-feedback groups. Another difference was that clients in the control group completed the ORS at the beginning of each session. However, results were not seen by the therapists in the control condition.	outcomes. There were no significant differences in number of sessions attended.
Reese et al. (2010) (59)	Randomized controlled trial	46 heterosexual couples receiving couple therapy at a graduate training clinic for a marriage and family therapy master's program in the	ORS	Intended as a replication of Anker et al. (55) with a sample from the United States. Couples were randomly assigned to an experimental (feedback) or control (treatment as usual) condition. All clients completed the ORS at the beginning of each session and the SRS at the end of each session. In the feedback group, ORS graphs were generated as feedback, and general guidelines were available on how the	Couples in the experimental (feedback) group received significantly larger and faster gains in ORS scores than clients in the control group, indicated improved outcomes.

		United States		therapist may proceed, although this was not monitored or managed.	
Schmidt et al. (2006) (60)	Randomized controlled trial	61 patients with bulimia nervosa or eating disorder not otherwise specified at a specialist eating disorder unit received guided self-help cognitive-behavioral therapy in the United Kingdom	- 6-item *Short Evaluation of Eating Disorders* (SEED) (92), a self-rated measure of severity of anorexia and bulimia symptoms - 14-item *Hospital Anxiety and Depression Scale* (HADS) (93), a self-rated assessment of anxiety and depression symptoms	Patients were randomly assigned to an experimental (feedback) or control (no feedback) group. Patients in the feedback group received a personalized letter after initial assessment, including feedback from physical examination and blood tests. A symptom feedback form was completed collaboratively by patient and therapist half way through treatment, and patients also received an end-of-treatment feedback letter from their therapist. All patients completed all PROMs prior to allocation to groups and at end of treatment, as well as the SEED only at 6-month follow-up. Throughout treatment, patients in the feedback group received two-weekly computerized PROM feedback. Patients in the control group completed the same number of within-treatment computerized assessments, but did not receive any of the feedback listed above. Feedback in the experimental group was also guided by an outcome monitoring and feedback system, providing automated feedback	Feedback did not have an effect on treatment up-take or drop-out. Feedback resulted in significantly larger improvements on scores for dietary restriction, but not for scores on bingeing, vomiting, or exercise.

				about progress (94).	
Simon et al. (2012) (61)	Randomized controlled trial	370 adults seeking psychotherapy services in a hospital-based out-patient clinic in the United States	OQ-45	Clients were randomly assigned to an experimental (feedback) or control (no feedback) condition. All clients completed the OQ-45 prior to each session. The CST tool was used for "not on track" cases in the feedback condition, which, for example, provided the therapists with decision trees for problem-solving, treatment suggestions, and progress alerts and tools to deal with "not on track patients". Therapists were instructed to present the PROM feedback to their clients, although this was not monitored.	OQ-45 scores of the feedback group improved significantly more than those of the no-feedback control group, albeit with a small effect size. The mean number of sessions was not significantly different between the two groups.
Simon et al. (2013) (62)	Randomized controlled trial	133 adults seeking in-patient treatment at an eating disorder hospital in the United States	OQ-45	The study's procedure was identical procedure to that of Simon et al. (61). The purpose of this study was to extend investigations of the effect of PROM feedback to a new population of clients.	PROM scores of the feedback group improved significantly more than those of the no-feedback control group, albeit with a small effect size. Body mass index scores increased in both conditions, with no significant group

					differences.
Slade et al. (2008) (31)	Quasi-experimental design with random assignment of participants to one of two feedback types, but with a comparison group from archival data	1,101 adult clients in a university counseling center in the United States, compared with archival data from 2,818 clients under no-feedback and feedback conditions in the same clinic (30,46,47,49). However, only data were presented from patients that were considered "not on track".	OQ-45	Clients were randomly assigned to one of two treatment conditions (feedback to both therapist and client or feedback to therapist only). Archival data from the same clinic and therapists allowed comparisons with no-feedback conditions and delayed feedback conditions. Compared to previous studies in the same clinic where feedback conditions gave feedback that was delayed by one week (30,46,47,49), this study used an immediate electronic feedback system, which provided instant automated PROM feedback. In the therapist-only feedback condition, therapists were encouraged to use the feedback in their treatment, but their reactions to the PROM feedback were not managed or monitored. In the client-therapist feedback condition, clients also received written feedback messages, and, if identified as not progressing, they were encouraged to discuss personal concerns about their progress and potential treatment modifications. CST feedback and decision trees were also provided to	There were no significant differences between the two treatment conditions (feedback to both therapist and client or feedback to therapist only), but significant improvements compared to treatment as usual. Immediate electronic feedback did not lead to significantly larger gains in outcome scores. Clients in the treatment as usual condition received significantly more treatment sessions.

