

Comparison of short- and long-term dynamic group psychotherapy: randomised clinical trial

Steinar Lorentzen, Torleif Ruud, Anette Fjeldstad and Per Høglend

Background

There are no randomised clinical trials comparing the outcomes of short- with long-term psychodynamic group psychotherapy.

Aims

To compare differences in outcome during and after short- and long-term group psychotherapy.

Method

In total, 167 out-patients with mood, anxiety and personality disorders were randomised to short- or long-term group therapy (20 or 80 weekly, 90 min sessions). Outcome measures were: symptoms (Symptom Checklist 90 – Revised), interpersonal problems (Inventory of Interpersonal Problems – Circumplex) and psychosocial functioning (Global Assessment of Functioning (GAF) split version: GAF-Symptom and GAF-Function). Change over the 3-year study period was

assessed using linear mixed models. The study was registered in clinicalTrials.gov as NCT00521417.

Results

Patients in both groups made significant gains. A significantly larger symptomatic change over time was found for long-term compared with short-term therapy, but no significant differences were detected for the three remaining outcome variables. There was a higher number of premature terminations in the long-term (33.3%) compared with the short-term group (8.6%).

Conclusions

Short- and long-term therapy seem equally effective for typical out-patients seeking group psychotherapy, except for symptomatic distress.

Declaration of interest

None.

Meta-analyses and reviews indicate that there are few differences in outcome between group therapies with different theoretical rationale.^{1,2} A problem with the existing research is that most group therapies studied are of short duration (<20 sessions),² although therapies lasting 1 year or more are quite common in clinical practice, at least in some European countries.^{3,4} There are only a few studies that demonstrate the effectiveness of long-term dynamic group psychotherapy with out-patients.^{5–7} Reviews of non-randomised studies on the importance of treatment duration^{8,9} have failed to give unequivocal answers. Several studies have indicated that short-term therapies offer insufficient help to many patients, such as those with personality disorders and patients with comorbid and chronic disorders.^{10–14} We do not know of any randomised studies on the effects of duration of group treatment. Since all observational studies are open to several interpretations,¹⁵ studies of patients randomised to short- v. long-term group therapy are urgently needed. Most research is restricted to patients with a single diagnosis. However, a substantial portion of out-patients meet criteria for several Axis I and Axis II diagnoses, and have to cope with a number of difficulties relating to long-term suboptimal functioning such as interpersonal problems, work problems, reduced physical health and higher overall mortality.¹⁶ Thus, we wanted to study out-patients within the public mental health services, who in general have heterogeneous and comorbid psychiatric disorders. In the present study, 6 months of group psychotherapy (20 weekly sessions) is compared with 2 years of group psychotherapy (80 weekly sessions). Over the 3-year study period outcomes relating to symptom severity, interpersonal problems and psychosocial functioning were measured. Based on the limited existing research, our main hypotheses were that symptom change would be similar in short- and long-term group psychotherapy, although patients would improve more on measures of interpersonal problems and psychosocial functioning in the long-term compared with the short-term group.

Method

Sites, patient referral and inclusion

This study was conducted in three urban areas in Norway (Ålesund, Sandnes/Stavanger and Oslo). One to two coordinators and two to four therapists were engaged at each site. The project was approved by The Data Inspectorate and The Regional Committee on Ethics in Health Research. The study was registered in clinicalTrials.gov as NCT00521417 and data collection took place between August 2005 and August 2010.

Community mental health centres, general practitioners and practising psychiatrists and psychologists were informed by mail of the clinical trial. They referred patients they considered to be in need of assistance from mental health specialist services to the local coordinators, who evaluated them for eligibility. Inclusion criteria were one or more Axis I and/or Axis II diagnoses, interest in working with problems and relationships in groups and willingness to accept randomisation. Exclusion criteria were psychosis, a main diagnosis of alcohol/drug addiction and organic brain disease. Patients with more than 6 months of previous group therapy or former psychotherapy that had ended within the past year were also excluded. Written informed consent was obtained from all participants.

In total, 175 patients were referred to the coordinators and 167 eligible patients were randomly allocated to the two treatments. The randomisation was carried out for each site by the local coordinator and was not influenced by the therapists.¹⁷ One short-term and one long-term group were gradually built simultaneously, with at least two women and two men in each group to secure mixed groups regarding gender (see online supplement DS1). The time from referral to evaluation/randomisation varied from 1 to 7 weeks. Time from randomisation to start of therapy varied from 2 to 12 weeks, with an average of 7 weeks. The coordinators were not involved in the treatment. The randomisation code was not opened until all data from the project had been collected.¹⁷

Groups, therapies and therapists

Eighteen psychotherapy groups, nine short-term and nine long-term groups, both with 90 min weekly sessions, were established. Our choice of therapy length was a compromise between enough time for the therapy to be effective and how long we anticipated most patients would be willing to participate. Each group consisted of eight patients and one therapist. Patients who terminated in the long-term groups during the first 6 months ($n = 7$ patients) were, according to protocol, replaced by the next patient of the same gender who was included in the project.

Both therapy formats were manualised psychodynamic group psychotherapies,¹⁸ built on psychodynamic understanding: a developmental perspective on personality, existence of internal representations of interpersonal relationships, psychological causation and influence of unconscious individual and group processes on behaviour. The therapist could use facilitating or supportive techniques to promote interaction, in combination with traditional psychodynamic techniques of confrontation and interpretation.¹⁹ Group members were asked to interact and to focus on conscious as well as potential derivatives of unconscious processes in themselves, others and the group (see online supplement DS1). Thus, the treatments studied are not primarily targeted to specific symptom disorders. In the short-term group, therapists were to be more active, to have a more circumscribed problem focus, to work more in the here-and-now and to be more attentive of the impending termination phase.

