BJPsych | The British Journal of Psychiatry (2009) 195, 15–22. doi: 10.1192/bjp.bp.108.053967

Review article

Effect of feedback of treatment outcome in specialist mental healthcare: meta-analysis

Carina Knaup, Markus Koesters, Dorothea Schoefer, Thomas Becker and Bernd Puschner

Background
Feedback of treatment outcome during the course of therapy (outcome management) is increasingly considered to be beneficial for improving the quality of mental healthcare.

Aims
To review the impact of feedback of outcome to practitioners and/or patients in specialist mental health services.

Method
A systematic search and meta-analysis of controlled trials using outcome management in mental health services published in English or German language.

Results
Twelve studies met inclusion criteria. Feeding back outcome showed a small, but significant ($d = 0.10$; 95% CI 0.01–0.19) positive short-term effect on the mental health of individuals that did not prevail in the long run. Subgroup analysis revealed no significant differences regarding feedback modalities. Outcome management did not contribute to a reduction of treatment duration.

Conclusions
Evidence on the effects of outcome management in mental healthcare is promising. More targeted research is needed in order to identify the effective ingredients of outcome feedback and to assess its cost-effectiveness.

Declaration of interest
None.

In many countries, predominantly in the USA, Australia and to a growing degree in Western Europe, the complex management of mental healthcare requires quality assurance and ongoing treatment monitoring. Although routine data collection in mental healthcare is becoming more widespread, these data are rarely fed back to practitioners – let alone to patients – in a meaningful way.

Continuous monitoring of patient outcome was recommended as a means to improve quality of health services 20 years ago.[1] Ten years later, Lyons *et al*[2] presented a concept on how to combine measurement with management of clinical outcomes in mental healthcare. This approach has been refined in the meantime, for example for quality management in psychotherapy.[3] Feedback of patient outcome to stakeholders constitutes the prime instrument of outcome management and can be regarded as the most direct link between research and practice since outcome data are being used immediately in order to improve service provision.

Two reviews have summarised the evidence on feedback of outcome in general healthcare[4,5] and found some positive effects on the process of care, but none on patient health status. By focusing on specific settings,[6] interventions[7] or diagnostic groups,[8] the scope of reviews on the impact of outcome management in mental healthcare has been rather restricted. Thus, evidence on the effectiveness of feedback in mental healthcare is inconclusive: although Gilbody *et al*[8] could not find any randomised controlled trials (RCTs) investigating the effect of use of outcome measures in the routine treatment of people with schizophrenia, Lambert *et al*[7] stated that in out-patient psychotherapy 'it may be time for clinicians to routinely track and formally monitor patient treatment response'. Furthermore, in many reviews, information from outcome measures for screening purposes has been mixed with continuous outcome management. Thus, there is a need for a systematic review of the effectiveness of feedback of outcome in specialist mental healthcare.

Method

Search strategy

A systematic literature search was conducted in the electronic databases Medline, PSYNDEX, PsycINFO, the Cochrane Central Register of Controlled Trials, the Cochrane Database of Systematic Reviews, the Current Controlled Trials Register, and the world wide web using Google and Google Scholar in November 2006. This search was updated in March 2008 using the same search criteria. A broad set of search terms was applied in order to achieve maximum sensitivity. Keywords entered were 'mental disorders' combined with 'feedback' and 'treatment outcome' as well as diverse synonyms of these search terms. A total of 1066 publications were identified in this first step. Additional hand searching was carried out using the reference lists of relevant articles and appropriate reviews.[4,6,9,10] If data were missing or not apt for use in meta-analysis, authors were contacted for further information (details of the search strategy and the search terms used can be obtained from B.P.).

Inclusion criteria

English- as well as German-language publications were included in this systematic review if they examined feedback of outcome in specialist mental healthcare settings. Study participants had to be of adult age with mental health problems treated in psychiatric or psychotherapeutic settings (community, in-patient, and out-patient mental health services). Feedback was defined as providing mental healthcare professionals and/or patients with individual information on treatment outcome based on standardised measures (e.g. mental health, unmet needs). Only studies with a controlled design (not necessarily randomised) that evaluated effects of feedback interventions on patient outcome such as mental health were selected for analysis. Two of the authors (C.K. and B.P.) independently decided about study inclusion. Disagreements were resolved by discussion.

Exclusion criteria

Duplicate publications, abstracts and studies in which outcome assessment was used exclusively for routine screening or diagnosis were excluded.

Data extraction

Study characteristics, features of the feedback systems and outcome data were extracted from each included article (by C.K. and subsequently checked by B.P.) using a standard coding form developed for this review. This form included authors, publication year, study type, country, setting, patient and therapist population, unit of randomisation, inclusion and exclusion criteria, groups, feedback and effect instruments, measurement points, feedback intervention (duration, type and content of feedback, recipient, frequency and timing), outcomes assessed, as well as results and method of analysis. Furthermore, cost data (e.g. treatment duration) were noted if available.

Procedures

Data synthesis was conducted using the following two-step procedure: first, a qualitative analysis was conducted in order to provide a narrative overview. Second, a quantitative data synthesis (meta-analysis) was carried out separately for all trials with sufficient data on mental health outcome and treatment duration as end-points of the intervention. All symptoms scales applied in the studies were used as a measure of mental health outcome. If studies reported more than one scale or compound dimensions of an outcome, a single mean effect size was calculated for each study in order to avoid bias in favour of trials with multiple outcome measures.