				client and therapists for clients considered “not on track”. The focus of this study was on patients “not on track” only.	

Table 2: Summary of reported effects of PROM feedback on PROM scores and duration of therapy. Note the following conventional cut-off values to determine effect sizes: d >.20 small, d >.50 medium, d >.80 large, and η^2 >.01 small, η^2 >.06 medium, η^2 >.14 large (95).

Study	Effect of feedback on PROM scores	Treatment length
Category 2:		
Berking et al. (2006) (43)	significant (*d*=0.47 to *d*=0.50)	not reported
Bickman et al. (2011) (44)	significant (*d*=0.18)	not reported
de Jong et al. (2012) (45)	significant positive effect only for "not on track" patients and when therapists reported use of feedback	not reported
Lambert et al. (2001) (46)	- significant for "not on track" clients (*d*=0.44) - not significant for "on track" clients	feedback associated with significantly longer treatment for "not on track" clients and significantly fewer days for "on track" clients
Lambert et al. (2002) (47)	- significant for "not on track" clients (*d*=0.40) - not significant for "on track" clients	feedback associated with significantly longer treatment for "not on track" clients
Lutz et al. (2012) (48)	not significant	feedback associated with significantly shorter treatment; "not on track" patients received longer treatment and "on track" patients less
Probst et al. (2013) (39)	- significant for "at risk" patients (*d*=0.54) - not significant for "on track" clients (40)	not significant
Whipple et al. (2003) (49)	- significant for "not on track" clients (*d*=0.70 and *d*=0.28) - not significant for "on track" clients	feedback associated with significantly longer treatment for "not on track" clients and significantly fewer days for "on track" clients

Category 3:		
Cheyne & Kinn (2001) (50)	not significant	no difference in number of appointments
de Jong et al. (in press) (32)	not significant	not significant
Hansson et al. (2013) (51)	not significant	no difference in number of clinic visits
Slade et al. (2006) (52)	not significant	feedback associated with significantly reduced in-patient days
Category 4:		
Newnham et al. (2010) (53)	significant only for clients "not on track" and for some of the measures	not applicable, as this was a 10-day program
Puschner et al. (2009) (54)	not significant	not reported
Category 5:		
Anker et al. (2009) (55)	significant (d=0.50)	not reported
Harmon et al. (2007) (30)	- significant (both categories 2 and 5 more improved outcome than category 1) (d=0.23 and d=0.33) - not significant for categories 2 vs 5	feedback associated with significantly longer treatment for "not on track" clients
Hawkins et al. (2004) (8)	- significant (category 5 more improved outcome than categories 2 and 1) (η^2=.02 and η^2=.04) - significant (data from categories 2 and 5 combined more improved outcome than category 1) (η^2=.02)	not significant
Murphy et al. (2012) (56)	significant only for a sub-group of the sample	not significant
Priebe et al. (2007) (57)	significant (d=0.20, or d=0.43 if including	not reported

	only participants with low initial PROM scores)	
Reese et al. (2009) (58)	significant (η^2=.07 and (η^2=.10)	not significant
Reese et al. (2010) (59)	significant (*d*=0.81)	not reported
Schmidt et al. (2006) (60)	significant only for some measures	not significant
Simon et al. (2012) (61)	significant (η^2=.02)	not significant
Simon et al. (2013) (62)	significant (*d*=0.30)	not significant
Slade et al. (2008) (31)	- significant (both categories 2 and 5 more improved outcome than category 1) (*d*=0.35 and *d*=0.48) - not significant for categories 2 vs 5	significantly more treatment sessions for category 1 control