Among the nine therapists (two men and seven women), there were two psychiatrists, three psychologists, three psychiatric nurses and one social worker. Their mean age was 52.7 years (s.d. = 3.7), time in practice 19.7 years (s.d. = 4.4) and formal postgraduate psychotherapy training 12.5 years (s.d. = 3.7). All therapists ran one short- and one long-term group. They were trained in both formats, and met regularly for supervision.

Assessment

The pre-randomisation evaluation (2–4 h) included a full psychiatric history, the diagnostic interviews Mini-International Neuropsychiatric Interview (MINI-PLUS),²⁰ Structured Clinical Interview for DSM-IV Axis II personality disorders (SCID-II)²¹ and Global Assessment of Functioning (GAF).²² The patients also completed self-reports, the Symptom Checklist 90 – Revised (SCL-90-R)²³ and the Inventory of Interpersonal Problems – Circumplex (IIP-C).²⁴

After randomisation, the patients met with their therapist individually for two sessions and received more information about the group and negotiated a treatment contract. The self-report outcome measures (SCL-90-R and IIP-C) were completed every 6 months for both groups as long as the long-term therapy was running and 12 months after long-term therapy termination, for all patients. Self-reports were also sent to individuals who did not start therapy and those who stopped attending prematurely. Selected sessions from both therapies (early, mid- and end phase) were audio-recorded.

The patients were interviewed again by the same coordinators 3 years after the beginning of therapy and diagnosed on Axes I and II, using MINI-PLUS and SCID-II respectively, and rated on GAF. According to the research protocol, this study is an ‘open label’ design. However, the coordinators were unlikely to remember the original allocation of the patients 3 years earlier, and the patients were instructed to conceal which group therapy they had attended.

Outcome measures

Since most patients had a high comorbidity and chronicity, we selected outcome measures for severity of symptoms, interpersonal functioning and psychosocial functioning.

Symptom Checklist 90 – Revised

This covers 90 items of different symptoms, rated from ‘0, not at all’ to ‘4, could not be worse’. The Global Severity Index (GSI) is the mean of all items. The instrument is well validated and the internal consistency was 0.96 in this study.

Inventory of Interpersonal Problems – Circumplex

The IIP-C is well validated and one of the most widely used self-report measures in psychotherapy research. Sixty-four items are rated on the same scale as the SCL-90-R. Internal consistency was 0.94 in this study. The IIP has demonstrated high test–retest stability across 10 weeks, but is still sensitive to change.²⁵

Global Assessment of Functioning

The GAF (Axis V, DSM-IV)²² was used: severity of psychopathology and level of social functioning were assessed using the split version of GAF,²⁶ with separate scores for symptom severity (GAF-S) and problems in functioning (GAF-F). Intraclass correlation (ICC) among five evaluators rating 20 cases was 0.76 for GAF-S and 0.80 for GAF-F.²⁷

Target Complaint

Target Complaint²⁸ is a self-report measure of the problems for which a person seeks therapy. The patient describes the primary problem and rates the severity on a scale ranging from ‘1, no problem’ to ‘12, could not be worse’. The duration of the target complaint was also recorded, as well as the expectation of change (from ‘1, change to the worse’ to ‘12, go away completely’), which was the main objective for using this measure in this study.

Statistical analysis

With a sample of 120 patients, a standard power calculation (end-point analysis) indicated that a moderate effect size (0.5) could be detected for an alpha threshold of 0.05 with a power of 0.80.²⁹ Our aim was to include 170 patients, to ensure an adequate sample size of 120 completers, in case of a substantial drop-out rate. Missing data were imputed using last observation carried forward (LOCF). Imputation was not done for longitudinal analyses, using linear mixed models (LMM).³⁰

Linear mixed models allow non-independence of repeated and nested data. It can handle incompleteness because of missing data. In order to account for non-independence of data within patients nested within groups nested within therapists, ‘patient’, ‘group’ and ‘therapist’ were each treated as random effects: randomly distributed intercepts and slopes were fitted for each patient, for each group and for each of the nine therapists. Treatment (group length) was coded long-term treatment group: 1, short-term treatment group: 0. Time was coded on an interval scale from 0 to 36 months.

To test the treatment effects over time (difference in slopes between short-term and long-term group), the following composite model equation was used:

$$Y_{ij} = B_0 + B_1 \text{time}_{ij} + B_2 (\text{time}_{ij} \times \text{grouplength}_i) + [\zeta_{0i} + \zeta_1 \text{time}_{ij} + \varepsilon_{ij}]$$

Y_{ij} is change in GSI, IIP-C, GAF-S or GAF-F over the 3-year study period. B_0 – B_2 are the fixed effects, and ζ_{0i} , $\zeta_1 \text{time}_{ij}$ and ε_{ij} are random intercept, random time and the error term respectively. Random effects for groups and for therapists were also tested. By design, treatment group means are equal at baseline. The statistical model forces both treatments to have a common intercept.

This model is more powerful and is routinely recommended for analysis of randomised clinical trials.³¹ The parameters are: B_0 intercept; B_1 slope (rate of change) in short-term therapy group; B_2 (time \times group length) is the difference in slopes between the long- and short-term group.

No longitudinal analyses were conducted on subgroups of patients. Effect sizes (converted to Cohen's d), derived from the F -test for the mixed effects model, were calculated as:

$$d = 2\sqrt{\frac{F}{d.f.}}$$

where F is the F -test statistic for the effect of interest in the repeated model, as well as other multilevel designs.³⁰ To explore whether or not missing data influenced the results, a pattern mixture approach was used.³² Drop-out status, defined as attending less than two-thirds of the scheduled sessions (i.e. less than 53 sessions in long-term therapy or 13 sessions in short-term therapy), was based on clinical consensus and determined before the statistical analyses were run. Clinically significant change (recovery) was computed according to the formulas outlined by Jacobson & Truax³³ and Christensen & Mendoza.³⁴

We used SPSS version 16.0 for Windows in our statistical analyses.