Given our focus on differences in post-treatment outcome in controlled studies, effect sizes were calculated on the basis of (raw-) post-score analysis using the formula $d=(M_1-M_2)/\text{s.d.}$, where M_1 and M_2 are the means of the outcome measures in the intervention (having received feedback) and control groups, and s.d. is the pooled standard deviation of the two groups' post-scores. These effect sizes were corrected in order to obtain unbiased estimators using procedures suggested by Hedges.[11] The direction of effect was standardised so that a positive value indicates a better outcome in the feedback condition. As larger sample sizes lead to more reliable estimators, all calculated effect sizes were weighted by the inverse of their variance.[12]

Some studies reported the results of more than two groups.[13–15] In these cases, effect sizes of each control group comparison were pooled, resulting in one effect size per trial. An exception was made in a study with subsequent phases[16] – each with an intervention and a control group – where both study phases were treated as two independent trials and effect sizes were calculated separately.

Heterogeneity between and within the selected studies was anticipated as a result of variation in settings, feedback interventions, outcome instruments, the mental disorders the participants had and mental health professionals' therapeutic orientation. Therefore, the meta-analysis was performed using a random-effects model as suggested by DeSimonian & Laird[17] that considers both intra- and inter-study variability. To quantify heterogeneity, the I^2 statistic, which represents an estimation of the percentage of total variation across studies due to heterogeneity, was calculated. I^2 values below 25% indicate low heterogeneity whereas those above 75% point towards high heterogeneity.[18]

To prevent dilution of effects as a result of differences in length of follow-up periods within one trial or in comparison with the others, we calculated two separate meta-analyses: short-term for follow-up intervals up to 9 weeks after intake (predominantly at the end of therapy), and long-term for those between 3 and 12 months.

Evidence of publication bias was analysed using graphical control techniques (funnel plots) and Egger's regression test.[19] A fail-safe number was calculated in view of the fact that non-significant studies tend to be unpublished ('file-drawer problem'). This parameter estimates the number of studies with an effect size of zero required to nullify the overall effect size. To test the robustness of the overall effect and to prove the influence of single trials on the overall effect an exclusion sensitivity plot was calculated.

The influence of variables such as the perspective of outcome rating (self *v.* staff), setting and feedback modalities, including recipient, content, frequency and timing of feedback, on the effects of feedback interventions was explored via *a priori* defined subgroup analyses. All analyses were performed with Comprehensive Meta Analysis 2.0 for Windows and the overall effects are presented as Hedges' *g* with 95% confidence intervals (CI).[12]

Results

Search flow

The search strategy initially identified 1078 publications. After initial title and abstract screening, 68 studies were rated as potentially relevant and screened in full text. Short-listed trials that were excluded were mostly descriptive reports dealing with outcome monitoring or with theoretical aspects, development, implementation, feasibility or acceptance of feedback systems. Of the remaining trials, only 21 fulfilled the inclusion criteria. Three of them could not be retrieved even after efforts to contact the authors, one was written in a language other than English or German, and five were still in progress with no published outcome data reported.

Characteristics of included studies

Finally, 12 studies published between 2001 and 2006 met all selection criteria and were available for meta-analysis (Tables 1 and 2). Half of them were conducted in the USA, four in the UK, and two in Germany. Two studies were dissertations.[16,20] Sample size varied between 127 and 1374 participants. Most studies were conducted in out-patient settings such as community-based mental health services (five studies), university counselling centres (three studies), and specialist mental health units (old age psychiatry, eating disorder; two studies). Only two trials were performed in in-patient psychotherapeutic clinics.[16,21]

Table 1 Study properties

Properties	*n*	%
Location		
USA	6	50
UK	4	33
Germany	2	17
Setting		
In-patient	2	17
Out-patient	10	83
Design		
Randomised controlled trial	10	83
Controlled trial	2	17
Patients		
Mental illness (mixed diagnoses)	10	83
Old age	1	8
Eating disorders	1	8
Unit of randomisation		
Patients	7	58
Professionals	1	8
Patients and professionals	2	17
None	2	17

https://doi.org/10.1192/bjp.bp.108.053967 Published online by Cambridge University Press

Table 2 Characteristics of the included studies and their participants

	Country and setting	Design; intervention group/control group *n*; follow-up time points	*n*	Age, years (mean)	Female, %[a]	Illness	Outcome measures	Feedback to
Ashaye *et al* (2003)[26]	UK day hospital	RCT; 1/1; 3 months	112	⩾65 (76.4)	64	Depression, dementia	HoNOS (65+), CAPE–BRS	Staff
Bauer (2004)[16]	Germany in-patient	CT; 2/2; discharge	391	18–79 (34.9)	71	Personality disorder, depression, anxiety	SCL–11, OQ–45, IS	Clinician
Berking *et al* (2006)[21]	Germany in-patient	RCT; 1/1; discharge	118	Adults (49.4)	62	Depression, anxiety	FEP, VEV, CGI	Clinician
Brodey *et al* (2005)[23]	USA out-patient	RCT; 1/1; 6 weeks	1374	⩾18	75	Depression, anxiety	SCL–11	Clinician
Hawkins *et al* (2004)[14]	USA out-patient	RCT; 2/1; discharge	201	Adults (30.8)	68	Mood and anxiety disorders	OQ–45	Clinician or clinician and patient
Lambert *et al* (2001)[15]	USA out-patient	RCT; 1/1; discharge	609	17–57 (22.2)	70	Personal concerns	OQ–45	Clinician
Lambert *et al* (2002)[13]	USA out-patient	CT; 1/1; discharge	1020	17–57 (22.3)	70	Personal concerns	OQ–45	Clinician
Marshall *et al* (2004)[27]	UK out-patient	RCT; 2/1; 12 months	304	Adults	–	Schizophrenia, depression	BPRS, WHO–DAS	Care coordinator
Schmidt *et al* (2006)[22]	UK out-patient	RCT; 1/1; discharge and 6 months	61	Adults	100	Eating disorder	SEED	Patient
Slade *et al* (2006)[24]	UK out-patient	RCT; 1/1; 7 months	160	18–64 (41.2)	22	Schizophrenia, affective disorder	BPRS, HoNOS, TAG	Staff and patient
Trudeau (2001)[20]	USA out-patient	RCT; 1/2; 2 and 4 months	127	Adults (33.9)	72	Mental health problems	OQ–45, RAND–36	Clinician
Whipple *et al* (2003)[25]	USA out-patient	RCT; 1/1; discharge	981	18–54 (22.9)	66	Personal concerns	OQ–45	Clinician