77 to the short-term group. However, 19 patients did not start treatment, 12 in the long- and 7 in the short-term group ($\chi^2 = 0.72$, d.f. = 1, $P = 0.47$). Reasons why these individuals did not start treatment included having second thoughts about the group project, deciding that they did not need therapy or starting alternative treatments (see online supplement DS1).

Among those who started treatment there were no differences in positive expectations about treatment between the two groups after randomisation ($t = -1.22$, d.f. = 144, $P = 0.23$, 95% CI -1.16 to 0.28). There were different rates of premature termination: 26 patients (33.3%) in the long-term group and 6 patients (8.6%) in the short-term group ($\chi^2 = 11.9$, d.f. = 1, $P < 0.001$) (online Fig. DS1). The numbers of premature terminations during the first 6 months were equal in both groups. We are uncertain whether the increased number of people dropping out of long-term therapy is a result of the difference in therapy length alone. We have information about why 26 of the 32 (81%) participants discontinued therapy prematurely: 10 thought the group was not helpful or felt they had deteriorated, 8 were dissatisfied with the therapist or the group, 4 got the help they needed and 4 left because of external events. The mean number of sessions in treatment for starters was 18.9 (s.d. = 3.8) in the short-term group and 57.7 (s.d. = 26.9) in the long-term group; an attendance rate of 94.5% and 72.1% respectively. In total, 79 of 90 patients (88%) completed the 3-year follow-up interview after long-term therapy compared with 71 of 77 patients (92%) after short-term therapy.

Results

The patient flow is shown in Fig. 1. A total of 175 patients were assessed, but 8 patients did not meet the inclusion criteria. Therefore 167 were randomised: 90 to the long-term group and

Patient characteristics at baseline

Table 1 summarises the patients' baseline characteristics for the whole sample ($n = 167$). There were significantly more patients with generalised anxiety disorder ($\chi^2 = 3.84$, d.f. = 1, $P = 0.050$)

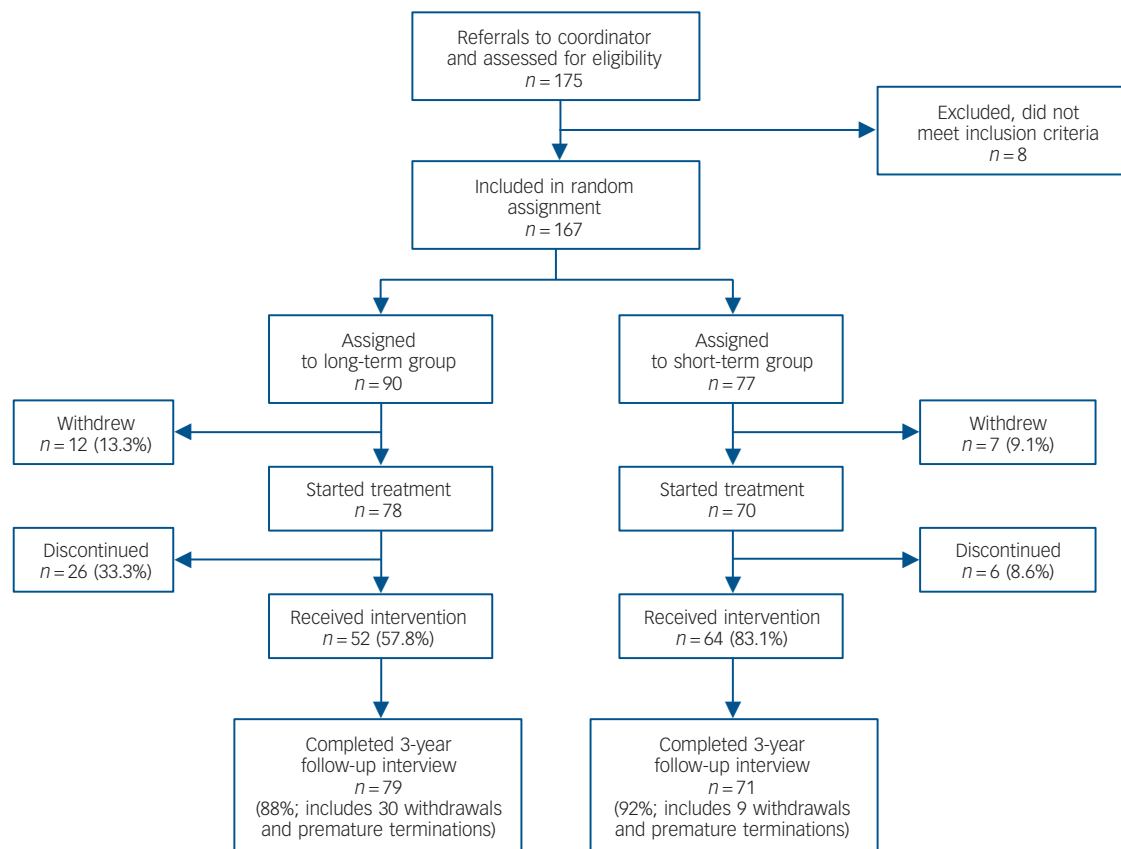


Fig. 1 Patient flow in a randomised clinical trial of short- and long-term psychodynamic group psychotherapy.