RCT, randomised controlled trial; CT, controlled trial; HoNOS, Health of the Nation Outcome Scales (working-age adults); HoNOS (65+), Health of the Nation Outcome Scales (older adults); CAPE–BRS, Clifton Assessment Procedures for the Elderly – Behaviour Rating Scales; SCL–11, Symptom-Check-List; OQ–45, Outcome Questionnaire; IS, Impairment Score; FEP, Questionnaire to Evaluate the Course of Psychotherapy; VEV, Questionnaire to Assess Changes in Experiencing and Behavior; CGI, Clinical Global Impression; BPRS, Brief Psychiatric Rating Scale; WHO–DAS, World Health Organization Disability Assessment Schedule; SEED, Short Evaluation of Eating Disorders; TAG, Threshold Assessment Grid; RAND–36, 36-item Short Form Health Survey.
a. Dashes indicate that no data were reported.

The majority of the studies (10, 83%) were RCTs, in 7 of which the unit of randomisation was the patient. The two non-randomised trials used a controlled prospective cohort design[13] and a prospective one with sequential study phases.[16] Most of the studies included chose treatment as usual (TAU) for the control groups. One study compared feedback with a control condition that consisted of additional mental health assessment.[22] In the case of another study that reported results of two control conditions, data abstraction was restricted to the assessment-only procedure.[20]

Procedures and features of the feedback systems

Characteristics of the feedback systems used are summarised in Table 3. In most studies, feedback was given continuously at 1- or 2-weekly intervals during the course of therapy. Others limited the number of feedback interventions to a maximum of two.[23,24] In general, standardised assessments of psychological functioning (e.g. Outcome Questionnaire–45.2, Symptom Checklist–11) or needs (e.g. Camberwell Assessment of Need for the Elderly, Camberwell Assessment of Need Short Appraisal Schedule, Cardinal Needs Schedule) were completed by participants and results were fed back to therapists or staff. Only three studies investigated the provision of a (identical or different) message to patients.[14,22,24]

Feedback mainly comprised information about current treatment status and changes over time. Thus, the basic component to illustrate treatment progress was graphical and/or numerical data, often complemented with further explanations or in some cases even with treatment recommendations. Only two studies[20,21] did not use any normative comparison with a reference group to evaluate individual progress in therapy.

Effect variables

A variety of outcomes were evaluated: mental health, met and unmet needs, physical impairment, social functioning, quality of life, patient satisfaction, acceptance or appraisal of feedback, as well as rates of significant clinical change and of treatment response.

Nine of the studies measured mental health status by using patient-reported generic symptom scales. Professional-rated measures of psychological functioning were used in five trials whereas three of them included both raters' perspectives. In half of the studies two or more mental health outcome questionnaires were administered (Table 2).

The parameter most frequently used in order to ascertain savings of direct costs as a result of the use of feedback was treatment duration in terms of number of treatment sessions[13–15,25] or length of stay.[16] Only one study relied on other indicators of service use (hospital admissions and drop-out rates).[24]

Short- and long-term effects on mental health outcomes

Most of the studies included measured the impact of feedback on individuals' mental health status at the end of treatment or within 9 weeks after intake. Two of these nine trials also

https://doi.org/10.1192/bjp.bp.108.053967 Published online by Cambridge University Press

Table 3 Feedback systems used[a]

	Measures used for feedback	Rated by	Modalities of feedback			
			Timing[b]	Frequency	Content	Features
Ashaye *et al* (2003)[26]	CANE	Clinician	–	Once	Unmet needs and suitable interventions	List
Bauer (2004)[16]	OQ–45, GBB, CSC, PAE, HAQ, FLZ	Patient	Timely	Once	Phase 1: Early treatment response; Phase 2: Additionally prognosis of treatment duration, treatment recommendation	Graphs, tables, text[c,d]
	HAQ, ICD–10	Clinician				
Berking *et al* (2006)[21]	FEP	Patient	Timely	Weekly	Progress, goal achievement	Progress graph, percentage of goal achievement
Brodey *et al* (2005)[23]	SCL–11	Patient	Timely	Twice	Progress, extreme answers, fill-in time	Graph, tables, text[c]
Hawkins *et al* (2004)[14]	OQ–45	Patient	Delayed	Weekly	Progress, treatment recommendation	Graph, coloured progress markers, text[c]
Lambert *et al* (2001)[15]	OQ–45	Patient	Delayed	Weekly	Progress, treatment recommendation	Graph, coloured progress markers, text[c]
Lambert *et al* (2002)[13]	OQ–45	Patient	Delayed	Weekly	Progress, treatment recommendation	Graph, coloured progress markers, text[c]
Marshall *et al* (2004)[27]	CNS	Patient	Delayed	Once	Needs, required interventions, access to intervention	Text
Schmidt *et al* (2006)[22]	TREAT–EAT, SEED, HADS	Patient	–	Biweekly	Physical and psychological status, variables facilitating or hindering change	–
Slade *et al* (2006)[24]	CANSAS, HAS, MANSA	Patient	Delayed	Twice	Progress, areas of disagreement	Graphs, text
	CANSAS, HAS, TAG	Staff				
Trudeau (2001)[20]	OQ–45	Patient	Delayed	Weekly	Progress	Graphs[e]
Whipple *et al* (2003)[25]	OQ–45	Patient	Delayed	Weekly	Progress and treatment response; CST for non-responders	Graph, coloured progress markers, text[c,f]
	CST based upon Haq–II, SCS, MSPSS	Patient	Delayed	–	Treatment guidelines, CST results	Text

CANE, Camberwell Assessment of Need for the Elderly; OQ–45, Outcome Questionnaire; GBB, Physical Complaints Questionnaire; CSC, Client Satisfaction Questionnaire; PAE, Treatment progress scale; HAQ, Helping Alliance Questionnaire (German version); FLZ, Life Satisfaction Questionnaire; FEP, Questionnaire to Evaluate the Course of Psychotherapy; SCL–11, Symptom Checklist; CNS, Cardinal Needs Schedule; TREAT–EAT, Outcome monitoring system; SEED, Short Evaluation of Eating Disorders; HADS, Hospital Anxiety and Depression Scale; CANSAS,Camberwell Assessment of Need Short Appraisal Schedule; HAS: Helping Alliance Scale; MANSA: Manchester Short Assessment of Quality of Life; TAG, Threshold Assessment Grid; CST, Client Support Tools (decision tree of possible interventions tailored to severity of impairment); Haq–II: Revised Helping Alliance Questionnaire; SCS, Stages of Change Scale; MSPSS, Multidimensional Scale of Perceived Social Support.
a. Dashes indicate that no data were reported.
b. Timely: within 1 week after assessment; Delayed: more than 1 week thereafter.
c. Reliable Change Index (RCI) and cut-offs between functional and dysfunctional based on Jacobson & Truax.[35]
d. Based on algorithms of the 'Stuttgart–Heidelberg model'.[16]
e. Community cut-off points.
f. Algorithms based on patient's intake impairment level and on the specific session change score.

evaluated long-term effects.[20,22] The remaining three studies[24,26,27] reported follow-up measurements only (Table 2).

In the meta-analysis of short-term outcomes, ten group comparisons with a total of 4009 participants could be included. Figure 1 shows the estimated mean effect sizes. The overall effect was $d = 0.10$ (95% CI 0.01–0.19) favouring the feedback intervention. Heterogeneity across studies was moderate ($I^2 = 31\%$, $P = 0.16$). There was no substantial change in results ($d = 0.11$, 95% CI 0.02–0.20, $P = 0.02$) when limiting this analysis to those trials[13–16,20,25] with the outcome measure used most frequently (Outcome Questionnaire–45.2).[28]

For the examination of long-term effects, five trials with a total of 573 participants were included in the analysis. All except one study[20] used diverse staff-rated outcome measures that were combined per trial for effect size calculations. As shown in Fig. 2 the mean weighted effect size ($d = -0.06$, 95% CI 0.22–0.11) was against the expected direction, but it was small and not significant. Heterogeneity between studies was very low ($I^2 = 0\%$, $P = 0.69$).

Effects on treatment duration

Data on length of treatment as an indicator of costs were available for 981 participants from five studies (42%). Figure 3 shows the results of the pooled meta-analysis. The overall effect size was $d = 0.05$ (95% CI −0.05 to 0.15) with moderate heterogeneity ($I^2 = 42.03\%$, $P = 0.12$).

Moderator analyses

There were no statistically significant differences in effect sizes between any of the subgroups analysed (Table 4). However, magnitude of subgroup differences was substantial in a number of instances. Feedback was more effective if it: was given to both patients and therapists or staff (*v.* only to therapists/staff); was reported at least twice (*v.* only once); comprised information on patient progress (*v.* on status only). Moreover, only the latter two subgroups showed effect sizes significantly different from zero within their subgroup analysis.

Heterogeneity within the clusters varied between very low and high particularly depending on number of included studies or group comparisons (I^2 ranged from 0 to 84%).

Publication bias and sensitivity analysis

Funnel plots for studies of short- (Fig. 4) or long-term mental health outcome as well as for studies evaluating treatment duration did not provide evidence of publication bias. Fail-safe numbers for the analyses pertaining to short-term effects showed that it would require 11 unpublished null result trials to nullify the

https://doi.org/10.1192/bjp.bp.108.053967 Published online by Cambridge University Press

Study name	Hedges' *g*	Lower limit	Upper limit	*P*	Intervention	Control
Bauer (2004) (phase 3)[16]	−0.12	−0.38	0.14	0.36	109	121
Bauer (2004) (phase 4)[16]	−0.03	−0.30	0.25	0.84	99	101
Berking *et al* (2006)[21]	0.45	0.06	0.85	0.02	51	53
Brodey *et al* (2005)[23]	0.09	−0.04	0.21	0.19	467	487
Hawkins *et al* (2004)[14]	0.14	−0.20	0.48	0.43	69	64
Lambert *et al* (2001)[15]	0.24	−0.12	0.61	0.19	154	151
Lambert *et al* (2002)[13]	0.21	0.09	0.33	0.00	528	492
Schmidt *et al* (2006)[22]	0.37	−0.31	1.06	0.28	17	16
Trudeau (2001)[20]	−0.09	−0.73	0.55	0.79	38	12
Whipple *et al* (2003)[25]	0.02	−0.11	0.14	0.78	499	482
Overall	**0.10**	**0.01**	**0.19**	**0.03**	**2030**	**1979**