Table 1 Pretreatment characteristics of randomised patients in short- and long-term dynamic group psychotherapy ($n = 167$)^a

	Short-term therapy group ($n = 77$)	Long-term therapy group ($n = 90$)	Whole sample ($n = 167$)
Age, years: mean (s.d.)	38.6 (9.4)	38.2 (9.4)	38.4 (9.4)
Education, years: mean (s.d.)	13.8 (3.3)	13.7 (3.1)	13.7 (3.2)
Expectations about treatment, ^b mean (s.d.)	7.9 (2.4)	7.2 (2.7)	7.7 (2.6)
Number of Axis I diagnoses, mean (s.d.)	3.2 (1.9)	3.4 (2.0)	3.3 (2.0)
Structured Clinical Interview for DSM-IV Axis II personality disorders, number of positive criteria: mean (s.d.)	7.5 (7.3)	8.5 (6.4)	8.0 (6.8)
Female, n (%)	50 (65)	55 (61)	105 (63)
Marital status: single, n (%)	34 (44)	41 (46)	75 (45)
Previous hospital admissions, n (%)	7 (9)	13 (14)	20 (12)
Axis I diagnosis, n (%)			
Major depression, single	5 (7)	10 (12)	15 (10)
Major depression, recurrent	53 (69)	61 (72)	114 (68)
Dysthymia	7 (9)	3 (20)	10 (6)
Panic disorder	23 (30)	34 (40)	57 (34)
Agoraphobia	8 (10)	4 (4)	12 (7)
Obsessive-compulsive disorder	8 (10)	20 (21)	28 (17)
Social phobia	24 (31)	31 (38)	55 (33)
Generalised anxiety disorder ^c	11 (14)	28 (30)	39 (23)
Post-traumatic stress disorder	6 (8)	1 (1)	7 (4)
Somatoform disorder	12 (16)	23 (26)	35 (21)
Other	29 (38)	28 (31)	57 (34)
No Axis I diagnosis	1 (1)	3 (3)	4 (2)
Axis II diagnosis, n (%)			
Avoidant	18 (23)	29 (32)	47 (28)
Dependent	3 (4)	5 (6)	8 (5)
Obsessive-compulsive ^d	3 (4)	11 (12)	14 (8)
Paranoid	7 (9)	4 (4)	11 (7)
Borderline	5 (7)	4 (4)	9 (5)
Personality disorder – not otherwise specified	4 (5)	9 (10)	13 (8)
Antisocial	0	2 (2)	2 (1)

a. All chi-squared analyses are with continuity correction.
b. After randomisation. Patients were asked how useful they thought treatment they had been offered would be on a Likert scale from 1 'not at all' to 12 'extremely useful'.
c. $\chi^2 = 3.84$, $P = 0.050$ ($n = 167$) and $\chi^2 = 5.66$, $P = 0.017$ ($n = 148$).
d. $\chi^2 = 2.74$, $P = 0.098$ ($n = 167$) and $\chi^2 = 4.50$, $P = 0.034$ ($n = 148$).

and obsessive-compulsive personality disorders ($\chi^2 = 4.50$, d.f. = 1, $P = 0.034$) in the long-term group. All other differences were statistically non-significant. The number of comorbid Axis I diagnoses was 3.3 on average. Patients reported that the problems they wanted therapy for had lasted on average 15 years. A total of 68% reported previous psychiatric treatment.

Therapist and group effects

Using GSI as the outcome variable, intercept and slope variances for therapists and groups indicated negligible non-independence in the data, all ICCs < 0.01. Using IIP-C as the outcome variable, intercept and slope variances for therapists and groups indicated negligible non-independence in the data, all ICCs < 0.02. Using GAF-S as the outcome variable, the random intercept variance for therapists indicated non-independence in the data, ICC = 0.11. Therapist intercept was therefore included as a random effect in the statistical model. The non-independence in random slopes within therapists was negligible, as was the random intercept and slopes within groups, all ICCs < 0.01. Using GAF-F as the outcome variable, the random intercept variance for therapists indicated some non-independence in the data, ICC = 0.05, and therapist intercept was included as random effect in the statistical model.

Treatment fidelity

Thirty-nine audio-recordings from the first 6 months of therapy (sessions 3, 10 and 17), from both the short- and long-term

groups, were drawn from the pool of 54 recordings. Two evaluators, who were masked to group, independently rated sessions on therapist-activity level, degree of focus, group work in the here-and-now and therapist competence, using Likert scales from 0 (not at all) to 4 (very much). Intraclass correlation on the therapy-process scales ranged from 0.70 to 0.94. There was significantly more work on a circumscribed problem focus ($t = 2.3$, d.f. = 15, $P = 0.036$) and work in the here-and-now in short-term therapy ($t = 2.1$, d.f. = 35, $P = 0.042$), as suggested in the treatment manuals. Level of therapist activity was equal in the two groups. Therapist mean competence was similar in both formats, ranging from moderate to high: 2.7 (s.d. = 0.5), range 1.7–3.4, in the short-term group and 2.3 (s.d. = 0.5), range 1.9–3.1, in the long-term group. We have previously reported that there were no significant differences (sessions 3, 10, 17) in therapeutic alliance or group cohesion in the two treatments.³⁵

Primary analysis of outcome variables

In the whole sample of patients ($n = 167$), descriptive statistics over time for the four outcome variables are presented in Table 2 (intention-to-treat analyses). We could detect no significant differences between the short- and long-term therapy groups at any time point, except that short-term therapy was superior to long-term therapy at 6 months, using the SCL-90-R as the outcome variable (see online supplement DS1). Figures 2–5 show the descriptive trajectories over time for the short- and long-term groups for the four outcome variables.