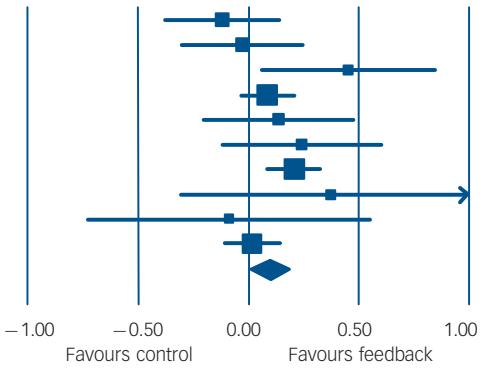

Fig. 1 Forest plot for short-term mental health outcome (random-effects model).[a]

a. Treatment outcomes were coded as short-term if they were measured within 9 weeks after initial assessment.

Study name	Hedges' *g*	Lower limit	Upper limit	*P*	Intervention	Control
Ashaye *et al* (2003)[26]	−0.11	−0.47	0.26	0.57	54	58
Marshall *et al* (2004)[27]	−0.03	−0.28	0.22	0.80	127	115
Schmidt *et al* (2006)[22]	0.17	−0.44	0.79	0.58	21	18
Slade *et al* (2006)[24]	−0.03	−0.37	0.32	0.88	93	49
Trudeau (2001)[20]	−0.50	−1.20	0.19	0.16	27	11
Overall	**−0.06**	**−0.22**	**0.11**	**0.49**	**322**	**251**

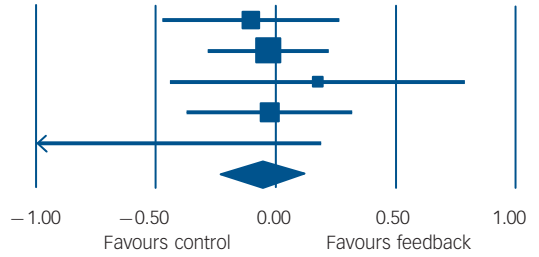

Fig. 2 Forest plot for long-term mental health outcome (random-effects model).[a]

a. Treatment outcomes were coded as long-term if they were measured between 3 and 12 months after initial assessment.

Study name	Hedges' *g*	Lower limit	Upper limit	*P*	Intervention	Control
Bauer (2004) (phase 3)[16]	0.09	−0.17	0.34	0.51	109	121
Bauer (2004) (phase 4)[16]	0.39	0.11	0.67	0.01	99	101
Hawkins *et al* (2004)[14]	0.09	−0.25	0.43	0.60	69	64
Lambert *et al* (2001)[15]	0.04	−0.12	0.20	0.63	307	302
Lambert *et al* (2002)[13]	−0.05	−0.17	0.07	0.44	528	492
Whipple *et al* (2003)[25]	−0.00	−0.13	0.12	0.97	499	482
Overall	**0.05**	**−0.15**	**0.15**	**0.33**	**1611**	**1562**

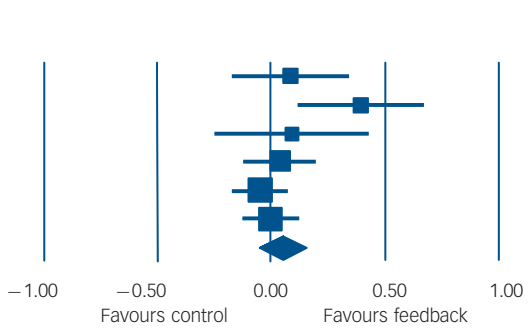

Fig. 3 Forest plot for treatment duration (random-effects model).

calculated effect. Moreover, Egger's regression test was not significant ($P=0.76$), suggesting no evidence of small-study bias.

Further sensitivity analyses revealed that the exclusion of any single study only marginally changed the overall effect sizes of all three analyses. The greatest difference in short-term outcome analysis occurred when the study of Lambert *et al*[13] was removed ($d=0.06$, 95% CI −0.02 to 0.15, $P=0.14$).

Discussion

Evidence base

Twelve studies fulfilling inclusion criteria for a systematic meta-analysis of the effectiveness of outcome management in mental healthcare were identified. Most of these were conducted in out-patient treatment settings in the USA or the UK. Outcome management in in-patient treatment has only been examined in Germany, which also reflects national differences in mental health service provision.[29] Sample size averaged 378 participants per study. Altogether, evidence is based upon 4540 adults, two-thirds of them female. The vast majority of study participants were people with affective disorders. However, participants' mental health problems did not always meet standard diagnostic criteria of a mental disorder, especially among student populations on which most of the research of the Lambert group is based. On the other hand, some studies focused on people with severe mental illness (e.g. Slade *et al*).[24] Given that two decades have passed since the introduction of the general principles of outcome

https://doi.org/10.1192/bjp.bp.108.053967 Published online by Cambridge University Press

Table 4 Results of subgroup analyses (random-effects model)

Subgroup	Short-term mental health outcome			
	k^a	Effect size	95% CI	P^b
Setting				
In-patient	3	0.07	−0.23 to 0.37	0.65
Out-patient	7	0.11	0.04 to 0.18	0.00
Rater of outcome				
Patient	11	0.11	0.03 to 0.19	0.01
Staff	3	0.01	−0.42 to 0.44	0.97
Feedback recipient				
Therapist/staff and patient	2	0.30	−0.01 to 0.60	0.06
Therapist/staff	9	0.09	−0.01 to 0.18	0.07
Feedback content				
Status information	3	−0.04	−0.23 to 0.14	0.63
Progress information	7	0.13	0.04 to 0.22	0.01
Feedback timing[c]				
Timely	4	0.06	−0.04 to 0.16	0.23
Delayed	5	0.06	−0.05 to 0.17	0.31
Feedback frequency				
Once	2	−0.08	−0.27 to 0.11	0.42
More than once	8	0.13	0.04 to 0.22	0.00

a. *k*, number of included group comparisons (weighted mean effect sizes).
b. $P<0.05$ indicates an effect size significantly different from zero.
c. Timely: within 1 week after assessment; delayed: more than 1 week thereafter.

management,[1] this suggests that some efforts have been made to provide systematic evidence of its effectiveness in specialist mental healthcare.

Short- and long-term effect

Results of this review of controlled studies show that feedback of outcome had a small, albeit statistically significant short-term effect on improving mental health outcomes ($d=0.10$, 95% CI 0.01–0.19). This effect was found to be consistent across a variety of outcome measures as analyses limited to studies using a popular single instrument (Outcome Questionnaire–45.2) or excluding the one study with a disorder-specific scale[22] showed similar results. Furthermore, sensitivity analyses revealed that the exclusion of any single study only marginally changed the overall effect sizes of all three analyses, implying relative stability of the results. This corresponds with the calculation of fail-safe numbers indicating that at least 11 null effect trials are needed to diminish the effect. This effect size is similar to that found in other reviews on the topic, for example by Sapyta *et al*[10] on various forms of feedback of client health status information (including screening information only) to health professionals (including general practitioners) in community settings ($d=0.21$) or by Lambert *et al*[7] on continuous feedback of outcome in out-patient psychotherapy ($d=0.09$). Bearing in mind differences in sampling and methods, such comparisons should be viewed cautiously.

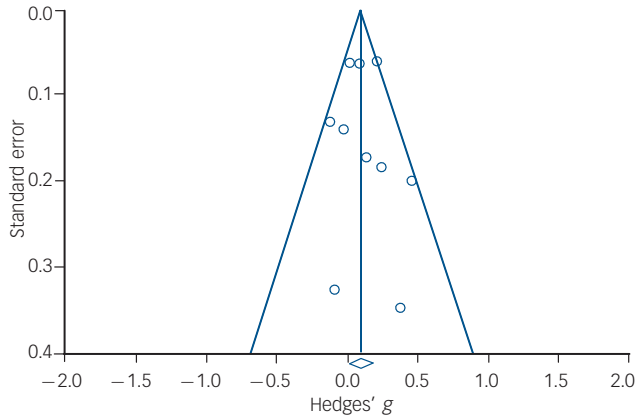

Fig. 4 Funnel plot of estimated effect sizes (Hedges' *g*) against standard error of the trials for short-term mental health outcome.

It was also found that this effect did not prevail in the long run, i.e. after termination of treatment and outcome feedback, symptomatic impairment in study participants allocated to control groups was no worse than in those who had received outcome management. On the one hand, this finding suggests that outcome management has no persistent effect on improving mental health. On the other, it could also be interpreted to suggest that it might be wise to continue using outcome feedback at least to some extent after the end of treatment in order to avoid wearing off of short-term benefits. Unfortunately, a clear interpretation of this finding is difficult since long-term effects have only rarely been studied (in only five trials).

Furthermore, no advantage of outcome management was found with respect to reduction of treatment costs or cost-offset. This finding should also be interpreted with caution since only very few researchers made efforts to collect comprehensive cost data. Therefore, statistical pooling was limited to studies reporting indicators of the amount of treatment utilised.

Moderators of effect

In addition, apart from specifying the general benefit of outcome management, it is also important to identify its active ingredients. Moderator analyses showed that the effect of outcome management on individuals' short-term mental health outcome is improved if: feedback comprises information on mental health progress (*v.* only on status); both patient and therapist receive feedback (*v.* only one of them); and feedback is given frequently (*v.* only once). Considering the small numbers of studies and substantial heterogeneity for some clusters (I^2 ranging from 0 to 84%), these results offer valuable clues but should be interpreted with caution.

https://doi.org/10.1192/bjp.bp.108.053967 Published online by Cambridge University Press

Limitations

Even though the aggregated sample size of included studies would appear sufficient, the number of studies these data were drawn from was small and not particularly representative. Additionally, the number of researchers conducting this research is very small, and could introduce bias through a number of avenues. This meta-analysis has a number of further drawbacks. First, assessed studies varied considerably with regard to certain patient characteristics, most notably illness severity and comorbid disorders. The majority of trials relied on data from people with rather mild mental illness treated with out-patient psychotherapy. Second, the number of studies decreased rapidly when we tried to answer more specific questions (e.g. long-term effects, moderators). As a result of variations in study designs, measurement points had to be pooled in order to be able to examine persistence of effect. Third, except in one study, feedback was based on patient-reported outcomes whereas it might be worthwhile to also have data obtained from independent raters. Fourth, in the vast majority of the trials, the unit of randomisation was the patient, which might entail cross-contamination,[8] i.e. 'real' effects of the intervention are watered down because of a general sensitisation since one clinician receives feedback for some patients but not for others.

Implications and outlook

Psychotherapy research has put forth a number of rather refined approaches to systematically feeding back patient outcome data to clinicians and also to patients in order to improve quality of mental healthcare, e.g. patient-focused research,[30] expected treatment response[31] or computer-assisted quality management.[32] All these outcome management systems heavily rely on professionals' endorsement and use of standardised outcome measures. Although there is a trend towards increased use of such instruments among psychologists in clinical practice,[33] psychiatrists have hardly picked up on this development.[34] A reason for this might be doubts about the benefits of outcome management among clinicians.