Table 2 Primary outcome measures over time in randomised patients in short- and long-term psychodynamic group psychotherapy^a

Outcome variables and time points, months	Short-term group		Long-term group		Whole sample	
	<i>n</i>	mean (s.d.)	<i>n</i>	mean (s.d.)	<i>n</i>	95% CI for difference short-term – long-term group
Symptoms (Symptom Checklist 90 – Revised)	76		88		164	
0		0.86 (0.50)		1.02 (0.64)		–0.33 to 0.03
6 ^b		0.77 (0.56)		0.96 (0.63)		–0.37 to –0.004
12		0.80 (0.56)		0.86 (0.62)		–0.24 to 0.13
18		0.86 (0.51)		0.85 (0.67)		–0.17 to 0.20
24		0.79 (0.51)		0.79 (0.65)		–0.18 to 0.18
36		0.70 (0.49)		0.71 (0.63)		–0.19 to 0.17
Interpersonal problems	76		88		164	
0		1.33 (0.56)		1.34 (0.51)		–0.18 to 0.15
6		1.18 (0.55)		1.32 (0.51)		–0.31 to 0.02
12		1.15 (0.56)		1.26 (0.55)		–0.28 to 0.07
18		1.22 (0.55)		1.22 (0.56)		–0.17 to 0.17
24		1.12 (0.60)		1.14 (0.54)		–0.20 to 0.15
36		0.98 (0.59)		1.04 (0.64)		–0.24 to 0.14
Global Assessment of Functioning – Symptoms	77		90		167	
0		58.2 (7.8)		58.1 (7.2)		–2.18 to 2.42
36		65.5 (11.0)		67.7 (13.3)		–6.0 to 1.6
Global Assessment of Functioning – Function	77		90		167	
0		60.3 (9.4)		59.3 (9.1)		–1.88 to 3.77
36		67.8 (11.7)		68.1 (14.2)		–4.39 to 3.66

a. Imputation method: last observation carried forward (LOCF).
 b. $t = -2.02$, d.f. = 162, $P = 0.045$ (independent sample t -test, two-tailed).

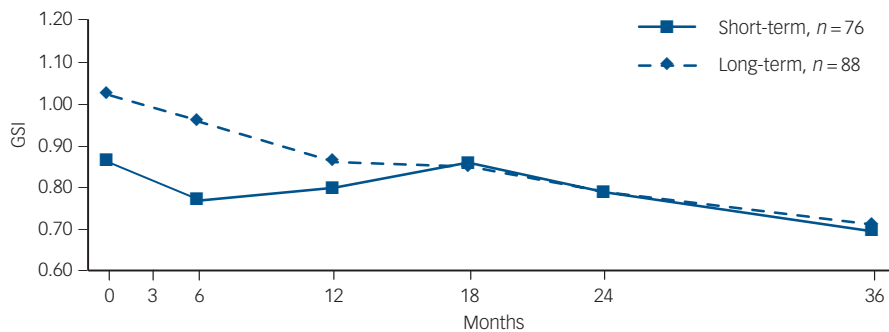


Fig. 2 Global Severity Index (GSI) score for patients in short- and long-term group therapy ($n = 164$).
 Imputed: last value carried forward. Measurements available for 164 of 167 patients.

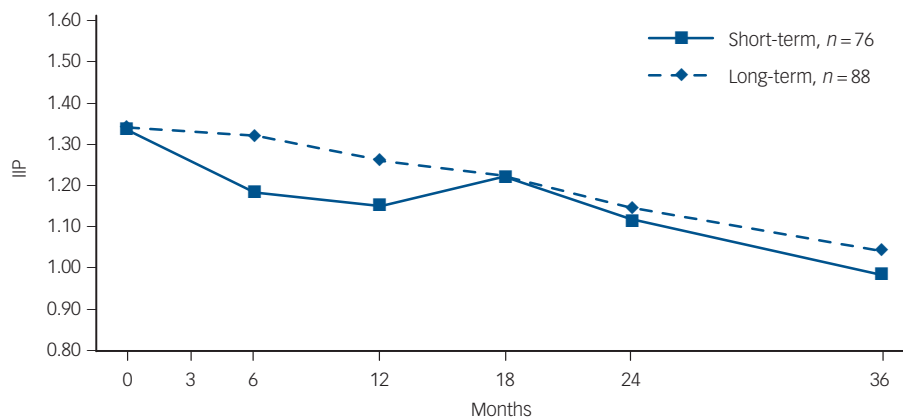


Fig. 3 Inventory of Interpersonal Problems (IIP) score for patients in short- and long-term group therapy ($n = 164$).
 Imputed: last value carried forward. Measurements available for 164 of 167 patients.

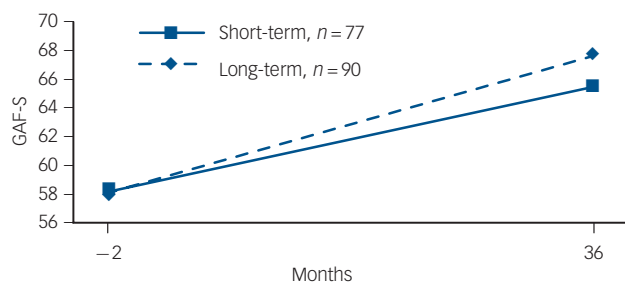


Fig. 4 Global Assessment of Functioning – Symptom (GAF-S) score for patients in short- and long-term group therapy ($n = 167$).

Imputed: last value carried forward.

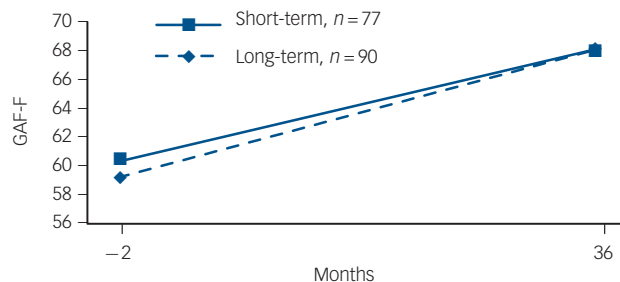


Fig. 5 Global Assessment of Functioning – Function (GAF-F) score for patients in short- and long-term group therapy ($n = 167$).