This review might contribute to reducing such doubts at least to some extent, since it has shown that feedback of outcome improves symptomatic impairment in people with mental illness treated in various specialist mental health services. Even though overall effect size was modest, some single studies yielded impressive effects. However, the evidence base is hardly sufficient to adequately analyse crucial more specific questions such as differential effectiveness in certain subgroups (e.g. by diagnosis). Nevertheless, this review provides some clues on issues that should be taken into account when implementing outcome management in routine practice, i.e. to feedback outcome continuously and regularly to both clinicians and patients, and to provide information on treatment progress.

Future research might pick up on these clues in order to move beyond trying to answer the question of a general effect of outcome management in mental healthcare. One direction should be to conduct much-needed comprehensive cost-effectiveness analyses based upon sound estimates of costs of feedback (staff, equipment, space and other resource costs), direct and indirect costs and treatment outcome. Most importantly, active ingredients of outcome management should be scrutinised in more detail in order to arrive at feedback systems tailored to the needs of specific subgroups of people with mental illness, but also to the requirements of clinicians who, at least during kick-off, are asked to devote substantial amounts of extra effort to make outcome management work.

Carina Knaup, Dipl-Psych, **Markus Koesters**, Dipl-Psych, **Dorothea Schoefer**, Dipl-Psych, **Thomas Becker**, MD, **Bernd Puschner**, PhD, Department of Psychiatry II, Ulm University, Ulm, Germany

Correspondence: Bernd Puschner, Ulm University, Department of Psychiatry II, Ludwig-Heilmeyer-Str. 2, 89312 Günzburg, Germany. Email: bernd.puschner@bkh-guenzburg.de

First received 21 Apr 2008, final revision 11 Sep 2008, accepted 10 Dec 2008

Funding

This review was carried out during the course of the study 'Outcome monitoring and outcome management in inpatient psychiatric care' (EMM), which received funding from the German Federal Ministry of Education and Research (grant number 01GL0504).

Acknowledgements

We thank two anonymous reviewers for their knowledgeable and helpful comments.

References

1 Ellwood PM. Shattuck Lecture – Outcomes management: a technology of patient experience. *N Engl J Med* 1988; **318**: 1549–56.

2 Lyons JS, Howard KI, O'Mahoney MT, Lish JD. *The Measurement and Management of Clinical Outcomes in Mental Health*. Wiley, 1997.

3 Kordy H, Bauer S. The Stuttgart-Heidelberg model of active feedback driven quality management: means for the optimization of psychotherapy provision. *Int J Clin Health Psych* 2003; **3**: 615–31.

4 Espallargues M, Valderas JM, Alonso J. Provision of feedback on perceived health status to health care professionals: a systematic review of its impact. *Med Care* 2000; **38**: 175–86.

5 Valderas JM, Kotzeva A, Espallargues M, Guyatt G, Ferrans CE, Halyard MY, et al. The impact of measuring patient-reported outcomes in clinical practice: a systematic review of the literature. *Qual Life Res* 2008; **17**: 179–93.

6 Gilbody SM, House AO, Sheldon TA. Routinely administered questionnaires for depression and anxiety: systematic review. *BMJ* 2001; **322**: 406–9.

7 Lambert MJ, Whipple JL, Hawkins EJ, Vermeersch DA, Nielsen SL, Smart DW. Is it time for clinicians to routinely track patient outcome? A meta-analysis. *Clin Psychol Sci Pr* 2003; **10**: 288–301.

8 Gilbody SM, House AO, Sheldon TA. Outcome measures and needs assessment tools for schizophrenia and related disorders. *Cochrane Database Syst Rev* 2003; issue 1: CD003081.

9 Marshall S, Haywood K, Fitzpatrick R. Impact of patient-reported outcome measures on routine practice: a structured review. *J Eval Clin Pract* 2006; **12**: 559–69.

10 Sapyta J, Riemer M, Bickman L. Feedback to clinicians: theory, research, and practice. *J Clin Psychol* 2005; **61**: 145–53.

11 Hedges LV. Distribution theory for Glass's estimator of effect size and related estimators. *J Educ Stat* 1981; **6**: 107–28.

12 Hedges LV, Oklin I. *Statistical Methods for Meta-analysis*. Academic Press, 1985.

13 Lambert MJ, Whipple JL, Vermeersch DA, Smart DW, Hawkins EJ, Nielsen SL, et al. Enhancing psychotherapy outcomes via providing feedback on client progress: a replication. *Clin Psychol Psychoth* 2002; **9**: 91–103.

14 Hawkins EJ, Lambert MJ, Vermeersch DA, Slade KL, Tuttle KC. The therapeutic effects of providing patient progress information to therapists and patients. *Psychother Res* 2004; **14**: 308–27.

15 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? *Psychother Res* 2001; **11**: 49–68.

16 Bauer S. *Outcome Monitoring and Feedback: Strategies for Optimisation of Inpatient Psychotherapy?* (in German). Doctoral Thesis (PhD), Faculty of Social and Behavioural Sciences, University of Tübingen, Germany, 2004.

17 DerSimonian R, Laird N. Meta-analysis in clinical trials. *Control Clin Trials* 1986; **7**: 177–88.

18 Higgins JPT, Thompson SG, Deeks JJ, Altmann GG. Measuring inconsistency in meta-analyses. *BMJ* 2003; **327**: 557–60.