Imputed: last value carried forward.

The patients in both treatment arms made significant gains. The within-group effect sizes (Cohen's d)³⁶ for the short- and long-term groups respectively were as follows: for GSI, 0.3 and 0.5; for IIP, 0.6 and 0.6; for GAF-S, 0.9 and 1.3; for GAF-F, 0.8 and 1.0. The average effect size across treatments and outcome measures was 0.8 (large).

Clinically significant change (recovery)^{33,34} as measured by the GSI was achieved by 33.3% of the patients in the short-term group and 36.6% in the long-term group. For IIP the corresponding numbers were 37.5% *v.* 34.8%. Using GAF-S, recovery was achieved by 20.3% in the short-term group *v.* 39.7% in the long-term group ($\chi^2 = 3.49$, *d.f.* = 1, $P = 0.06$). The corresponding figures for GAF-F were 39.0% *v.* 39.4%. This means that 35% of the patients are recovered, across measures and treatments.

The multilevel analyses of longitudinal data on the GSI, IIP-C, GAF-S and GAF-F were performed on the whole sample of patients (intention-to-treat analyses). Using the GSI as the outcome variable, there was a significant difference in change over time (slopes) between the short- and long-term group, in favour of long-term therapy. Time \times group length was $B = -0.004$, $F = 4.20$, *d.f.* = 1,154, $P = 0.042$, 95% CI -0.0036 to -0.0002 , and the effect size of this term was 0.3 (small).

Using the IIP-C as the outcome variable, the difference in slopes between the short- and long-term group was not significant ($B = -0.0007$, $F = 0.09$, *d.f.* = 1,159, $P = 0.76$, 95% CI -0.005 to 0.004). With GAF-S as the outcome variable, the difference in slopes between the short- and long-term group was not significant, although we observed a trend in favour of long-term therapy ($B = 0.09$, $F = 2.76$, *d.f.* = 1,149, $P = 0.10$, 95% CI -0.02 to 0.19), and the effect size of this term was 0.3 (small). Using the GAF-F as the outcome variable the difference in slopes between the short- and long-term group was not significant $B = 0.06$, $F = 1.23$, *d.f.* = 1,149, $P = 0.27$, 95% CI -0.04 to 0.15).

When we included starter status (1, 0) and/or premature termination status in the statistical models,³² these terms were not significant and the pattern of results did not change. The analyses of the sample of patients that started therapy ($n = 148$) and the analyses of completers ($n = 116$) gave the same pattern of findings. Burden and adverse effects for patients were small. There were no suicide attempts or suicides carried out. The number of patients who deteriorated was relatively small (5.8%; see online supplement DS1).

The patients were interviewed about additional treatment and sick leave during the 3 years. We were unable to detect any statistically significant differences between the two groups regarding use of health services. Use of additional treatment was probably not a confounding factor in this study (see online supplement DS1). Positive and negative life events were recorded annually by the patients (before therapy start, and at 1, 2 and 3

years) and there were no significant differences between the two groups.

Discussion

Main findings

To the best of our knowledge, this is the first randomised clinical trial comparing the efficacy of short- and long-term psychodynamic group psychotherapy. Our hypotheses in this study were that change in symptoms throughout the 3-year study period would be similar for both short- and long-term therapy, although patients would improve more on interpersonal problems and psychosocial functioning with long-term compared with short-term therapy, throughout the 3-year study period. Our hypothesis about similar change in symptomatic distress (GSI) was not supported, as we found a significant treatment effect in favour of long-term therapy. The effect size was 0.3 (small). However, there was no difference in end-point values on symptom distress after 3 years between the two therapies. The results are difficult to interpret because of the initial between-group differences, and the increased rate of change in the long-term group could reflect regression to the mean. Contrary to our hypothesis, we observed that short- and long-term therapy were equally effective across 3 years, using IIP, GAF-S and GAF-F as the outcome variables. However, there was a trend in favour of long-term therapy ($P = 0.10$) using GAF-S as the outcome variable. The effect size was small (0.3), but this result was also supported by the fact that more patients were recovered in the long-term ($P = 0.06$) than the short-term group. The number needed to treat was five in favour of long-term therapy, which means that, on average, five patients have to be treated for one to benefit compared with a control (contrast) in a clinical trial.

The average within-group effect size in the study was 0.8, comparable to, or larger, than what has been reported in major meta-analyses of psychotherapy studies.^{37,38} The proportion of recovered individuals (clinically significant change) across treatments and measures was 35%, which is also comparable to those reported in the psychotherapy literature.^{1,2,8} We could not detect any significant difference in effectiveness between the nine therapists. This may be because the therapists were very experienced, they had been specifically trained before the groups started and they were using treatment manuals.³⁹

The patients in the short-term group did not seek more additional treatment than those in the long-term group during the 3-year study period. At the end of therapy all patients had been recommended not to seek additional treatment during the following year, unless treatment was urgently needed, since delayed effects after treatment often might be experienced.

It is a common experience that painful feelings connected to termination and separation (from the therapist and other group members) may be an impetus to prematurely seek additional therapy. Our impression is that our recommendation to abstain from additional treatment did not lead to more problems and suffering for the patients. This is supported by the fact that the patients in the short-term group, by and large, seem to have been helped just as much as those in the long-term group over the 3-year study period.

Strengths and limitations

There are several strengths to this study. First, the treatment took place in ordinary treatment settings, out-patient departments and private practices, with regular patients and therapists, which allows us to generalise our findings to these clinical situations. Other strengths are that there was a high inclusion rate from referral and that self-reports were sent regularly to all randomised patients, including at termination of the group they were allocated to. This, and the fact that about 90% of the patients participated in a 3-year follow-up, has given a near complete data-set, allowing for reliable intention-to-treat analyses.

Despite these strengths, there are several limitations that should be noted. Although randomisation is the 'gold standard' from a methodological point of view, it may in some respects be a drawback from a clinical point of view. Therapists were not allowed to select their own patients and 'compose' their groups in the way they usually would, which may restrict the generalisability of our findings to contemporary practice to some extent.⁴⁰ On the other hand, therapists were not allowed to exclude patients included in the study, which may in fact increase generalisability, and possibly reduce the magnitude of treatment effects, thereby not overestimating them. The patient sample had a wide variety of diagnoses, which may make it difficult to evaluate the effects of group therapy duration on a specific psychiatric disorder, such as depression. However, the wide variety of diagnoses in this sample may actually increase generalisability to patients seeking out-patient group psychotherapy. It may also be noted as a limitation that the study protocol was registered in clinicalTrials.gov after the randomisation had started. Another limitation may be the fact that self-report questionnaires, while patients were in therapy, were collected by the therapists. Some patients may have felt inhibited in answering openly. The purpose of this procedure, however, was primarily to increase the response rate and not to inform therapists, who were instructed to send the forms directly to the project leader.

This study indicates that short- and long-term group therapy are similar in effectiveness across 3 years for 'the average patient', and thus do not offer guidelines for who should enrol in short- or long-term therapy respectively. We have so far, only presented the primary outcome analyses in this study, but the trial also encompasses data on possible moderators of treatment effects (degree of initial distress, personality pathology, quality of object relationships) and potential mechanisms behind change (mediators). Moderator analyses may offer additional information about differential efficacy of short- and long-term therapy, for subgroups of patients (see online supplement DS1).

Implications

Short- and long-term group psychotherapy were similarly effective for the typical patient in this study. Our hypotheses that long-term therapy would be superior in areas of interpersonal problems and psychosocial functioning, were not supported. Offering long-term therapy to most patients seems unnecessary and may be a waste of

resources, but further analyses of potential moderators of treatment in short- and long-term group therapy are needed.

Steinar Lorentzen, MD, PhD, Department for Research and Development, Clinic for Mental Health and Addiction, Oslo University Hospital, and Institute of Clinical Medicine, University of Oslo; **Torleif Ruud**, MD, PhD, Department for Research and Development, Division of Mental Health Services, Akershus University Hospital, Lørenskog, and Institute of Clinical Medicine, University of Oslo; **Anette Fjeldstad**, MD, Department of Research and Development, Clinic for Mental Health and Addiction, Oslo University Hospital; **Per Høglend**, MD, PhD, Institute of Clinical Medicine, University of Oslo, Oslo, Norway

Correspondence: Steinar Lorentzen, University of Oslo, Institute of Clinical Medicine, Clinic for Mental Health and Addiction, PO Box 1039 Blindern, 0315 Oslo, Norway. Email: steinar.lorentzen@medisin.uio.no

First received 20 Apr 2012, final revision 5 Apr 2013, accepted 8 May 2013

Funding

The study has been partly funded through a grant from the Health authority Helse Øst HRF during 2004–2007, and has received some economical support from the Oslo University Hospital-Aker and the University of Oslo.

Acknowledgements

We want to express gratitude to the following institutions and individuals for important contributions: out-patient clinics at community mental health centres in Ålesund, Alna (Oslo), in Sandnes and Ryfylke (Stavanger); and Laila Hjulstad, Knut Skjøstad, Vibeke Lohne, Allan Larsen, Kirsten Høbye, Helge Knudsen, Marthe Horneland, Dagny Sande Børnes, Ole Inge Gjøen, Wibeke Kløvning, Synnøve A. Kristiansen, Martin Mydske Nilssen and Ørjan Berg.

References

- Lambert MJ, Bergin AE. The effectiveness of psychotherapy. In *Handbook of Psychotherapy and Behavior Change* (eds SL Garfield, AE Bergin): 143–89. Wiley, 1994.
- Burlingame GM, Strauss B, Joyce AS. Change mechanisms and effectiveness of small group treatments. In *Bergin and Garfield's Handbook of Psychotherapy and Behavior Change* (ed. MJ Lambert): 640–89. Wiley, 2013.
- Lorentzen S, Ruud T. Group therapy in public mental health services: approaches, patients and group therapists. *J Psychiatr Ment Health Nurs* 2013; April 15 (Epub ahead of print).
- Rüger U, Bell K. Historische entwicklung und aktueller stand der richtlinien-psychotherapie in Deutschland [The history of psychotherapy in the German system of compulsory national health insurance]. *Z Psychosom Med Psychother* 2004; **50**: 127–52.
- Lorentzen S, Bøgwald, KP, Høglend P. Change during and after long-term analytic group psychotherapy. *Int J Group Psychother* 2002; **32**: 419–29.
- Tschuschke V, Anbeh T, Kiencke P. Evaluation of long-term analytic outpatient group therapies. *Group Anal* 2007; **40**: 140–59.
- Bateman A, Fonagy P. Randomized controlled trial of outpatient mentalization-based treatment versus structured clinical management for borderline personality disorders. *Am J Psychiatry* 2009; **166**: 1355–64.
- Orlinsky DE, Grawe K, Parks BK. Process and outcome in psychotherapy: noch einmal. In *Handbook of Psychotherapy and Behavior Change* (eds SL Garfield, AE Bergin): 270–376. John Wiley, 1994.
- Orlinsky DE, Rønnestad MH, Willutzki U. Fifty years of psychotherapy process-outcome research: continuity and change. In *Bergin and Garfield's Handbook of Psychotherapy and Behavior Change* (ed. MJ Lambert): 307–89. John Wiley, 2004.
- Reich JH, Green AI. Effect of personality disorders on outcome of treatment. *J Nerv Ment Dis* 1991; **179**: 74–82.
- Perry JC, Banon E, Ianni F. Effectiveness of psychotherapy for personality disorders. *Am J Psychiatry* 1999; **156**: 1312–21.
- Gabbard GO. Psychotherapy of personality disorders. *J Psychother Pract Res* 2000; **9**: 1–6.
- Piper WE. Client variables. In *Handbook of Group Psychotherapy: An Empirical and Clinical Synthesis* (eds A Fuhrman, GM Burlingame): 83–113. John Wiley, 1994.
- Lorentzen S, Høglend P. Moderators of the effects of treatment length in long-term psychodynamic group psychotherapy. *Psychother Psychosom* 2008; **77**: 321–2.