19 Egger M, Davey Smith G, Schneider M, Minder CE. Bias in meta-analysis detected by a simple, graphical test. *BMJ* 1997; **315**: 629–34.

20 Trudeau LS. *Effects of a Clinical Feedback System on Client and Therapist Outcomes in a Rural Community Mental Health Center*. Doctoral Thesis

https://doi.org/10.1192/bjp.bp.108.053967 Published online by Cambridge University Press

(PhD), Institute for Social and Behavioral Research, Iowa State University, USA, 2001.

21 Berking M, Orth U, Lutz W. Effects of systematic feedback to the therapist on patient progress. An empirical study in a cognitive-behavioral inpatient setting (in German). *Z Kl Psych Psychot* 2006; **35**: 21–9.

22 Schmidt U, Landau S, Pombo-Carril MG, Bara-Carril N, Reid Y, Murray K, et al. Does personalized feedback improve the outcome of cognitive-behavioural guided self-care in bulimia nervosa? A preliminary randomized controlled trial. *Brit J Clin Psychol* 2006; **45**: 111–21.

23 Brodey BB, Cuffel B, McCulloch J, Tani S, Maruish M, Brodey I, et al. The acceptability and effectiveness of patient-reported assessments and feedback in a managed behavioral healthcare setting. *Am J Manag Care* 2005; **11**: 774–80.

24 Slade M, McCrone P, Kuipers E, Leese M, Cahill S, Parabiaghi A, et al. Use of standardised outcome measures in adult mental health services. Randomised controlled trial. *Br J Psychiatry* 2006; **189**: 330–6.

25 Whipple JL, Lambert MJ, Vermeersch DA, Smart DW, Nielsen SL, Hawkins EJ. Improving the effects of psychotherapy: the use of early identification of treatment and problem-solving strategies in routine practice. *J Couns Psychol* 2003; **50**: 59–68.

26 Ashaye OA, Livingston G, Orrell MW. Does standardized needs assessment improve the outcome of psychiatric day hospital care for older people? A randomized controlled trial. *Aging Ment Health* 2003; **7**: 195–9.

27 Marshall M, Lockwood A, Green G, Zajac-Roles G, Roberts C, Harrison G. Systematic assessments of need and care planning in severe mental illness: cluster randomised controlled trial. *Br J Psychiatry* 2004; **185**: 163–8.

28 Lambert MJ, Burlingame GM, Umphress V J, Hansen N B, Vermeersch D A, Clouse G, et al. The reliability and validity of the Outcome Questionnaire. *Clin Psychol Psychoth* 1996; **3**: 106–16.

29 Becker T, Vazquez-Barquero JL. The European perspective of psychiatric reform. *Acta Psych Scand* 2001; **104**: 8–14.

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

31 Lueger RJ, Howard KI, Martinovich Z, Lutz W, Anderson EE, Grissom G. Assesssing treatment progress of individual patients using expected treatment response models. *J Consult Clin Psych* 2001; **69**: 150–8.

32 Kordy H, Hannöver W, Richard M. Computer-assisted feedback-driven quality management for psychotherapy: the Stuttgart-Heidelberg model. *J Consult Clin Psych* 2001; **69**: 173–83.

33 Hatfield DR, Ogles BM. The use of outcome measures by psychologists in clinical practice. *Prof Psychol Res Pr* 2004; **35**: 485–91.

34 Gilbody SM, House AO, Sheldon TA. Psychiatrists in the UK do not use outcome measures. National survey. *Br J Psychiatry* 2002; **180**: 101–3.

35 Jacobson NS, Truax P. Clinical significance: a statistical approach to defining meaningful change in psychotherapy research. *J Consult Clin Psych* 1991; **59**: 12–9.

Psychiatrists in 19th-century fiction

The Yellow Wallpaper (1892), Charlotte Perkins Gilman

Fiona Subotsky

The narrator of *The Yellow Wallpaper* has been taken by her medical husband to 'a colonial mansion, a hereditary estate', perhaps 'a haunted house', to recover from 'a temporary nervous depression – a slight hysterical tendency'. She has been forbidden to write, though she does so secretly, and is increasingly confined to her bedroom at the top of the house, an old nursery with barred windows, a heavy bed fastened to the floor and badly damaged yellow wallpaper with a strange pattern and smell which becomes increasingly hideous yet fascinating to her. Weeks go by. She seems to see behind the pattern a woman, or many women, creeping, as if behind bars, and tries to release her or them by creeping round the walls herself and stripping off the paper. The ending is both shocking and uncertain, in itself deranged.

Where is the psychiatrist in this? The narrator's husband says that if she does not improve, he will send her 'to Weir Mitchell in the fall'. Silas Weir Mitchell was in fact an eminent contemporary physician in the United States, author both of scientific books and papers and also of a considerable amount of fiction. He had treated Charlotte Perkins Gilman for depression, in her view disastrously. He was especially famous for his 'rest cure', which required his patients, usually women considered to be suffering from hysterical nervous complaints, to be confined to bed and do nothing exciting at all but passively undergo electrical and massage treatments and consume a large quantity of food. He described his method in *Fat and Blood, and How to Make Them* (1877). While highly fashionable for a while, by 1896 our own *Journal* concluded that this treatment 'has frequently proved a most conspicuous failure'.

Charlotte Gilman, however, did escape her confines and became an active feminist lecturer and writer; her novel has become a feminist classic.

The British Journal of Psychiatry (2009) 195, 22. doi: 10.1192/bjp.195.1.22

https://doi.org/10.1192/bjp.bp.108.053967 Published online by Cambridge University Press