- 15 Stiles WB, Shapiro DA. Disabuse of the drug metaphor: psychotherapy process-outcome correlations. *J Consult Clin Psychol* 1994; **62**: 942–5.
- 16 Holt-Lundstad J, Smith TB, Layton JB. Social relationships and mortality risk: a meta-analytic review. *PLoS Med* 2010; **7**: e1000316.
- 17 Schulz KF, Grimes DA. Blinding in randomised trials: hiding who got what. *Lancet* 2002; **359**: 696–700.
- 18 Lorentzen S. *Group Analytic Psychotherapy: Working with Affective, Anxiety and Personality Disorders*. Routledge, 2013.
- 19 Foulkes SH, Anthony EJ. *Group Psychotherapy: The Psychoanalytical Approach*. Maresfield reprints, 1984.
- 20 Sheehan D, Janavs J, Baker R, Harnett-Sheehan K, Knapp E, Sheehan M. *Mini International Neuropsychiatric Interview*. Tampa University of South Florida Press, 2002.
- 21 First MB, Gibbon M, Spitzer RL, Williams JBW, Benjamin LS. *Structured Clinical Interview for DSM-IV Axis I Personality Disorders (SCID-I)*. American Psychiatric Press, 1997.
- 22 Endicott J, Spitzer RL, Fleiss JL, Cohen J. The Global Assessment Scale. A procedure for measuring overall severity of psychiatric disturbance. *Arch Gen Psychiatry* 1979; **33**: 766–71.
- 23 Derogatis LR. *The SCL-90-R: Administration, Scoring and Procedures. Manual I*. Clinical Psychometric Research Unit, Johns Hopkins University School of Medicine, 1977.
- 24 Alden LE, Wiggins JS, Pincus AL. Construction of circumplex scales from the Inventory of Interpersonal Problems. *J Pers Assess* 1990; **55**: 521–36.
- 25 Horowitz LM, Rosenberg SE, Baer BA, Ureno G, Villasenor VS. Inventory of Interpersonal Problems: psychometric properties and clinical applications. *J Consult Clin Psychol* 1988; **56**: 885–92.
- 26 Goldman HH, Skodol AE, Lave TR. Revising axis V for DSM-IV: a review of measures of social functioning. *Am J Psychiatry* 1992; **149**: 1148–56.
- 27 Shrout PE, Fleiss JL. Intraclass correlations: uses in assessing rater reliability. *Psychol Bull* 1991; **86**: 420–8.
- 28 Battle C, Imber S, Hoehn-Saric R, Stone AR, Nash EH, Frank JD. Target Complaints as criteria of improvement. *Am J Psychother* 1966; **20**: 184–92.
- 29 Altman DG. *Practical Statistics for Medical Research*: 456. Chapman & Hall, 1991.
- 30 Verbeke G, Molenberg G. *Linear Mixed Models for Longitudinal Data*. Springer, 2000.
- 31 Fitzmaurice G, Laird NM, Ware JH. *Applied Longitudinal Analysis*. Wiley, 2004.
- 32 Gibbons RD, Hedeker D. Random effects probit and logistic regression models for three-level data. *Biometrics* 1997; **53**: 1527–37.
- 33 Jacobson NS, Truax P. Clinical significance: a statistical approach to defining meaningful change in psychotherapy research. *J Consult Clin Psychol* 1991; **59**: 12–9.
- 34 Christensen L, Medoza JL. A method of assessing change in a single subject: an alteration of the RC index. *Behav Ther* 1986; **17**: 305–8.
- 35 Bakali JV, Wilberg T, Hagtvet KA, Lorentzen S. Sources accounting for alliance and cohesion at three stages in group psychotherapy: variance component analysis. *Group Dyn* 2010; **14**: 368–83.
- 36 Cohen J. *Statistical Power Analysis for the Behavioral Sciences*. Academic Press, 1988.
- 37 Crits-Cristoph P. The efficacy of brief dynamic psychotherapy: a meta-analysis. *Am J Psychiatry* 1992; **149**: 151–8.
- 38 Lipsey MW, Wilson DB. The efficacy of psychological, educational, and behavioral treatment: confirmation from meta-analysis. *Am Psychol* 1993; **48**: 1181–209.
- 39 Crits-Cristoph P, Mintz J. Implications of therapist effects for the design and analysis of comparative studies of psychotherapies. *J Consult Clin Psychol* 1991; **59**: 20–6.
- 40 Horneland M, Sande DB, Høbye K, Knutsen H, Lorentzen S. Can the clinician-researcher gap be bridged? Experiences from a randomized clinical trial in analytic/dynamic group psychotherapy. *Group Anal* 2012; **25**: 84–92.

